

Essays in Asset Pricing Theory

Nicolo Ceneda

Imperial College London

Department of Finance

Submitted in partial fulfilment of the requirements for the degree of Doctor of
Philosophy at Imperial College London.

April 2026

Statement of Originality

I herewith certify that this thesis constitutes my own work and that all material, which is not my own work, has been properly acknowledged. I confirm that Chapter 2 was jointly written with Harjoat Bhamra. I confirm that Chapter 3 was jointly written with Walter Distaso and Valentina Corradi.

Nicolo Ceneda

Copyright Declaration

The copyright of this thesis rests with the author. Unless otherwise indicated, its contents are licensed under a Creative Commons Attribution-Non Commercial 4.0 International Licence (CC BY-NC). Under this licence, you may copy and redistribute the material in any medium or format. You may also create and distribute modified versions of the work. This is on the condition that: you credit the author and do not use it, or any derivative works, for a commercial purpose. When reusing or sharing this work, ensure you make the licence terms clear to others by naming the licence and linking to the licence text. Where a work has been adapted, you should indicate that the work has been changed and describe those changes. Please seek permission from the copyright holder for uses of this work that are not included in this licence or permitted under UK Copyright Law.

Acknowledgements

I am forever grateful to my advisors, Harjoat Bhamra, Ziang Li, and Franklin Allen, whose expertise and guidance have allowed me to grow as a researcher and to push my intellectual limits beyond what I thought possible. A special thank you goes to Harjoat for his relentless dedication to my development and for teaching me so much (especially mathematics).

Thank you to all my friends for their patience and support throughout these years of sacrifice.

I dedicate this thesis to my mother, Mirella; to the memory of my father and brother, Valerio and Jacopo; and to my girlfriend, Nora, for standing by my side throughout this challenging journey.

Abstract

This dissertation consists of three essays in asset pricing theory. The first two essays study unconventional monetary policy, focusing respectively on monetary policy disagreement and market failure. The third essay studies portfolio sorting and the econometric distortions that arise when portfolios are formed using observable or estimated sorting variables.

Chapter 1 develops a preferred-habitat model of the yield curve that links conventional and unconventional monetary policy tools to disagreement within the monetary policy committee. It studies the transmission of policy-rate changes, forward guidance, and quantitative easing along the curve, and uses FOMC dot-plot data to discipline disagreement about future monetary policy. The main contribution is to show that greater monetary policy disagreement weakens forward guidance but strengthens quantitative easing, especially at longer maturities, implying that when disagreement is high the central bank should lean more toward quantitative easing, particularly when aiming to influence the long end of the curve.

Chapter 2 develops a production-economy model and shows that a financial crisis that depletes intermediary capital can destroy competitive equilibrium. Quantitative easing becomes a necessary condition for market clearing, and the central bank acts as a risk-bearer of last resort. However, quantitative easing can become permanent or persist for centuries in a QE Trap, with fiscal recapitalisation as the only practical escape. In the unconstrained economy, a shortage of risk-bearing capacity instead inflates asset prices above fundamental values, so this class of economies cannot deliver both correct asset prices and market clearing in every state.

Chapter 3 studies how portfolio-sorting tests are affected by the gap between sample order statistics and true quantiles, and, when the sorting variable is estimated (such as beta), by measurement error. This error can be decomposed into contamination errors and information loss errors. It then proposes an adaptive trimming rule under which tests for equality of returns between portfolios have correct asymptotic size and unit power.

Contents

Chapter 1: Disagreement in the Term Structure of Interest Rates	18
1.1 Introduction	18
1.1.1 Literature Review	23
1.2 Model	25
1.2.1 Bond Market	25
1.2.2 Agents	27
1.2.3 Equilibrium Without Arbitrageurs	29
1.2.4 Equilibrium With Arbitrageurs	29
1.3 Deterministic Target Short-Rate	33
1.3.1 Interest Rate Policy	33
1.3.2 Risk Premia	37
1.4 Stochastic Target Short-Rate	40
1.4.1 Interest Rate Policy	43
1.4.2 Risk Premia	45
1.5 Calibration	46
1.5.1 Target Short Rate and Monetary Policy Disagreement	46
1.5.2 Current Short Rate and Bond Supply	49
1.6 Policy Analysis	53
1.6.1 Unconventional Monetary Policy	53
1.6.2 Effect of Monetary Policy Disagreement	55
1.7 Conclusion	58
Chapter 2: Market Failure, the QE Trap and Fiscal Escape	60
2.1 Introduction	60
2.1.1 Related Literature	63

2.2	Model	64
2.2.1	Environment	64
2.2.2	Leverage Constraint	66
2.2.3	Benchmarks and Baseline Dynamics	67
2.3	Capital Depletion Causes Equilibrium Failure	67
2.3.1	Dynamics of ω and the Onset of Failure	71
2.3.2	Which Assumptions Are Essential?	72
2.4	Quantitative Easing, the QE Trap, and Asset Prices	73
2.4.1	Central Bank as Risk-Bearer of Last Resort	73
2.4.2	Equilibrium with QE	73
2.4.3	Dynamics with the Leverage Constraint	75
2.4.4	Two Steady States: Permanent QE and the QE Trap	75
2.4.5	Asset Prices and Risk Premia	79
2.5	The Intermediation Gap	80
2.5.1	A Three-Act Narrative	80
2.5.2	Martingale Criterion	81
2.5.3	Connection to the Rational Bubble Literature	81
2.5.4	The Exact Gap Formula in the Production Economy	81
2.5.5	The Intermediation Gap with and without Leverage Constraints	83
2.5.6	The Leverage Paradox	83
2.5.7	The Intermediation Gap as a Sufficient Statistic for the Output Gap	85
2.6	Real Effects and Policy	86
2.6.1	The Unified State Variable	86
2.6.2	Phantom Wealth	86
2.6.3	Large-Scale Asset Purchases	87
2.6.4	Fiscal Recapitalisation	87
2.6.5	Mapping to Policy	89
2.6.6	Empirical Support	89
2.7	Empirical Evidence	89
2.7.1	Data	90
2.7.2	Calibration	92
2.7.3	Test 1: Real Economic Activity	94

2.7.4	Test 2: Dealer Intermediation and Central Bank Substitution	98
2.7.5	Test 3: Pre-QE Intermediation Gap as a Sufficient Statistic	103
2.7.6	Test 4: Fiscal Recapitalization and Recovery	106
2.7.7	Test 5: QE Trap	110
2.8	Conclusion	111
Chapter 3: The Econometrics of Portfolio Sorts		113
3.1	Introduction	113
3.2	Set-Up	116
3.3	Sorting Without Measurement Error	117
3.3.1	Asymptotic Results Under Weak Correlation of the Sorting Variable	117
3.3.2	Asymptotic Results Under Strong Correlation of the Sorting Variable	119
3.4	Sorting With Measurement Error	120
3.4.1	Classification Errors	121
3.4.2	Trimming Rule	122
3.4.3	Asymptotic Results	124
3.5	Subsampling	128
3.6	A Motivating Example: Sorting on Stocks' Betas	130
3.6.1	Unconditional Betas	131
3.6.2	Conditional Betas	133
3.6.3	Test Statistic and Critical Values	136
3.7	Simulations	137
3.7.1	Data	137
3.7.2	Simulations for Classification Errors	138
3.7.3	Simulations for Hypothesis Testing	144
3.8	Conclusion	149
Bibliography		151
Appendix A: Disagreement in the Term Structure of Interest Rates		161
A.1	Proofs	161
A.1.1	Proposition 1.1	161
A.1.2	Proposition 1.2	162

A.1.3	Proposition 1.3	163
A.1.4	Proposition 1.4	164
A.1.5	Proposition 1.5	165
A.1.6	Lemma 1.1	167
A.1.7	Lemma 1.2	169
A.1.8	Theorem 1.1	170
A.1.9	Corollary 1.1	175
A.1.10	Proposition 1.6	178
A.1.11	Proposition 1.7	181
A.2	Overview of the Expectations Hypothesis	194
A.3	Calibration	195
A.4	Policy Analysis	198
A.5	Model Implementation	205
Appendix B: Market Failure, the QE Trap and Fiscal Escape		211
B.1	Complete-Markets Benchmark	211
B.2	Baseline Equilibrium: Limited Participation Without Leverage Constraints	211
B.2.1	State Variable and Equilibrium Characterisation	212
B.2.2	Dynamics of the Wealth Distribution	214
B.2.3	Two Invariant Regimes and Convergence	215
B.2.4	Asset Prices and Real Variables	215
B.3	Proofs for Limited Participation Without Leverage Constraint	217
B.4	Equilibrium with Leverage Constraints but Without QE	220
B.4.1	Equilibrium Above the Crisis Threshold	220
B.4.2	The Crisis Threshold	221
B.4.3	Market Failure Below the Crisis Threshold	221
B.4.4	Crisis Probability	221
B.4.5	Comparison of the Three Cases	222
B.5	Proof of Equilibrium Failure	223
B.6	Derivation of σ_ω With Leverage Constraint	224
B.7	Q-Martingale Property	226
B.8	First-Passage Analysis on the Pre-Failure Domain	228

B.8.1	Preliminaries	229
B.8.2	Hitting Probability	229
B.8.3	Unconditional Expected Hitting Time	231
B.8.4	Conditional Expected Hitting Time	232
B.8.5	Closed-Form Case: $\varphi = 0$	234
B.8.6	Sanity Check: Constant-Diffusion Brownian Limit	235
B.8.7	Tail Distribution of τ_c	236
B.8.8	Economic Interpretation	247
B.9	Boundary Classification	248
B.10	Convergence	250
B.11	Proof of Intermediation Gap Classification	251
B.12	Empirical Appendix	252
B.12.1	Constructing the Central Bank Risk-Bearing Measure	252
B.12.2	Structural Breaks Additional Evidence	253
Online Appendix B: Market Failure, the QE Trap and Fiscal Escape		256
BO.1	Proofs for the Complete-Markets Benchmark	256
BO.2	Auxiliary Results for the QE Dynamics	259
BO.2.1	Properties of the Volatility Function	260
BO.2.2	Three Invariant Regimes	260
BO.2.3	Boundary Classification	260
BO.3	Proof of Asset Pricing Formulae	261
BO.4	Equilibrium Quantities as Functions of the Wealth Share	262
BO.4.1	The Wealth-Share Volatility: The Fundamental Object	263
BO.4.2	Portfolio Share and Wealth-Share Volatility	264
BO.4.3	Asset Pricing Quantities	264
BO.4.4	Real Variables and the Intermediation Gap	265
BO.5	Intermediation Gap Formula Derivation	265
BO.6	Franchise Value and Dynamic Incentive Compatibility	272
BO.7	Employment Multiplier and Welfare Analysis	274
BO.7.1	Employment Multiplier of Recapitalisation	274
BO.7.2	Welfare Analysis	275

BO.8 Proofs of Policy Results	276
BO.9 Translation of Hugonnier's Results	276
BO.10 Partial Participation and CRRA Robustness	278
BO.10.1 Extended Model	278
BO.10.2 Robustness of Equilibrium Failure	279
BO.10.3 Dynamics and Preservation of Results	280
BO.10.4 Bubble Classification	280
BO.10.5 CRRA / Power Utility	281
BO.11 Endogenous Non-Participation via Ambiguity	282
BO.11.1 Household Preferences under Ambiguity	282
BO.11.2 Robust Equilibrium Failure	284
BO.11.3 Endogenous Non-Participation	284
BO.11.4 Permanent QE, the QE Trap, and Bubble Classification	285
BO.11.5 Three-Region Equilibrium	285
BO.12 Robustness under Epstein–Zin Preferences	285
BO.12.1 Equilibrium Failure Is Preference-Independent	286
BO.12.2 Permanent QE and the QE Trap Under EZ	286
BO.13 Leverage vs. Equity Constraints	287
BO.13.1 Setup: the Equity Constraint	288
BO.13.2 Crisis Threshold and Central Bank Intervention	288
BO.13.3 Wealth Dynamics Under the Equity Constraint	291
BO.13.4 No Permanent QE Under the Equity Constraint	292
BO.13.5 Intermediation Gap and the Leverage Paradox	294
BO.13.6 General Characterisation: When Does Permanent QE Exist?	295
BO.13.7 Comparison Table	298
BO.13.8 Verification: Special and Limiting Cases	299
BO.13.9 Economic Interpretation and Policy Implications	301
BO.14 Drift-Time Analysis for Recovery Dynamics	301
BO.14.1 Drift-Time Lower Bound on Recovery	301
BO.14.2 Exact Antiderivative via Partial Fractions	302
BO.14.3 Leading-Order Approximation Near $\underline{\omega}$	304
BO.14.4 Error Bound	305

BO.14.5 Recovery Timescale Table	306
BO.15 Empirical Evidence Robustness Checks	307
BO.15.1 Test 1: Real Economic Activity	307
BO.15.2 Test 2: Dealer Intermediation and Central-Bank Substitution	319
BO.15.3 Test 3: Pre-QE Intermediation Gap as a Sufficient Statistic	326
Appendix C: The Econometrics of Portfolio Sorts	330
C.1 Proofs	330
C.1.1 Theorem 1	330
C.1.2 Theorem 2	332
C.1.3 Lemma 1	332
C.1.4 Theorem 3	333
C.1.5 Corollary 1	335
C.1.6 Theorem 4	339
C.1.7 Proposition B	341
C.1.8 Proposition B'	342
C.1.9 Theorem B1	343
C.2 Simulations for Classification Errors	343
C.2.1 True (Unobservable) Sorting Variable	343
C.2.2 Classification Errors	345

List of Tables

Chapter 1: Disagreement in the Term Structure of Interest Rates	18
1.1 Calibrated parameters of target short-rate	49
1.2 Calibrated parameters of current short-rate and bond supply	51
1.3 Elasticities of model-generated moments	53
Chapter 2: Market Failure, the QE Trap and Fiscal Escape	60
2.1 Calibrated parameters	93
2.2 Test 1A – theory-implied level fit, output gap	95
2.3 Test 1B – predictive regressions, output gap	97
2.4 Test 1B – predictive regressions, negative unemployment gap	97
2.5 Test 1B – predictive regressions, cumulative payroll growth	98
2.6 Test 1B – predictive regressions, cumulative industrial production growth	98
2.7 Test 2A – predictive regressions, treasury dealer financing	100
2.8 Test 2A – predictive regressions, agency MBS dealer financing	100
2.9 Test 2B – central bank substitution, treasury dealer gross positions	102
2.10 Test 2B – central bank substitution, agency MBS dealer gross positions	102
2.11 Test 3A – theory-implied level fit, exact intermediation gap	104
2.12 Test 3B – pooled regressions, exact intermediation gap	106
2.13 Test 4A – predictive regressions, intermediary capital recovery	107
2.14 Test 4B – predictive regressions, output gap	108
2.15 Test 4B – predictive regressions, negative unemployment gap	109
2.16 Test 4B – predictive regressions, cumulative GDP growth	109
2.17 Test 5 – regime summary, HKM intermediary capital ratio	110

Chapter 3: The Econometrics of Portfolio Sorts	113
3.1 Summary statistics – 10 portfolios, 12 months	139
3.2 Classification errors – 10 portfolios, 12 months, without trimming	143
3.3 Classification errors – 10 portfolios, 12 months, with trimming	143
3.4 Size and power of the (un)trimmed test statistics	149
Appendix B: Market Failure, the QE Trap and Fiscal Escape	211
B.1 Equilibrium properties under three model configurations	222
B.2 Tail probabilities at the empirical calibration	246
B.3 Bai-Perron break-count selection for the mean of ω_t	254
B.4 Modern-sample robustness for the mean of ω_t	255
Online Appendix B: Market Failure, the QE Trap and Fiscal Escape	256
BO.1 Numerical solution of exact gap ODE	271
BO.2 Failure-region comparison	289
BO.3 Leverage constraint vs. equity constraint	299
BO.4 Exact drift time	306
BO.5 Approximate drift time	307
BO.6 Test 1A – sample robustness, output gap	308
BO.7 Test 1A – sample robustness, negative unemployment gap	308
BO.8 Test 1B – sample robustness, output gap	310
BO.9 Test 1B – sample robustness, negative unemployment gap	311
BO.10 Test 1B – sample robustness, cumulative payroll growth	312
BO.11 Test 1B – sample robustness, cumulative industrial production growth	313
BO.12 Test 1B – control robustness, output gap	314
BO.13 Test 1B – control robustness, negative unemployment gap	315
BO.14 Test 1B – control robustness, cumulative payroll growth	316
BO.15 Test 1B – control robustness, cumulative industrial production growth	317
BO.16 Test 1B – bootstrap inference, output gap	318
BO.17 Test 1B – bootstrap inference, negative unemployment gap	318
BO.18 Test 1B – bootstrap inference, cumulative payroll growth	319
BO.19 Test 1B – bootstrap inference, cumulative industrial production growth	319
BO.20 Test 2A – sample robustness, treasury dealer financing	320

BO.21 Test 2A – sample robustness, agency MBS dealer financing	320
BO.22 Test 2A – control robustness, treasury dealer financing	321
BO.23 Test 2A – control robustness, agency MBS dealer financing	322
BO.24 Test 2A – bootstrap inference, treasury dealer financing	323
BO.25 Test 2A – bootstrap inference, agency MBS dealer financing	323
BO.26 Test 2B – sample robustness, treasury and agency MBS dealer gross position . .	324
BO.27 Test 2B – control robustness, treasury and agency MBS dealer gross position . .	324
BO.28 Test 2B – bootstrap inference, treasury and agency MBS dealer gross position .	325
BO.29 Test 2B – regressor robustness, treasury and agency MBS dealer financing . . .	326
BO.30 Test 3A – sample robustness, output gap	326
BO.31 Test 3A – sample robustness, negative unemployment gap	327
BO.32 Test 3B – sample robustness, exact intermediation gap	328
BO.33 Test 3B – bootstrap inference, exact intermediation gap	329
BO.34 Test 3B – horse race, cyclically adjusted price-earnings ratio	329

Appendix C: The Econometrics of Portfolio Sorts 330

C.1 Summary statistics – 10 portfolios, 12 months	344
C.2 Summary statistics – 10 portfolios, 24 months	344
C.3 Summary statistics – 10 portfolios, 60 months	344
C.4 Summary statistics – 5 portfolios, 12 months	344
C.5 Summary statistics – 5 portfolios, 24 months	345
C.6 Summary statistics – 5 portfolios, 60 months	345
C.7 Classification errors – 10 portfolios, 12 months, without trimming	345
C.8 Classification errors – 10 portfolios, 12 months, with trimming	346
C.9 Classification errors – 10 portfolios, 24 months, without trimming	346
C.10 Classification errors – 10 portfolios, 24 months, with trimming	346
C.11 Classification errors – 10 portfolios, 60 months, without trimming	346
C.12 Classification errors – 10 portfolios, 60 months, with trimming	347
C.13 Classification errors – 5 portfolios, 12 months, without trimming	347
C.14 Classification errors – 5 portfolios, 12 months, with trimming	347
C.15 Classification errors – 5 portfolios, 24 months, without trimming	347
C.16 Classification errors – 5 portfolios, 24 months, with trimming	348

C.17	Classification errors – 5 portfolios, 60 months, without trimming	348
C.18	Classification errors – 5 portfolios, 60 months, with trimming	348

List of Figures

Chapter 1: Disagreement in the Term Structure of Interest Rates	18
1.1 Time series of dot plot projections	20
1.2 Deterministic target short-rate – carry trades	34
1.3 Deterministic target short-rate – underreaction of yields	36
1.4 Deterministic target short-rate – impact on yields	37
1.5 Deterministic target short-rate – effect on risk premia	38
1.6 Example of dot plot	47
1.7 Moments of 1-year-ahead dot-plot	48
1.8 Model-generated and empirical moments	52
1.9 Forward-guidance announcement	54
1.10 Quantitative easing announcement	55
1.11 Impact of monetary policy disagreement	56
1.12 Impact of monetary policy disagreement based on policy specifications	57
1.13 Impact of monetary policy disagreement at different maturities	58
Chapter 2: Market Failure, the QE Trap and Fiscal Escape	60
2.1 Equilibrium-failure mechanism	69
2.2 Equilibrium dynamics	77
2.3 Five regions of ω	78
2.4 Asset prices and real variables	79
2.5 Exact intermediation gap-to-price ratio	82
2.6 HKM intermediary capital ratio	93
Chapter 3: The Econometrics of Portfolio Sorts	113
3.1 Contamination error and information loss error	122

3.2	Trimming procedure	124
3.3	Number of common stocks in CRSP U.S. stock database	138
3.4	Average size of (un)trimmed portfolios and trimmed quantity	142
3.5	Contamination error of (un)trimmed portfolios	144
3.6	Empirical distribution of market betas	145
Online Appendix B: Market Failure, the QE Trap and Fiscal Escape		256
BO.1	Wealth-share volatility	263
BO.2	Portfolio share and volatility	264
BO.3	Asset pricing quantities	264
BO.4	Real variables and intermediation gap	265

Chapter 1

Disagreement in the Term Structure of Interest Rates

Nicolo Ceneda

1.1 Introduction

Since late 2008, when short-term interest rates reached the zero lower bound, central banks have expanded their traditional monetary policy toolkit – previously centered around adjustments to the monetary policy rate – to include large-scale asset purchases (buying and selling of long-term government bonds and other long-term securities) and forward guidance (shaping of market expectations about the future trajectory of short-term interest rates). The former acts through a risk premium channel: by altering the maturity structure of outstanding government bonds, quantitative easing changes the amount of duration risk that market participants must bear, thus impacting bond risk premia and long-term interest rates. Instead, the latter works through an expectations channel: by conveying information about the central bank's willingness to keep short-rates at a given level in the future, forward guidance impacts expectations and thus interest rates.

At the same time, the way central banks have tried to communicate their intentions and influence the markets has kept evolving, and forward guidance has become ever more important moving beyond being just a crisis tool. In December 2008, the Federal Open Market Committee (FOMC) introduced qualitative forward guidance, signalling that weak conditions would warrant exceptionally low rates "for some time". In August 2011, it shifted to calendar-based guidance, warranting low levels for the federal funds rate "at least through mid-2013". Finally, in December 2012 it adopted state-contingent guidance, indicating liftoff would wait as long as

the "unemployment rate remains above 6.5%" and "inflation [...] is projected to be no more than 0.5pp above the longer-run goal". To further strengthen its influence on the markets, in January 2012, the Federal Reserve started releasing FOMC participants' projections for the appropriate federal funds rate at several horizons. These projections show where each participant thinks the policy rate should be at the end of the current calendar year, at the end of each of the following two calendar years, and in the longer-run. The resulting chart plotting these individual projections is informally known as the "dot plot". Despite being closely monitored by market participants, it remains unclear how the information conveyed by these forward interest-rate publications affects unconventional monetary policy tools, and how large those effects are.

These plots become especially important when financial markets are uncertain about the path of monetary policy. Yet, monetary policy disagreement, as conveyed by the dispersion of these projections, may cause even higher uncertainty. The left axis of Figure 1.1 plots the mean, minimum and maximum values of the FOMC's dot-plot projections at each FOMC meeting and for different projection horizons, while the right axis plots their standard deviation. The figure shows that dispersion is not simply a post-pandemic phenomenon: for the one- and two-year-ahead projections, it was high in the early dot-plot sample, narrowed gradually in the years leading up to Covid-19, fell sharply during the pandemic, and then rose again as the Federal Reserve moved away from the zero lower bound. As an illustration of why dispersion in the dots matters, consider the 16-17 September 2025 meeting: the FOMC cut the target range by 25bp to 4.00-4.25%, yet the end-2025 projections spanned 2.88%-4.38% (a 150bp range, with a mean of about 3.81%) with lone dots at 4.38% (essentially the pre-meeting midpoint level) and at 2.88% (75bp below the next-lowest projection). This episode led to loud and negative market feedback about how this heterogeneity blurs the information content of forward guidance, significantly diminishing its power¹. This begs the question: what is the impact of monetary policy disagreement on the effectiveness of unconventional monetary policy tools?

To answer this question qualitatively, I develop a model of the yield curve that explicitly

¹On this occasion, Anna Wong, Chief US Economist at Bloomberg LP and former Principal Economist of the Federal Reserve Board said "I have not seen a meeting with so many contradictions. This is a low conviction Federal Reserve with limited visibility and a very wide range of views on the future. And this is why this market is so confused". Similarly, Lisa Abramowicz, co-host of Bloomberg Surveillance, stated "That sound you hear is the idea of forward guidance dying, because ultimately, how can a Federal Reserve with this level of dispersion in views really give any sense of their reaction function". Tom Keen, also co-host of Bloomberg Surveillance, stated "And you look at the dot plot and the dispersion looks like a B-52 bomber wingspan". Mike McKee, Executive Director at Goldman Sachs, stated "When you look at the dot plot itself, it is so narrowly divided that you can't take any signal from it". Jeffrey Rosenberg, Managing Director at BlackRock, stated "The bond market and the year end [...] are basically aligned. It's when you look further out into those projections [...] that there is a lot of disagreement and it gets to this notion of the spread [...] in the dots as you look at the dispersion there [...] That's part of what you are seeing in terms of the disconnect between the dots today and bond market expectations for the path of policy rates in 2026".

links quantitative easing and forward guidance with monetary policy disagreement. I then calibrate the model using dot-plot data and answer the same question quantitatively.

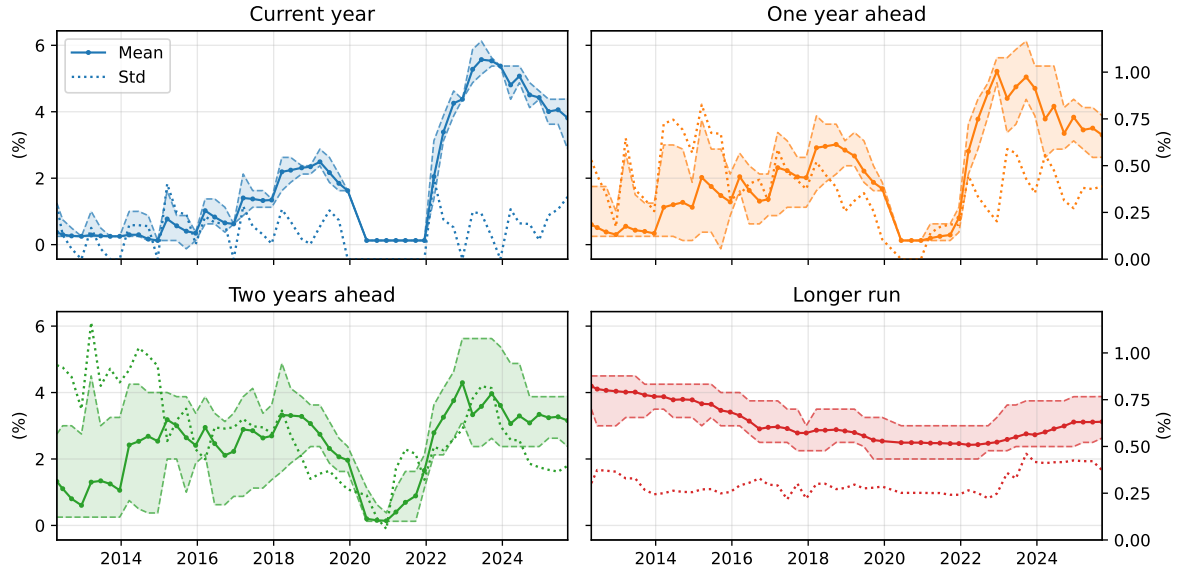


Figure 1.1: Time series of (left axis) mean, maximum and minimum values and (right axis) standard deviation of the (a) current year, (b) 1-year-ahead, (c) 2-year-ahead and (d) longer-run projections of the Fed dot plots.

In Section 1.2, I introduce a preferred-habitat model of the term structure of interest rates à la [Vayanos and Vila \(2021b\)](#) with three key ingredients. First, bond yields are an affine function of two risk factors, which evolve exogenously according to an Ornstein-Uhlenbeck process: the current short-rate, which represents the short-rate that is observable in the market and captures the interest rate policy of the central bank, and the target short-rate, which represents the central bank’s intermediate policy target and captures the forward-guidance policy of the central bank. Crucially, the diffusion coefficient of the target short-rate captures the level of disagreement within the policy committee about the future level of the short-rate at the intermediate horizon or, equivalently, how wide the distribution of the target-short-rate is. Second, the economy features two types of agents: preferred-habitat investors, who have demand for government bonds with specific maturities, and arbitrageurs, who can invest in bonds of any maturity and in the current short-rate, and are mean-variance maximisers over changes in wealth. Third, arbitrageurs hold the residual supply of bonds after accounting for preferred-habitat investors’ demand, which captures the quantitative-easing policy of the central bank. I model the net supply in reduced form and assume that it responds elastically to bond prices. Thus, bond yields are determined endogenously in equilibrium through the trading of preferred-habitat investors and arbitrageurs.

In Section 1.3, I solve the model under the simplifying assumption that the target short-rate evolves deterministically. This is equivalent to a scenario where there is no disagreement within the policy committee about the future level of the short-rate at the intermediate horizon. The price of risk of the target short-rate becomes zero and the only source of risk is the current short-rate. This benchmark deliberately preserves the carry-trade mechanism in [Vayanos and Vila \(2021b\)](#), which I use to show how that familiar risk-bearing channel operates when the central bank's policy stance is split into a current short-rate and a target short-rate. I evaluate three aspects of this mechanism. The first aspect is: how are shocks to the current and target short-rate transmitted along the yield curve? Suppose that a negative shock causes the current or target short-rate to fall. Arbitrageurs face a lower immediate and/or expected future cost of funding bond positions, and therefore have an incentive to borrow at the current short-rate and buy bonds. Their demand pushes bond prices up and yields down, so the curve moves in the direction implied by the lower current or target short-rate. Because these positions lose value when the current short-rate increases, risk-averse arbitrageurs stop short of the trades that would make the yield response fully match the change in expected short-rates. Forward rates therefore adjust only partially to future expected short-rates. The second aspect is: how do shocks to the current and target short-rate impact different parts of the yield curve? For a shock to the current short-rate, the impact on the yield curve decreases with maturity, is one-to-one as the time-to-maturity tends to zero, and tends to zero as the time-to-maturity tends to infinity. For a shock to the target short-rate, the impact on the yield curve is hump-shaped with maturity, and tends to zero as the time-to-maturity tends to zero or infinity. Intuitively, the former has the largest impact on very short maturities, while the latter has the largest impact on intermediate maturities as the current short rate catches up with the new level of the target, and then both rates mean-revert to the natural rate. The third aspect is: how do bond risk premia depend on the shape of the yield curve? Consistently with the evidence in [Fama and Bliss \(1987\)](#) and [Campbell and Shiller \(1991\)](#), the model generates a positive link between bond risk premia and the slope of the yield curve. Suppose that a negative shock causes the current or target short-rate to fall. Then bonds earn a positive risk premium so that risk-averse arbitrageurs are compensated for the extra duration risk they take when engaging in the carry trades.

In Section 1.4, I solve the model under the more realistic assumption that the target short-rate evolves stochastically. I reconsider the previous aspects, and use the model to study the impact of disagreement within the policy committee. Formally, this section is close to the two-factor

analysis in [Vayanos and Vila \(2021b\)](#): once another priced risk factor is present, arbitrageurs no longer evaluate carry trades only through their exposure to the current short-rate, and long-maturity positions can also serve hedging purposes. The contribution here is to interpret that second source of risk as disagreement about the central bank's intermediate target short-rate, rather than as a demand or supply factor, and to study how this disagreement changes the transmission of forward guidance in a model with price-elastic bond supply. Regarding the first and second aspects, the presence of target short-rate risk weakens the pass through of policy shocks and, for sufficiently long maturities, can overturn the sign of the response. To understand this, suppose that a negative shock causes the current short-rate to fall. The resulting carry trades expose arbitrageurs not only to a future reversal in the current short-rate, but also to movements in the target short-rate. Because this second exposure is relatively more important for long bonds, arbitrageurs limit their long positions in those maturities, and may take short positions there to hedge the risk embedded in intermediate-maturity holdings. In that case, long-maturity yields can rise even after an expansionary short-rate shock. Regarding the third aspect, target short-rate risk can reinforce the model-implied comovement between bond risk premia and the slope of the yield curve, but only under additional restrictions on the relative current- and target-short-rate loadings.

In [Section 1.5](#), I calibrate the model in two steps. First, I calibrate the parameters of the target short-rate using dot plot data. Notably, I use the standard deviation of the 1-year-ahead projections to discipline the magnitude of monetary policy disagreement within the FOMC. Then, I calibrate the parameters of the current short-rate and the supply of bonds using data on U.S. Treasury yields and trading volume, and estimates of demand elasticity from the literature. To do so, I match seven empirical moments to their model-generated counterparts.

In [Section 1.6](#), I use the calibrated model to conduct an analysis of the effectiveness of unconventional monetary policy instruments. First, I show that in the model forward guidance is especially effective in changing the yields for intermediate-maturity bonds, but less so for long-term yields. Then, I show that QE is especially effective in changing yields when implemented via purchases of long-maturity bonds. Lastly, I explore how monetary policy disagreement affects the effectiveness of forward guidance and quantitative easing and I present three main results. (i) As monetary policy disagreement increases, forward guidance becomes less effective, while QE becomes more effective. Therefore, the policy implication is that the higher the level of monetary policy disagreement, the more the central bank should lean toward quantitative

easing over forward guidance when aiming to influence the term structure of interest rates. The mechanism is the arbitrageurs' risk-bearing constraint. Higher disagreement raises the compensation required for exposure to target short-rate risk. A forward-guidance announcement changes expected future short rates, but transmitting that change to longer-maturity yields requires arbitrageurs to hold positions that are relatively more exposed to target short-rate risk. When that risk is larger, arbitrageurs reduce those positions, and at long maturities may use them mainly as hedges. As a result, guidance-induced movements in expected short rates pass through less strongly to the long end of the curve. Quantitative easing works through the same risk-bearing constraint in the opposite direction. By reducing the supply of duration risk that arbitrageurs must absorb, QE lowers their exposure to current- and target-rate risk, and therefore lowers the risk premia embedded in yields. When disagreement is high, those premia are larger, so the same reduction in duration risk has a larger effect. (ii) The impact of monetary policy disagreement depends on the design of the policy intervention: in the calibration it can reduce the effect of forward guidance by up to 15.4% and increase the effect of QE by up to 21.0%. (iii) The impact of disagreement – whether positive or negative – is larger for longer-maturity bonds. Therefore, the policy implication is that, when disagreement is high, the tilt in favor of quantitative easing should be even stronger when aiming to influence the long end of the curve.

1.1.1 Literature Review

This paper contributes to two main strands of the literature: the preferred-habitat modeling of the term structure and the growing research on the role of monetary policy uncertainty. [Culbertson \(1957\)](#) and [Modigliani and Sutch \(1966\)](#) were the first to propose the preferred-habitat theory of the term structure. [Vayanos and Vila \(2021b\)](#) and [Vayanos and Vila \(2021c\)](#) formalised the theory by embedding it into a modern no-arbitrage framework. Since the first draft of the model in 2007, the theoretical and empirical literature on the role of supply and demand forces shaping the term structure of interest rates has ballooned, and the framework has been extended in numerous directions. [Greenwood et al. \(2023b\)](#) provided a pedagogical explanation of the model and surveyed several theoretical extensions. These extensions include [King \(2019\)](#) who introduces an effective lower bound on the short rate, [Greenwood et al. \(2018\)](#) and [Hanson et al. \(2021\)](#) who introduce a slow-moving response by arbitrageurs, [Kekre et al. \(2024\)](#) who introduce wealth-dependent risk aversion of arbitrageurs, [He et al. \(2022\)](#), [Hanson](#)

et al. (2024) and Greenwood et al. (2023a) who introduce constraints on the balance-sheet of arbitrageurs, Krishnamurthy and Vissing-Jorgensen (2012) who introduce a non-risk-based convenience yield on government bonds, Campbell et al. (2009) who show how to model both nominal and real curves for Treasury yields, and Ray et al. (2024) who embed the preferred-habitat model in a New Keynesian framework. I contribute to this literature by introducing forward guidance about the short-rate. I do so by distinguishing between the current short-rate, which should be interpreted as the short-rate that is observable in the market, and the target short-rate, which should be interpreted as the central bank's intermediate policy target (e.g. at a one- or two-year horizon), not as the current operating target for the short-rate.

Most relevant for my analysis is the preferred-habitat model of Greenwood et al. (2015a), who introduce both forward guidance about the short-rate and forward guidance about bond supply. Their key idea is that expectations about future short rates and bond supply can change even if their current values do not. For example, when central banks engage in forward guidance or quantitative easing, they communicate their intentions about the future path of short-rates or bond purchases without necessarily changing the short-rate or the bond supply instantaneously. Their modelling approach is the same as the one used in this paper, namely allowing the short-rate and the supply factor to revert to a time-varying mean. This results in a model with four risk factors. Instead, in my model I choose not to model forward guidance about quantitative easing, and retain only two risk factors. This allows me to differentiate my model along a crucial dimension, namely making the supply of bonds price elastic. This allows me to introduce time-varying risk premia that respond to shocks to the current and target short rate, while still solving the model analytically.

While there is a wide body of literature on monetary policy uncertainty, central bank transparency and central bank communications, there are very few studies that relate these topics to the FOMC dot plots released in the Summary of Economic Projections (SEP). Bongard et al. (2021) show that market participants adjust their expectations by assimilating the forward guidance provided in the SEP's projections and in the FOMC's statements. Detmers (2016) considers the effect of dispersion in the SEP's projections of future interest rates on the effectiveness of forward guidance. He finds that forward guidance reduces the sensitivity of interest rates to macroeconomic shocks but that this effect is reduced when there are high levels of disagreement about the projections. He concludes that disagreement lowers the information content of forward guidance and weakens the Fed's commitment as perceived by market participants.

Machado et al. (2015) use a standard New Keynesian model to quantify the uncertainty around interest rate forecasts provided in the FOMC's dot projections. Other papers study the topic of monetary policy disagreement, without relating it to the FOMC dot plots. Banerthghansa and McCracken (2009) analyse disagreement among FOMC members prior to the introduction of the dot plots and highlight that disagreement on the economic outlook is a precursor to disagreement on the appropriate policy path. Riboni and Ruge-Murcia (2014) argue that dissenting votes in monetary policy committees are not just noise, but can be valuable signals that help in forecasting future policy decisions. Instead, Ehrmann and Fratzscher (2013) suggests that dispersed, individualistic communication can be less effective than communicating a clear consensus or majority view. I contribute to this literature by linking the empirical literature that studies the SEP's projections with the theoretical literature on preferred-habitat investors, and by studying the impact of monetary policy disagreement on the effectiveness of unconventional monetary policy tools.

1.2 Model

In this section, I develop a preferred-habitat model of the term structure of interest rates à la Vayanos and Vila (2021b). The economy features a bond market, which consists of default-free but otherwise risky zero-coupon government bonds, and two types of agents, namely preferred-habitat investors, who have preferences for bonds with specific maturities, and arbitrageurs, who can invest in bonds of any maturity and in the current short-rate, and are mean-variance maximisers over changes in wealth.

1.2.1 Bond Market

Time t is continuous and infinite. At each time t , the yield curve consists of a continuum of default-free zero-coupon government bonds that have face value one and mature at time $t + \tau$, where $\tau \in (0, \infty)$. Assuming that the interval of times to maturity is infinite is without loss of generality because I can set the supply of bonds to be zero for bonds with sufficiently long maturities. Let $P_t(\tau)$ be the time- t price of the bond that matures at time $t + \tau$, and denote its logarithm by $p_t(\tau) \equiv \ln P_t(\tau)$. The time- t continuously compounded yield to maturity of the

τ -year bond is:

$$y_t(\tau) = -\frac{p_t(\tau)}{\tau}.$$

Similarly, the time- t continuously compounded forward rate between maturities $\tau - \Delta\tau$ and τ is:

$$f_t(\tau - \Delta\tau, \tau) = -\frac{p_t(\tau) - p_t(\tau - \Delta\tau)}{\Delta\tau}.$$

The short-rate is defined as the limit of the yield as time to maturity goes to zero, $r_t = \lim_{\tau \rightarrow 0} y_t(\tau)$. Similarly, the instantaneous forward rate for maturity τ is defined as the limit of the forward rate between maturities $\tau - \Delta\tau$ and τ as the time interval goes to zero, $f_t(\tau) = \lim_{\Delta\tau \rightarrow 0} f_t(\tau - \Delta\tau, \tau) = -\partial_\tau p_t(\tau) = y_t(\tau) + \tau \partial_\tau y_t(\tau)$.

I assume that the short-rate follows a stochastic mean process (see e.g. [Balduzzi et al. \(1998\)](#) and [Chen \(1996\)](#)), whereby both the short-rate – henceforth current short-rate – and its mean-reversion level – henceforth target short-rate – evolve exogenously according to Ornstein-Uhlenbeck processes:

$$dr_t = \kappa_r (\eta_t - r_t) dt + \sigma_r dz_{r,t},$$

$$d\eta_t = \kappa_\eta (r^* - \eta_t) dt + \sigma_\eta dz_{\eta,t},$$

where $(\kappa_r, \kappa_\eta, \sigma_r, \sigma_\eta, r^*)$ are positive constants, and $(z_{r,t}, z_{\eta,t})$ are two independent standard Brownian motions under the physical probability measure.

The current short-rate r_t should be interpreted as the short-rate that is observable in the market, and it captures the interest rate policy of the central bank. The target short-rate η_t should be interpreted as the central bank's intermediate policy target (e.g. at a one- or two-year horizon), not as the current operating target for the short-rate, and it captures the forward-guidance policy of the central bank. The mean-reversion level of the target short-rate r^* should be interpreted as the natural rate of interest, i.e. the rate that would prevail when the output gap is zero and inflation is stable. Over time, the current short-rate mean reverts to the target short-rate announced by the central bank at the speed κ_r , and the target short-rate in turn mean

reverts to the natural rate at the speed κ_η . The diffusion coefficient σ_r captures the volatility of the current short-rate. The Brownian increments $dz_{r,t}$ capture both shocks due to trading of short-term bonds by arbitrageurs, and shocks (exogenous from the perspective of arbitrageurs) due to interest rate policy announcements by the central bank. The diffusion coefficient σ_η captures the disagreement within the monetary policy committee about the future level of the short-rate at the intermediate horizon or, equivalently, how wide the distribution of the target-short-rate is. The Brownian increments $dz_{\eta,t}$ capture shocks due to the central bank receiving new information about the economy, noise due to converting this new information into forward-guidance policy announcements, and shocks due to the policy announcements themselves that provide forward-guidance on the future path of the short-rate.

1.2.2 Agents

There are two types of agents. Arbitrageurs (e.g. market makers, and hedge funds) are competitive and form a continuum with unit mass. They can invest in bonds of any maturity, and in the current short-rate. They choose their bond portfolio to trade off the mean and variance of the instantaneous changes in their time- t wealth W_t . The arbitrageur's optimisation problem is²:

$$\begin{aligned} \max_{\{X_t(\tau)\}_{\tau \in (0, \infty)}} \quad & \mathbb{E}_t(dW_t) - \frac{a}{2} \text{Var}_t(dW_t), \\ \text{s.t.} \quad & dW_t = \left(W_t - \int_0^\infty X_t(\tau) d\tau \right) r_t dt + \int_0^\infty X_t(\tau) \frac{dP_t(\tau)}{P_t(\tau)} d\tau, \end{aligned}$$

where $a \geq 0$ is arbitrageurs' risk aversion coefficient, and $X_t(\tau)$ is the time- t value of arbitrageurs' holdings of bonds with maturity τ (expressed in present-value terms). Note that the second term in the budget constraint is the return from investing in bonds, and the first term is the return from investing the remaining wealth in the current short-rate.

Preferred-habitat investors have preferences for bonds with specific maturities, and hence they are active in a particular segment of the term structure. These investors refer to all bond market participants other than arbitrageurs: they include both buyers and sellers of bonds (e.g. households, financial institutions, and the central bank), as well as issuers (e.g. the fiscal authority). For instance, pension funds typically have demand for long-maturity bonds, because

²See [Vayanos and Vila \(2021b\)](#) for a brief micro-foundation of the objective function based on overlapping generations and von Neumann-Morgenstern preferences.

their duration matches that of their pension liabilities; insurance companies typically demand long- and intermediate-maturity bonds, because their duration matches that of their insurance liabilities; in contrast, money-market funds operate under mandates that require them to hold short-maturity bonds. Investors with habitat in $[\tau, \tau + d\tau]$ are in measure $d\tau$. At time- t , they hold a position $Z_t(\tau)$ in the τ -year bond (expressed in present-value terms) and hold no other bonds.

I denote the time- t gross supply of the τ -year bond by $S_{G,t}(\tau)$. Thus, market clearing requires that arbitrageurs hold the net supply $S_t(\tau)$ – henceforth supply – coming from preferred-habitat investors³:

$$X_t(\tau) = S_t(\tau) \equiv S_{G,t}(\tau) - Z_t(\tau).$$

I model the supply in reduced form⁴. I assume that the time- t supply of the τ -year bond (expressed in present-value terms and measured in dollars) responds elastically to bond prices and, for simplicity, does not depend on supply shocks. In particular, I assume that the supply function is linear and increasing in the logarithm of the bond price:

$$S_t(\tau) = \alpha_0(\tau) + \beta(\tau)p_t(\tau),$$

where $\alpha_0(\tau)$ and $\beta(\tau)$ are functions that depend only on the time to maturity τ . The intercept coefficient $\alpha_0(\tau) \geq 0$ is the component of supply that is independent of the current bond price. This could reflect the government's regular debt issuance requirements less any price-inelastic demand from preferred-habitat investors who have fixed-maturity needs. The slope coefficient $\beta(\tau) \geq 0$ is the sensitivity of the supply of the τ -year bond to changes in its log price. Its sign implies that the bond supply increases in its price: this could be because gross bond supply increases with price (e.g. because governments issue more debt when interest rates are low), or because bond demand from preferred-habitat investors decreases with price (e.g. investors substitute away from bonds and towards other asset classes when interest rates are low). Note that by setting $\alpha_0(\tau) = \beta(\tau) = 0$ for $\tau > T$, I can set an upper limit on the maximum maturity of

³Equivalently, one could assume that government bonds are in zero gross supply, $S_{G,t}(\tau) = 0$, and redefine the preferred-habitat investor demand $Z_t(\tau)$ as net demand, i.e. the demand by preferred-habitat investors for the τ -year bond, net of the government gross supply for that bond. The market-clearing condition would be written $X_t(\tau) = S_t(\tau) \equiv -Z_t(\tau)$.

⁴See [Vayanos and Vila \(2021c\)](#) for an optimizing foundation of the reduced form model of the demand of preferred-habitat investors.

the bonds that are traded in the market.

1.2.3 Equilibrium Without Arbitrageurs

As a benchmark, I derive the segmentation equilibrium that would prevail in the absence of arbitrageurs, when the market for bonds of a given maturity must be cleared solely by the actions of preferred-habitat investors in that maturity segment. As explained previously, the yield $y_t(\tau) = -\frac{p_t(\tau)}{\tau}$ for maturity τ is determined by setting the net supply by preferred-habitat investors to zero:

$$y_t(\tau) = \frac{\alpha_0(\tau)}{\tau\beta(\tau)}.$$

The yield $y_t(\tau)$ for maturity τ is constant over time and is disconnected from the time-varying current and target short-rates. It depends only on the supply intercept $\alpha_0(\tau)$ and the supply slope $\beta(\tau)$ for maturity τ . An increase in $\alpha_0(\tau)$ increases the supply of the τ -year bond by preferred-habitat investors, and hence increases its yield. The effect is weaker the larger $\beta(\tau)$ is because the supply by preferred-habitat investors is more price-elastic.

1.2.4 Equilibrium With Arbitrageurs

To derive the equilibrium with arbitrageurs, I follow a three-step procedure: (1) Conjecture a functional form for equilibrium prices; (2) Derive the arbitrageurs' first-order condition given the conjectured prices; (3) Impose the market clearing condition in the first-order condition, solve the resulting differential equation, and verify that the prices are as conjectured.

(Step 1) I look for an equilibrium where bond yields are affine functions of the current and target short-rate risk factors, r_t and η_t . In the conjectured equilibrium, the time- t price of the τ -year bond is:

$$P_t(\tau) = e^{-(A_r(\tau)r_t + A_\eta(\tau)\eta_t + C(\tau))},$$

where $(A_r(\tau), A_\eta(\tau), C(\tau))$ are functions that depend only on the time to maturity τ . The functions $A_r(\tau)$ and $A_\eta(\tau)$ characterize the sensitivity of bond prices to r_t and η_t as a function of bond time to maturity τ , where I define factor sensitivity or dollar duration with respect to the factor as the percentage decline in price per unit increase in the factor.

(Step 2) I compute the instantaneous return of $P_t(\tau)$ in an infinitesimal time interval $[t, t+dt]$. Note that $P_t(\tau)$ depends on r_t , η_t and t , since $\tau = \mathcal{T} - t$, where $\mathcal{T}$ denotes the maturity of the bond. Notice also that since $\frac{d\tau}{dt} = -1$, I can write $\frac{\partial P_t(\tau)}{\partial t} = \frac{\partial P_t(\tau)}{\partial \tau} \frac{d\tau}{dt} = -\frac{\partial P_t(\tau)}{\partial \tau}$. Applying Ito's lemma and using the process for r_t and η_t , the change in the bond price over dt is:

$$dP_t(\tau) = -\partial_\tau P_t(\tau)dt + \partial_{r_t} P_t(\tau)dr_t + \frac{1}{2}\partial_{r_t}^2 P_t(\tau)dr_t^2 + \partial_{\eta_t} P_t(\tau)d\eta_t + \frac{1}{2}\partial_{\eta_t}^2 P_t(\tau)d\eta_t^2$$

$$\frac{dP_t(\tau)}{P_t(\tau)} = \mu_t(\tau)dt - A_r(\tau)\sigma_r dz_{r,t} - A_\eta(\tau)\sigma_\eta dz_{\eta,t},$$

where the instantaneous expected return is:

$$\begin{aligned} \mu_t(\tau) &= \partial_\tau A_r(\tau)r_t + \partial_\tau A_\eta(\tau)\eta_t + \partial_\tau C(\tau) - A_r(\tau)\kappa_r(\eta_t - r_t) - A_\eta(\tau)\kappa_\eta(r^* - \eta_t) \\ &\quad + \frac{1}{2}A_r^2(\tau)\sigma_r^2 + \frac{1}{2}A_\eta^2(\tau)\sigma_\eta^2. \end{aligned}$$

Substituting $\frac{dP_t(\tau)}{P_t(\tau)}$ into the arbitrageurs' budget constraint:

$$\begin{aligned} dW_t &= \left(W_t - \int_0^\infty X_t(\tau)d\tau\right)r_t dt + \int_0^\infty X_t(\tau) \left(\mu_t(\tau)dt - A_r(\tau)\sigma_r dz_{r,t} - A_\eta(\tau)\sigma_\eta dz_{\eta,t}\right) d\tau \\ &= \left(W_t r_t + \int_0^\infty X_t(\tau) (\mu_t(\tau) - r_t) d\tau\right) dt - \left(\sigma_r \int_0^\infty X_t(\tau) A_r(\tau) d\tau\right) dz_{r,t} \\ &\quad - \left(\sigma_\eta \int_0^\infty X_t(\tau) A_\eta(\tau) d\tau\right) dz_{\eta,t}. \end{aligned}$$

Substituting dW_t into the arbitrageurs' objective:

$$\begin{aligned} \max_{\{X_t(\tau)\}_{\tau \in (0, \infty)}} W_t r_t + \int_0^\infty X_t(\tau) (\mu_t(\tau) - r_t) d\tau - \frac{a\sigma_r^2}{2} \left(\int_0^\infty X_t(\tau) A_r(\tau) d\tau\right)^2 \\ - \frac{a\sigma_\eta^2}{2} \left(\int_0^\infty X_t(\tau) A_\eta(\tau) d\tau\right)^2. \end{aligned}$$

Pointwise maximisation yields the arbitrageurs' first-order condition:

$$\mu_t(\tau) - r_t = A_r(\tau)\lambda_{r,t} + A_\eta(\tau)\lambda_{\eta,t},$$

where the prices of risk of the current and target short-rate risk factors are:

$$\lambda_{f,t} = a\sigma_f^2 \int_0^\infty X_t(\tau) A_f(\tau) d\tau \quad \text{for } f \in \{r, \eta\}.$$

The left hand side of the first-order condition is the increase in the expected return of the arbitrageurs' portfolio if they shift one unit of the numeraire from the current short-rate to the τ -year bond, $\mu_t(\tau) - r_t$. This increase is equal to the difference between the expected return of the bond $\mu_t(\tau)$ and the current short-rate r_t . The right hand side is the product of the arbitrageurs' risk aversion coefficient a times the increase in the risk of the arbitrageurs' portfolio due to the current short-rate risk factor, $\sigma_r^2 A_r(\tau) \int_0^\infty X_t(\tau) A_r(\tau) d\tau$, and the target short-rate risk factor, $\sigma_\eta^2 A_\eta(\tau) \int_0^\infty X_t(\tau) A_\eta(\tau) d\tau$. This increase is equal to the covariance between the return on the additional investment in the bond and the return on the portfolio. Since the two risk factors are independent, this covariance reduces to the sum over the factors of the product of the factor's variance σ_f^2 times the sensitivities of the two returns to the factor $A_f(\tau) \int_0^\infty X_t(\tau) A_f(\tau) d\tau$. The sensitivity of the bond's return to factor f is $A_f(\tau)$ and the sensitivity of the portfolio's return is $\int_0^\infty X_t(\tau) A_f(\tau) d\tau$. The sensitivity of the portfolio to factor f (or dollar duration with respect to factor f) is obtained by multiplying the arbitrageurs' position $X_t(\tau)$ in the τ -year bond by the bond's sensitivity $A_f(\tau)$ to the factor and then summing across maturities. Alternatively, the first-order condition can be interpreted as following from the absence of arbitrage, which requires that there exists a price for each risk factor, common across assets, such that the expected excess return of any asset is equal to the sum across factors of the asset's sensitivity to each factor times the price of risk of that factor. With multiple factors, the no arbitrage condition requires that the asset's expected excess return multiplied by the inverse of the vector of the asset's sensitivity to the factors is equal to the vector of prices of risk, which is the same for all assets. If the product differed across bonds, then it would be possible to construct riskless arbitrage portfolios. However, the absence of arbitrage alone imposes no restrictions on the prices of risk $\lambda_{f,t}$, which need to be determined by equilibrium conditions.

(Step 3) Using the market-clearing condition and the conjectured price of the bonds, I determine that the prices of risk of the current and target short-rate risk factors are:

$$\begin{aligned}
\lambda_{f,t} &= a\sigma_f^2 \int_0^\infty (\alpha_0(\tau) + \beta(\tau)p_t(\tau)) A_f(\tau) d\tau \\
&= a\sigma_f^2 \int_0^\infty (\alpha_0(\tau) - \beta(\tau) (A_r(\tau)r_t + A_\eta(\tau)\eta_t + C(\tau))) A_f(\tau) d\tau \\
&= \lambda_f + \lambda_{fr}r_t + \lambda_{f\eta}\eta_t,
\end{aligned}$$

where:

$$\lambda_f = a\sigma_f^2 \int_0^\infty (\alpha_0(\tau) - \beta(\tau)C(\tau)) A_f(\tau) d\tau,$$

$$\lambda_{fr} = -a\sigma_f^2 \int_0^\infty \beta(\tau) A_r(\tau) A_f(\tau) d\tau,$$

$$\lambda_{f\eta} = -a\sigma_f^2 \int_0^\infty \beta(\tau) A_\eta(\tau) A_f(\tau) d\tau.$$

Note that the prices of risk are time-varying because the elastic supply assumption implies that arbitrageurs' bond positions respond to changes in the current and target short-rate, and hence to bond prices. Substituting $\mu_t(\tau)$, $\lambda_{r,t}$ and $\lambda_{\eta,t}$ into the first-order condition, I find an affine equation in r_t and η_t which I use to solve for the functions $A_r(\tau)$, $A_\eta(\tau)$ and $C(\tau)$:

$$\begin{aligned} & \partial_\tau A_r(\tau) r_t + \partial_\tau A_\eta(\tau) \eta_t + \partial_\tau C(\tau) - A_r(\tau) \kappa_r (\eta_t - r_t) - A_\eta(\tau) \kappa_\eta (r^* - \eta_t) \\ & + \frac{1}{2} A_r^2(\tau) \sigma_r^2 + \frac{1}{2} A_\eta^2(\tau) \sigma_\eta^2 - r_t = A_r(\tau) (\lambda_r + \lambda_{rr} r_t + \lambda_{r\eta} \eta_t) + A_\eta(\tau) (\lambda_\eta + \lambda_{\eta r} r_t + \lambda_{\eta\eta} \eta_t). \end{aligned}$$

The differential equation must hold for all values of r_t and η_t . Hence, identifying the linear terms in r_t and η_t , and the terms independent of r_t and η_t , I find a system of ODEs:

$$\partial_\tau A_r(\tau) + A_r(\tau) (\kappa_r - \lambda_{rr}) - A_\eta(\tau) \lambda_{\eta r} - 1 = 0,$$

$$\partial_\tau A_\eta(\tau) + A_\eta(\tau) (\kappa_\eta - \lambda_{\eta\eta}) - A_r(\tau) (\kappa_r + \lambda_{r\eta}) = 0,$$

$$\partial_\tau C(\tau) - A_r(\tau) \lambda_r + \frac{1}{2} A_r^2(\tau) \sigma_r^2 - A_\eta(\tau) (\kappa_\eta r^* + \lambda_\eta) + \frac{1}{2} A_\eta^2(\tau) \sigma_\eta^2 = 0.$$

The last step required to determine equilibrium bond yields is to solve the above system of three differential equations for $(A_r(\tau), A_\eta(\tau), C(\tau))$. The system needs to be solved with the terminal conditions $A_r(0) = 0$, $A_\eta(0) = 0$ and $C(0) = 0$ respectively, because maturing bonds are worth their face value of one.

1.3 Deterministic Target Short-Rate

Suppose that the target short-rate evolves deterministically. This is equivalent to setting $\sigma_\eta = 0$ in the diffusion process of the target short-rate. The three components λ_η , $\lambda_{r\eta}$ and $\lambda_{\eta\eta}$ of the price of risk of the target short-rate become zero, and the system of equations reduces to:

$$\partial_\tau A_r(\tau) + A_r(\tau)(\kappa_r - \lambda_{rr}) - 1 = 0,$$

$$\partial_\tau A_\eta(\tau) + A_\eta(\tau)\kappa_\eta - A_r(\tau)(\kappa_r + \lambda_{r\eta}) = 0,$$

$$\partial_\tau C(\tau) - A_r(\tau)\lambda_r + \frac{1}{2}A_r^2(\tau)\sigma_r^2 - A_\eta(\tau)\kappa_\eta r^* = 0.$$

Proposition 1.1 *The functions $(A_r(\tau), A_\eta(\tau))$ are given by:*

$$A_r(\tau) = \frac{1 - e^{-\tau\kappa_{rr}^*}}{\kappa_{rr}^*} \geq 0.$$

$$A_\eta(\tau) = \frac{(1 - e^{-\tau\kappa_\eta})\kappa_{rr}^*\kappa_{r\eta}^* - (1 - e^{-\tau\kappa_{rr}^*})\kappa_\eta\kappa_{r\eta}^*}{(\kappa_{rr}^* - \kappa_\eta)\kappa_{rr}^*\kappa_\eta} \geq 0.$$

where $\kappa_{rr}^* \equiv \kappa_r - \lambda_{rr} > \kappa_r$ satisfies the fixed-point condition:

$$\kappa_{rr}^* = \kappa_r + a\sigma_r^2 \int_0^\infty \beta(\tau) \left(\frac{1 - e^{-\tau\kappa_{rr}^*}}{\kappa_{rr}^*} \right)^2 d\tau,$$

and $\kappa_{r\eta}^* \equiv \kappa_r + \lambda_{r\eta} < \kappa_r$ satisfies the fixed-point condition:

$$\kappa_{r\eta}^* = \frac{\kappa_r}{1 + a\sigma_r^2 \int_0^\infty \beta(\tau) \left(\frac{(1 - e^{-\tau\kappa_\eta})\kappa_{rr}^* - (1 - e^{-\tau\kappa_{rr}^*})\kappa_\eta}{(\kappa_{rr}^* - \kappa_\eta)\kappa_{rr}^*\kappa_\eta} \right) \left(\frac{1 - e^{-\tau\kappa_{rr}^*}}{\kappa_{rr}^*} \right) d\tau}.$$

1.3.1 Interest Rate Policy

In the segmentation equilibrium, yields are disconnected from the current and target short-rate and each maturity clears within its own habitat. Arbitrageurs break this separation by transmitting short-rate shocks to bond yields via carry trades. First, suppose that a negative shock causes the current short-rate to fall below the value that the yield of the τ -year bond would

take in the segmentation equilibrium. The benefit to arbitrageurs is twofold: first, a lower short-rate corresponds to an immediate reduction in the funding cost of their bond positions; second, given that the current short-rate mean reverts to the (unchanged) target short-rate, arbitrageurs also revise downwards their expectations of the future funding cost (see Figure 1.2 (a)). Bonds therefore become more attractive relative to investing at the current short-rate. Arbitrageurs respond by borrowing at r_t and purchasing the τ -year bond, which raises its price and lowers its yield. Next, suppose that a negative shock causes the target short-rate to fall. While arbitrageurs do not benefit from an immediate reduction in the funding cost, they do revise downwards their expectations of the future cost (see Figure 1.2 (b)). This again increases the attractiveness of bond positions, inducing arbitrageurs to buy bonds and pushing yields down.

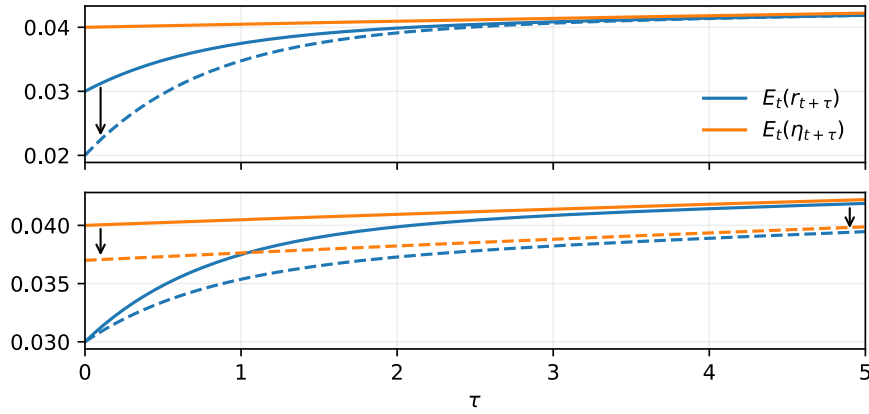


Figure 1.2: The effect of a reduction in (a) the current short-rate and (b) the target short-rate on the expected future current short-rate $\mathbb{E}_t(r_{t+\tau})$ at time $t + \tau$ and the expected future target short-rate $\mathbb{E}_t(\eta_{t+\tau})$ at time $t + \tau$. The solid lines represent values pre-shock, while the dashed lines represent values post-shock.

To quantify the extent to which arbitrageurs transmit short-rate shocks to bond yields, I measure how a time- t shock to the current and target short-rate affects the expected future current short-rate $\mathbb{E}_t(r_{t+\tau})$ at time $t + \tau$ and the instantaneous forward rate $f_t(\tau)$ for maturity τ . Recall that when the expectations hypothesis of the term structure holds, the forward-rate response coincides with the response of expected future current short-rates (see Appendix A.2 for a review). The difference between the two responses is therefore a direct measure of incomplete transmission along the yield curve.

Proposition 1.2 *Holding constant the other risk factor:*

- A unit shock to the current short-rate r_t raises the expected future current short-rate $\mathbb{E}_t(r_{t+\tau})$ at time $t + \tau$ by $\partial_{r_t}\mathbb{E}_t(r_{t+\tau}) = e^{-\tau\kappa_r}$ and the instantaneous forward rate $f_t(\tau)$ for maturity τ by $\partial_{r_t}f_t(\tau) = e^{-\tau\kappa_{rr}^*}$.

- A unit shock to the target short-rate η_t raises the expected future current short-rate $\mathbb{E}_t(r_{t+\tau})$ at time $t + \tau$ by $\partial_{\eta_t} \mathbb{E}_t(r_{t+\tau}) = \frac{(e^{-\tau\kappa_\eta} - e^{-\tau\kappa_r})\kappa_r}{\kappa_r - \kappa_\eta}$ and the instantaneous forward rate $f_t(\tau)$ for maturity τ by $\partial_{\eta_t} f_t(\tau) = \frac{(e^{-\tau\kappa_\eta} - e^{-\tau\kappa_{rr}^*})\kappa_{r\eta}^*}{\kappa_{rr}^* - \kappa_\eta}$.

The forward rate underreacts (and the expectations hypothesis fails) if arbitrageurs are risk-averse ($a > 0$) and the supply coming from preferred-habitat investors is price-elastic ($\beta(\tau) > 0$). The magnitude of the underreaction increases in σ_r^2 , a and $\beta(\tau)$.

The proposition shows that the extent to which arbitrageurs transmit short-rate shocks to bond yields depends on three parameters: the variance of the current short-rate σ_r^2 , the arbitrageurs' risk-aversion coefficient a , and the price sensitivity of preferred-habitat supply $\beta(\tau)$. When $a > 0$ and $\beta(\tau) > 0$, arbitrageurs require compensation for taking positions whose value can move against them. As a result, the forward curve does not adjust as much as the expected path of future current short-rates, and hence, arbitrageurs transmit short-rate shocks to bond yields only partially. For current-short-rate shocks, the extent of the underreaction is captured by $\kappa_{rr}^* > \kappa_r$: under the risk-neutral measure, the current short-rate mean reverts more quickly than under the physical measure. For target-short-rate shocks, the extent of the underreaction is captured by $\kappa_{rr}^* > \kappa_r$ and $\kappa_{r\eta}^* < \kappa_r$: the influence of the target short-rate on the current short-rate under the risk-neutral measure is smaller than under the physical measure. The limiting cases make the role of risk bearing clear. If $a = 0$, arbitrageurs do not require to be compensated for the risk involved in carry trades (e.g. an increase in the current short-rate when they borrow short term to buy bonds), and take positions large enough to restore full transmission of short-rate shocks to bond yields. If $\beta(\tau) = 0$ for all $\tau \in (0, \infty)$, the quantity that arbitrageurs must absorb does not respond to prices; yields can then adjust without changing arbitrageurs' exposure to carry-trade risk. Transmission is again complete, but for a different reason. More generally, pass-through weakens when arbitrageurs are more risk-averse, current short-rate risk is larger, or preferred-habitat supply is more price-elastic. Figure 1.3 shows the resulting gap between the instantaneous forward-rate response and the response of expected future current short-rates.

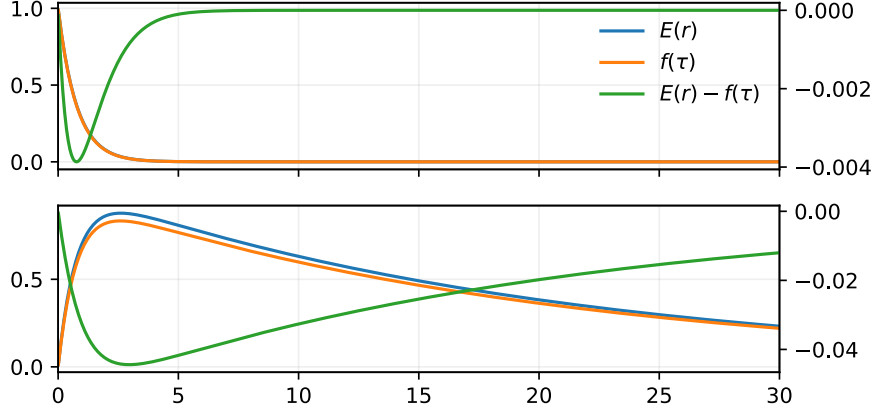


Figure 1.3: The effect of a unit shock to (a) the current short-rate and (b) the target short-rate on (left axis) the expected future current short-rate $\mathbb{E}_t(r_{t+\tau})$ at time $t + \tau$ and the instantaneous forward rate $f_t(\tau)$ for maturity τ , and (right axis) the resulting underreaction of forward rates.

While studying the effect of shocks on the expected future current short-rate and the instantaneous forward rate provides insight into the transmission mechanism, I now quantify the effects of short-rate shocks on bond yields.

Proposition 1.3 *Holding constant the other risk factor:*

- A unit shock to the current short-rate r_t raises the yield of the τ -year bond by:

$$\partial_{r_t} y_t(\tau) = \frac{A_r(\tau)}{\tau} \geq 0.$$

The effect decreases with time to maturity, is equal to one for $\tau = 0$, and to zero for $\tau \rightarrow \infty$.

- A unit shock to the target short-rate η_t raises the yield of the τ -year bond by:

$$\partial_{\eta_t} y_t(\tau) = \frac{A_\eta(\tau)}{\tau} \geq 0.$$

The effect is hump-shaped with time to maturity, and is equal to zero for $\tau = 0$ and $\tau \rightarrow \infty$.

The proposition shows that the effect of shocks to the current short-rate decreases with maturity. Intuitively, since r_t mean-reverts, the effect of shocks to r_t on the expected future current short-rate is largest in the near future. On the other hand, the effect of shocks to the target short-rate is hump-shaped with maturity. Intuitively, the effect of shocks to η_t on $\mathbb{E}_t(r_{t+\tau})$ is small for short maturities because these shocks do not affect r_t , increases with maturity as $\mathbb{E}_t(r_{t+\tau})$ catches up with the new value of η_t , and decreases again to zero because η_t mean-reverts. Figure 1.4 shows the effect of a unit shock to the current and target short-rate on bond yields.

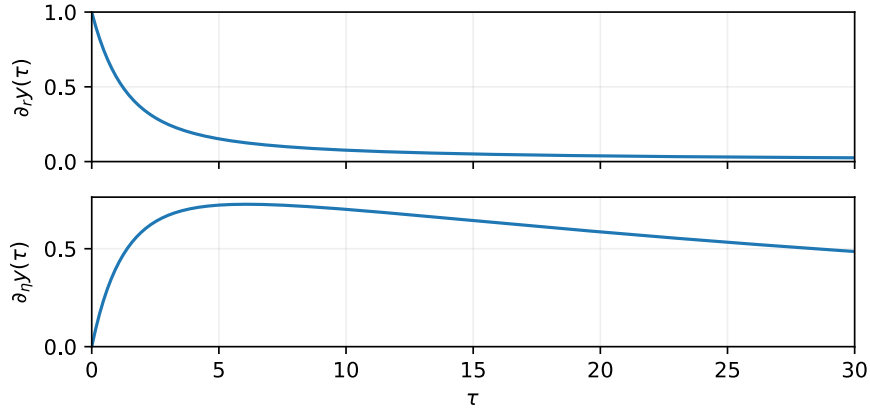


Figure 1.4: The effect of a unit shock to (a) the current short-rate and (b) the target short-rate on bond yields.

These results give a simple interpretation of monetary-policy transmission. If r_t is the rate paid on bank reserves and the relevant arbitrageurs are banks, then changes in the reserve rate affect long-maturity yields through banks' willingness to absorb duration risk. The pass-through is weaker when that risk is more costly to hold: arbitrageurs may be more risk-averse, the current short-rate may be more volatile, or the supply coming from preferred-habitat investors may react more strongly to bond prices. The same channel implies that arbitrageurs are compensated for transmitting policy shocks. The compensation is the expected return on the carry trades, and it appears in equilibrium as bond risk premia.

1.3.2 Risk Premia

Under the expectations hypothesis, bonds earn the risk-free rate in expectation (or equivalently, the instantaneous holding premium is zero). Instead, when $a > 0$ and $\beta(\tau) > 0$, expected returns must compensate the positions that clear the market. Suppose that a negative shock causes the current or target short-rate to fall. As shown above, in either case, yields fall, as price-elastic preferred-habitat investors sell bonds. Then bonds earn a positive expected return in excess of the risk-free rate (or equivalently, the instantaneous holding premium is positive) so that risk-averse arbitrageurs are induced to buy them. This is because risk-averse arbitrageurs engage in carry trades only if they expect to earn high returns as compensation for the extra duration risk they take. In the setting with deterministic target short-rate, the instantaneous holding premium arises to compensate arbitrageurs for their exposure to a single source of risk, namely innovations in the current short-rate. The magnitude of this premium for a given bond, $\mu_t(\tau) - r_t$, depends on the bond's sensitivity to the short-rate, $A_r(\tau)$, and the price of risk of the

current short-rate, $\lambda_{r,t}$, where in turn the price of risk depends on both state variables, r_t and η_t . Figure 1.5 plots the effect of a unit shock to the current short-rate, $\partial_{r_t}(\mu_t(\tau) - r_t) = A_r(\tau)\lambda_{rr}$, and to the target short-rate, $\partial_{\eta_t}(\mu_t(\tau) - r_t) = A_r(\tau)\lambda_{r\eta}$, on the instantaneous holding premia.

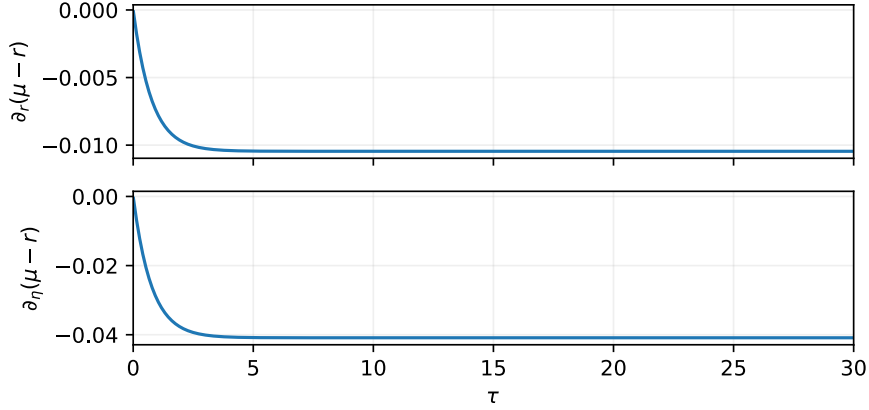


Figure 1.5: The effect of a unit shock to (a) the current short-rate and (b) the target short-rate on instantaneous holding premia.

Next, I study how my model aligns with two established empirical tests that have been widely cited in the literature as evidence of the failure of the expectations hypothesis. The first empirical test I consider comes from [Fama and Bliss \(1987\)](#), who estimate the following regression to study the relationship between the slope of the term structure and holding premia:

$$\frac{p_{t+\Delta\tau}(\tau - \Delta\tau) - p_t(\tau)}{\Delta\tau} - y_t(\Delta\tau) = a_{FB} + b_{FB} (f_t(\tau - \Delta\tau, \tau) - y_t(\Delta\tau)) + e_{t+\Delta\tau},$$

The dependent variable is the difference between two competing investment strategies over a fixed holding period $\Delta\tau$: a risky strategy (with unknown return), which entails purchasing a τ -year bond at time t and selling it at time $t + \Delta\tau$; and a risk-free strategy (with known return), which entails purchasing a $\Delta\tau$ -year bond at time t and holding it to maturity. This ex-post excess return quantifies the holding premium (i.e. the observable equivalent of the instantaneous holding premium $\mu_t(\tau) - r_t$), and represents the compensation earned for bearing the risk of holding the long-term bond over the short-term one. The independent variable is the slope of the term structure as measured by the forward-spot spread, which is computed as the difference between the forward rate between maturities $\tau - \Delta\tau$ and τ , and the yield of the $\Delta\tau$ -year bond. Recall that the expectations hypothesis predicts that holding premia are constant. Therefore, there should be no relation between the unknown quantity on the left-hand side and the known quantity on the right-hand side, and the regression coefficient should be zero. Instead, the

authors find a positive relationship: the coefficient b_{FB} is positive, greater than one for most maturities, and increasing with maturity.

The second empirical test I consider comes from [Campbell and Shiller \(1991\)](#), who estimate the following regression to study the relationship between expected and actual changes in bond yields:

$$y_{t+\Delta\tau}(\tau - \Delta\tau) - y_t(\tau) = a_{CS} + b_{CS} \frac{\Delta\tau}{\tau - \Delta\tau} (y_t(\tau) - y_t(\Delta\tau)) + e_{t+\Delta\tau},$$

The dependent variable is the change in the yield of the τ -year bond over the time interval $\Delta\tau$. The independent variable is the slope of the term structure as measured by the yield spread, which is computed as the difference between the yield of the τ -year bond and the yield of the $\Delta\tau$ -year bond, scaled so that the regression coefficient $b_{CS} = 1$ under the expectations hypothesis. It can be shown that the expectations hypothesis implies that the independent variable is equivalent to the expected value of the dependent variable. Therefore, the regression coefficient should be one. Instead, the authors find a negative relationship: the coefficient b_{CS} is smaller than one, negative for most maturities, and decreasing with maturity. This finding is related to the positive relationship between the slope of the term structure and holding premia. Indeed, suppose that the term structure has a slope higher than average. Because bonds earn positive expected excess returns, their yields increase by less than under the expectations hypothesis, implying $b_{CS} < 1$.

Proposition 1.4 *For $\Delta\tau \rightarrow 0$ and for all τ :*

- *The FB regression coefficient is $b_{FB} = \frac{\kappa_{rr}^* - \kappa_r}{\kappa_{rr}^*}$. It is positive if arbitrageurs are risk-averse ($a > 0$), and the supply by preferred-habitat investors is price-elastic ($\beta(\tau) > 0$ in a positive-measure subset of $(0, \infty)$). It increases in a , σ_r^2 and $\beta(\tau)$.*
- *The CS regression coefficient is $b_{CS} = 1 - \frac{A_r(\tau)\tau(\kappa_{rr}^* - \kappa_r)}{\tau - A_r(\tau)}$. It is smaller than one if arbitrageurs are risk-averse ($a > 0$), and the supply by preferred-habitat investors is price-elastic ($\beta(\tau) > 0$ in a positive-measure subset of $(0, \infty)$). It increases in τ .*

The proposition shows that the deterministic-target-short-rate benchmark reproduces the qualitative failure of the expectations hypothesis: $b_{FB} > 0$ and $b_{CS} < 1$ when arbitrageurs are risk-averse and preferred-habitat net supply is price-elastic. The FB coefficient increases when arbitrageurs' risk aversion a , current-short-rate variance σ_r^2 , or the supply slope $\beta(\tau)$ increase,

while the CS coefficient increases with maturity τ .

1.4 Stochastic Target Short-Rate

I now assume that the target short-rate evolves stochastically, and I go back to the system of ODEs derived in section 1.2.4. Recognising that the first two equations constitute a coupled system of differential equations, I re-write the system of three ODEs as a system of two ODEs in vector-matrix form:

$$\partial_\tau A(\tau) + \Psi A(\tau) = \iota,$$

$$\partial_\tau C(\tau) - A(\tau)^\top \psi + \frac{1}{2} A(\tau)^\top \Sigma \Sigma^\top A(\tau) = 0,$$

with:

$$\Psi \equiv K^\top + a \left(\int_0^\infty \beta(\tau) A(\tau) A(\tau)^\top d\tau \right) \Sigma \Sigma^\top,$$

$$\psi \equiv r^* K \mathbf{1} + a \Sigma \Sigma^\top \int_0^\infty (\alpha_0(\tau) - \beta(\tau) C(\omega, \tau)) A(\tau) d\tau,$$

where $A(\tau) \equiv [A_r(\tau), A_\eta(\tau)]^\top$, $\iota \equiv [1, 0]^\top$, $\mathbf{1} \equiv [1, 1]^\top$ and:

$$K \equiv \begin{bmatrix} \kappa_r & -\kappa_r \\ 0 & \kappa_\eta \end{bmatrix}, \quad \Sigma \equiv \begin{bmatrix} \sigma_r & 0 \\ 0 & \sigma_\eta \end{bmatrix}.$$

Proposition 1.5 *The functions $(A_r(\tau), A_\eta(\tau))$ are given by:*

$$A_r(\tau) = \phi_r \left(\frac{1 - e^{-\tau v_2}}{v_2} - \frac{1 - e^{-\tau v_1}}{v_1} \right) + \frac{1 - e^{-\tau v_1}}{v_1},$$

$$A_\eta(\tau) = \phi_\eta \left(\frac{1 - e^{-\tau v_2}}{v_2} - \frac{1 - e^{-\tau v_1}}{v_1} \right),$$

where (v_1, v_2) are the eigenvalues of Ψ , and (ϕ_r, ϕ_η) are scalars computed from the elements of the

eigenvectors. The function $C(\tau)$ is given by:

$$C(\tau) = \left(\int_0^\tau A(u)^\top du \right) \psi - \frac{1}{2} \int_0^\tau A(u)^\top \Sigma \Sigma^\top A(u) du$$

Next, I explain how to derive the values of Ψ and ψ . Under the assumption that the function $\beta(\tau)$ is an exponential, or a linear combination of exponentials, the integrals in Ψ become Laplace transforms of the functions $(A_r(\tau), A_\eta(\tau))$ and of these functions' pairwise products. Moreover, by multiplying the ODE system by the exponentials in $\beta(\tau)$ and by the product of these exponentials with the functions $(A_r(\tau), A_\eta(\tau))$, I find equations that involve the same Laplace transforms. This yields a system of equations in the Laplace transforms. While the system remains non-linear, the Laplace transform approach has the advantage that it does not require one to compute the eigenvalues and eigenvectors of Ψ , which can be real or complex.

Thus, I next derive the system of equations in the Laplace transforms. I assume $\beta(\tau) = \beta e^{-\tau \delta_\beta}$ where (β, δ_β) are positive scalars, and I define:

$$\ell \equiv \begin{bmatrix} \ell_r \\ \ell_\eta \end{bmatrix} \equiv \int_0^\infty \beta(\tau) A(\tau) d\tau, \quad L \equiv \begin{bmatrix} \ell_{r,r} & \ell_{r,\eta} \\ \ell_{r,\eta} & \ell_{\eta,\eta} \end{bmatrix} \equiv \int_0^\infty \beta(\tau) A(\tau) A(\tau)^\top d\tau.$$

Given that the vector ℓ has dimension 2×1 , and that the matrix L has dimension 2×2 and is symmetric, there are a total of 5 distinct elements in the set (ℓ, L) . These elements are Laplace transforms of the functions $(A_r(\tau), A_\eta(\tau))$ and of those functions' pairwise products. Moreover, note that, using L , I can write the matrix Ψ as:

$$\Psi = \begin{bmatrix} \kappa_r + a\sigma_r^2 \ell_{r,r} & a\sigma_\eta^2 \ell_{r,\eta} \\ -\kappa_r + a\sigma_r^2 \ell_{r,\eta} & \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta} \end{bmatrix} = K^\top + aL\Sigma\Sigma^\top.$$

Lemma 1.1 Suppose that $\beta(\tau) = \beta e^{-\tau \delta_\beta}$ where (β, δ_β) are positive scalars. The 5 elements of (ℓ, L) solve the system given by the vector equation $(\delta_\beta \mathbf{I}_2 + \Psi) \ell = \frac{\beta}{\delta_\beta} \mathbf{1}$, which yields the two scalar equations:

$$(\delta_\beta + \kappa_r + a\sigma_r^2 \ell_{r,r}) \ell_r + a\sigma_\eta^2 \ell_{r,\eta} \ell_\eta = \frac{\beta}{\delta_\beta},$$

$$(\delta_\beta + \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta}) \ell_\eta - (\kappa_r - a\sigma_r^2 \ell_{r,\eta}) \ell_r = 0,$$

and the matrix equation $(\delta_\beta \mathbf{I}_2 + \mathbf{\Psi}) \mathbf{L} + \mathbf{L} \mathbf{\Psi}^\top = \mathbf{\iota} \mathbf{\iota}^\top + \mathbf{\ell} \mathbf{\ell}^\top$, which yields the three scalar equations:

$$\left(\frac{\delta_\beta}{2} + \kappa_r + a\sigma_r^2 \ell_{r,r} \right) \ell_{r,r} + a\sigma_\eta^2 \ell_{r,\eta}^2 = \ell_r,$$

$$\left(\delta_\beta + \kappa_r + \kappa_\eta + a\sigma_r^2 \ell_{r,r} + a\sigma_\eta^2 \ell_{\eta,\eta} \right) \ell_{r,\eta} - \left(\kappa_r - a\sigma_r^2 \ell_{r,\eta} \right) \ell_{r,r} + a\sigma_\eta^2 \ell_{r,\eta} \ell_{\eta,\eta} = \ell_\eta,$$

$$\left(\frac{\delta_\beta}{2} + \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta} \right) \ell_{\eta,\eta} - \left(\kappa_r - a\sigma_r^2 \ell_{r,\eta} \right) \ell_{r,\eta} = 0.$$

These constitute a system of 5 equations in the 5 unknowns $(\ell_r, \ell_\eta, \ell_{r,r}, \ell_{\eta,\eta}, \ell_{r,\eta})$. The first system is linear in (ℓ_r, ℓ_η) and its solution is:

$$\ell_r = \frac{\frac{\beta}{\delta_\beta} \left(\delta_\beta + \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta} \right)}{(\delta_\beta + \kappa_r + a\sigma_r^2 \ell_{r,r}) \left(\delta_\beta + \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta} \right) + (\kappa_r - a\sigma_r^2 \ell_{r,\eta}) a\sigma_\eta^2 \ell_{r,\eta}},$$

$$\ell_\eta = \frac{\frac{\beta}{\delta_\beta} (\kappa_r - a\sigma_r^2 \ell_{r,\eta})}{(\delta_\beta + \kappa_r + a\sigma_r^2 \ell_{r,r}) \left(\delta_\beta + \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta} \right) + (\kappa_r - a\sigma_r^2 \ell_{r,\eta}) a\sigma_\eta^2 \ell_{r,\eta}}.$$

Substituting back ℓ_r and ℓ_η , I obtain a system of 3 equations in the 3 unknowns $(\ell_{r,r}, \ell_{\eta,\eta}, \ell_{r,\eta})$.

Given a solution to the system, I can determine (ℓ_r, ℓ_η) .

Lemma 1.2 *The elements of $\mathbf{\Psi}$ are provided in Appendix A.1.7.*

Theorem 1.1 *Suppose that $\beta(\tau) = \beta e^{-\tau \delta_\beta}$ where (β, δ_β) are positive scalars. Then an equilibrium exists if $\kappa_r^2 \sigma_\eta^2 \delta_\beta (\delta_\beta + \kappa_r) < a\sigma_r^4 \beta$. Moreover, in equilibrium $\Psi_{r,r} > \kappa_r$, $0 < \Psi_{r,\eta} < \frac{\kappa_r \sigma_\eta^2}{\sigma_r^2}$, $-\kappa_r < \Psi_{\eta,r} < 0$, $\Psi_{\eta,\eta} > \kappa_\eta$.*

I complement the equilibrium existence result by computing the equilibrium in closed form when the arbitrageurs' risk-aversion coefficient is close to zero, and the other parameters can take any values. Thus, I next derive the asymptotic behavior of $(v_1, v_2, \phi_r, \phi_\eta)$ when $a \approx 0$.

Corollary 1.1 *Suppose that $\beta(\tau) = \beta e^{-\tau \delta_\beta}$ where (β, δ_β) are positive scalars. Then the asymptotic*

behavior of $(v_1, v_2, \phi_r, \phi_\eta)$ when $a \approx 0$ is $(v_1, v_2, \phi_r, \phi_\eta) = (\kappa_r, \kappa_\eta, a_{\underline{c}_r}, \underline{c}_\eta)$, where:

$$\underline{c}_r = -\frac{\kappa_r^2 \sigma_\eta^2}{(\kappa_r - \kappa_\eta)^3} \left(\int_0^\infty \beta(\tau) \frac{1 - e^{-\tau \kappa_r}}{\kappa_r} \frac{1 - e^{-\tau \kappa_\eta}}{\kappa_\eta} d\tau - \int_0^\infty \beta(\tau) \left(\frac{1 - e^{-\tau \kappa_r}}{\kappa_r} \right)^2 d\tau \right),$$

$$\underline{c}_\eta = \frac{\kappa_r}{\kappa_r - \kappa_\eta}.$$

1.4.1 Interest Rate Policy

When the target short-rate is stochastic, arbitrageurs face an additional source of risk when they transmit short-rate shocks to bond yields. This weakens transmission because the carry trades that link the current short-rate to longer maturities become riskier. In response, arbitrageurs reduce the size of these trades, and, when hedging motives are strong enough, can reshape them into butterfly trades by taking the opposite position at long maturities. As a result, a shock to the current short-rate can move long-maturity yields in the opposite direction from the shock itself. To see the mechanism, suppose that a negative shock lowers the current short-rate below the value that the yield of the τ -year bond would take in the segmentation equilibrium. As in the deterministic-target case, arbitrageurs have an incentive to borrow at the current short-rate and buy the τ -year bond. This position gains from the gap between the current short-rate and the bond yield, but it also loses value if either the current short-rate or the target short-rate increases. The relative importance of target short-rate risk grows with maturity. This is already suggested by the deterministic-target model: the effect of current-short-rate shocks on yields declines with maturity, whereas the effect of target-short-rate shocks is hump-shaped. Since long-maturity bonds are relatively more exposed to target short-rate risk, arbitrageurs may short-sell them to hedge the target short-rate exposure created by their overall portfolio. When arbitrageurs are sufficiently risk-averse, this hedging demand can be strong enough to raise long-maturity yields even after a fall in the current short-rate and in expected future current short-rates. The resulting position, long at intermediate maturities and short at very short (borrowing) and long maturities, is a butterfly trade.

I now quantify the effects of short-rate shocks on bond yields. The following proposition assumes $-\kappa_r < \Psi_{\eta,r} < 0$, which holds in the equilibrium in Theorem 1.1. The other assumptions in the theorem are not needed as long as this property holds.

Proposition 1.6 Suppose $-\kappa_r < \Psi_{\eta,r} < 0$:

- If the two eigenvalues of Ψ are real, then $A_\eta(\tau) > 0$, $\partial_\tau A_\eta(\tau) > 0$, and $\partial_\tau \frac{A_r(\tau)}{A_\eta(\tau)} < 0$. Moreover, $A_r(\tau) > 0$ for $\tau \in (0, \bar{\tau})$ and $A_r(\tau) < 0$ for $\tau \in (\bar{\tau}, \infty)$, where $\bar{\tau} = \infty$ when $a \approx 0$ or $\beta(\tau) = 0$, and $\bar{\tau} < \infty$ when $a \approx \infty$.
- If the two eigenvalues of Ψ are complex, then $A_\eta(\tau) > 0$ for τ in a maximal interval $(0, \hat{\tau}_2)$, $\partial_\tau A_\eta(\tau) > 0$ for τ in a maximal interval $(0, \hat{\tau}_1)$, and $\partial_\tau \frac{A_r(\tau)}{A_\eta(\tau)} < 0$ for $\tau \in (0, \hat{\tau}_1)$, where $\hat{\tau}_2 > \hat{\tau}_1 > 0$. Moreover, if $\hat{\tau}_2 < \infty$, then $A_r(\tau) > 0$ for τ in a maximal interval $(0, \bar{\tau})$, where $\bar{\tau} \in (0, \hat{\tau}_2)$.

The proposition separates the behaviour of the sensitivity functions according to whether the eigenvalues of Ψ are real or complex. When the eigenvalues are real, the maturity profile is straightforward. The target-short-rate sensitivity $A_\eta(\tau)$ is positive and increasing in maturity, so a higher target short-rate raises yields at every maturity, with a larger effect farther out on the curve. The current-short-rate sensitivity $A_r(\tau)$ is positive at short maturities but can become negative after the threshold $\bar{\tau}$. In that case, an increase in the current short-rate raises short-maturity yields but lowers yields for maturities above $\bar{\tau}$. The exception is the limiting case in which arbitrageurs are risk-neutral or preferred-habitat supply is price-inelastic, in which case $A_r(\tau)$ need not turn negative. Finally, the ratio $\frac{A_r(\tau)}{A_\eta(\tau)}$ falls with maturity, showing that target-short-rate shocks become relatively more important than current-short-rate shocks at longer maturities. This is the formal counterpart of the hedging logic described above.

When the eigenvalues of Ψ are complex, the same forces are present but the maturity profile can oscillate. A positive shock to the current short-rate still raises yields at short maturities. At longer maturities, however, the yield response can turn negative because arbitrageurs may buy long-maturity bonds as hedges against target short-rate risk. These bonds then provide insurance against target-short-rate risk in the arbitrageurs' equilibrium portfolio and can carry negative expected excess returns. This sign reversal is also present in the real-eigenvalue case when $A_r(\tau)$ eventually becomes negative. What is specific to the complex-eigenvalue case is the behaviour of very long-maturity bonds, which generates the oscillation. When the target short-rate is low, arbitrageurs hold short bond positions in long-maturity bonds (hedging); through the accumulation of negative expected returns, this can raise the prices of bonds with even longer maturities. At those very long maturities, bonds no longer serve as hedges against target short-rate risk, so their prices can increase after a fall in the current short-rate. The response of bond prices to the current short-rate therefore alternates across maturities. Over

the first half-cycle of the oscillation, which may extend beyond the maximum traded maturity T , the real-eigenvalue intuition continues to apply: both $A_r(\tau)$ and $A_\eta(\tau)$ initially increase with maturity, $A_r(\tau)$ reaches its maximum first, and $A_\eta(\tau)$ reaches its maximum at the larger maturity.

1.4.2 Risk Premia

Target short-rate risk can reinforce the positive comovement between bond risk premia and the slope of the term structure under restrictions on the relative factor loadings of r_t and η_t . This qualification matters because the second factor is not a stochastic preferred-habitat demand or supply shifter, as in [Vayanos and Vila \(2021b\)](#). A high target short-rate raises expected future current short-rates and therefore shifts the expectations component of yields upward, but it does not mechanically force arbitrageurs to absorb more bonds. With deterministic but price-elastic supply, the target short-rate affects arbitrageurs' positions only indirectly through bond prices. Hence the premium component need not move in the same direction as the slope of the term structure unless the sensitivity functions satisfy the loading restrictions stated below.

Proposition 1.7 Suppose $0 < \Psi_{r,\eta} < \frac{\kappa_r \sigma_\eta^2}{\sigma_r^2}$ and $\Delta\tau \rightarrow 0$:

- The FB coefficient is positive for $\tau < \min\{\bar{\tau}, \hat{\tau}_1\}$ if:

$$-\frac{\partial_\tau A_\eta(\tau)}{\partial_\tau A_r(\tau) - 1} < \frac{\kappa_r}{\kappa_r + \kappa_\eta}.$$

When $a \approx 0$, the FB coefficient is positive for τ such that:

$$\frac{e^{-\tau\kappa_\eta} - e^{-\tau\kappa_r}}{1 - e^{-\tau\kappa_r}} < \frac{\kappa_r - \kappa_\eta}{\kappa_r + \kappa_\eta}.$$

- The CS coefficient is smaller than 1 for $\tau < \min\{\bar{\tau}, \hat{\tau}_1\}$ if:

$$-\frac{A_\eta(\tau)}{A_r(\tau) - \tau} < \frac{\kappa_r}{\kappa_r + \kappa_\eta}.$$

When $a \approx 0$, the CS coefficient is smaller than 1 for τ such that:

$$\frac{\frac{1-e^{-\tau\kappa_\eta}}{\tau\kappa_\eta} - \frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r}}{1 - \frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r}} < \frac{\kappa_r - \kappa_\eta}{\kappa_r + \kappa_\eta}.$$

The proposition therefore gives a conditional premia-slope result rather than an unconditional

strengthening result. It shows that $b_{FB} > 0$ and $b_{CS} < 1$ for maturities in the range where the current- and target-short-rate loadings have the appropriate signs, provided that specific inequality conditions on the bond sensitivity functions and their derivatives are met. Similar conditions are provided for the asymptotic case of $a \approx 0$.

1.5 Calibration

The equilibrium term structure is determined by the risk-aversion coefficient a of arbitrageurs, the parameters (κ_r, σ_r) of the current short-rate, the parameters $(r^*, \kappa_\eta, \sigma_\eta)$ of the target short-rate, and the functions $(\alpha_0(\tau), \beta(\tau))$ that describe the supply intercept and supply slope of preferred-habitat investors. I divide my calibration into two parts. In the first part, I calibrate the parameters of the target short-rate, tying the diffusion coefficient to a measure of monetary policy disagreement I construct. In the second part, I calibrate the parameters of the current short-rate and the supply of bonds coming from preferred-habitat investors, given the parameters of the target short-rate. Throughout the calibration, I work with nominal data and I express yields and their moments in percentage terms.

1.5.1 Target Short Rate and Monetary Policy Disagreement

I calibrate the parameters $(\kappa_\eta, \sigma_\eta)$ using the dot plot data, extracted from the Summary of Economic Projections (SEP) released in conjunction with each Federal Open Market Committee (FOMC) meeting. The data has been released at a quarterly frequency from the first quarter of 2012 to the present. The first quarter of 2020 is the only exception, when data was not released due to the Covid-19 pandemic. Each dot plot provides the FOMC participants' assessment of the appropriate level of the federal funds rate⁵ at the end of a specified calendar year, where "appropriate monetary policy" is defined as the future path of policy that each participant deems most likely to achieve the employment and price stability mandates, given the information available at the time of the meeting. The plots are usually made up of 19 dots – the 7 members of the Board of Governors and the 12 presidents of the Federal Reserve Banks – for each projection year, but on rare occasions the number is as low as 15. Moreover, in the SEP releases of the first and second quarters of the year, FOMC participants provide three projections: end of the current year Y , end of year $Y + 1$, and end of year $Y + 2$. In the SEP releases of

⁵More precisely, the dots represent the value (rounded to the nearest 1/8 percentage point) of the midpoint of the appropriate target range for the federal funds rate or the appropriate target level for the federal funds rate.

the third and fourth quarters of the year, they provide a fourth projection for end of year $Y + 3$. The main challenge when working with this data is therefore that the time interval between the moment a projection is made and the target year varies significantly⁶. Figure 1.6 shows an example of a dot plot: each column corresponds to a projection year, each dot represents one FOMC participant's projected federal funds rate for that year, and the vertical spread within a column captures disagreement across participants.

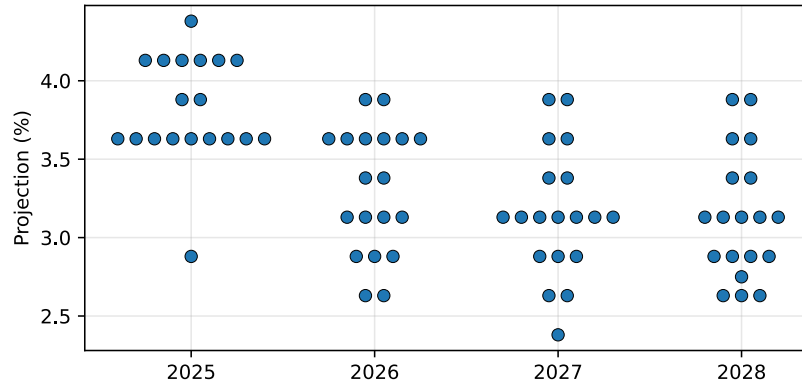


Figure 1.6: Dot plot released in conjunction with the FOMC meeting that took place on 09-17-2025.

First, I build time series for the median value and the cross-sectional standard deviation of the 1-year-ahead projection. To retain a quarterly frequency, while accounting for the variable length of the projection horizon, I use a mixture of densities approach. In particular, for an FOMC meeting taking place on day t in calendar year Y , let w_t be the percentage of the year that has elapsed, and $m_t(j)$ be the probability mass function of the projections for end of year j . Then, for each FOMC meeting, I form the convex mixture:

$$m_{t,1y} = (1 - w_t)m_t(Y) + w_t m_t(Y + 1),$$

and I compute the median and the cross-sectional standard deviation from this distribution. Figure 1.7 shows the resulting time series.

⁶Consider the FOMC meeting taking place in March of year Y . There are three projections for end of year Y (9 months into the future), $Y + 1$ (9+12m) and $Y + 2$ (9+24m). Next, consider the FOMC meeting taking place in June of year Y . There are three projections for end of year Y (6 months into the future), $Y + 1$ (6+12m) and $Y + 2$ (6+24m). Next, consider the FOMC meeting taking place in September of year Y . There are four projections for end of year Y (3 months into the future), $Y + 1$ (3+12m), $Y + 2$ (3+24m) and $Y + 3$ (3+36m). Finally, consider the FOMC meeting taking place in December of year Y . There are four projections for end of year Y (0 months into the future – this is no longer a projection but rather the perceived optimal level of the federal funds rate), $Y + 1$ (12m), $Y + 2$ (24m) and $Y + 3$ (36m).

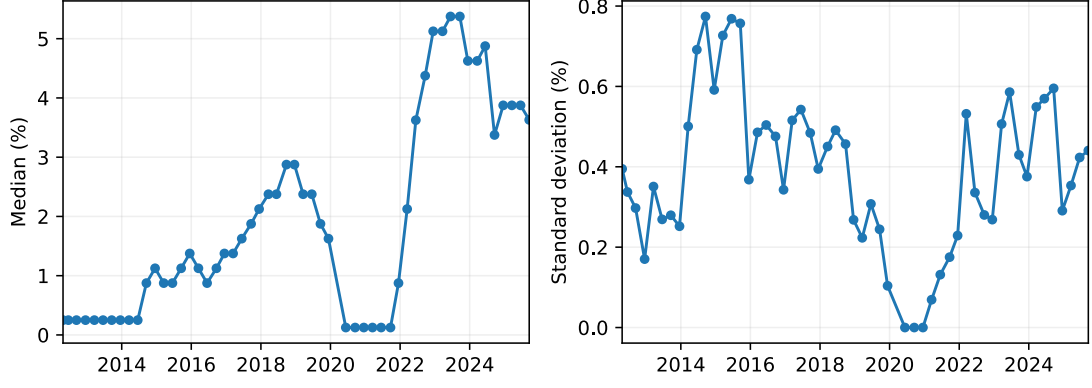


Figure 1.7: Time series of (a) the median value and (b) the cross-sectional standard deviation of the 1-year-ahead projection.

I take the median value of the 1-year-ahead projections as a proxy for the target short rate η_t . More precisely, let $\{t_i\}_{i=0}^N$ denote the FOMC meeting dates and $\Delta_i \equiv t_i - t_{i-1}$ the irregular time interval between consecutive meetings. I treat the computed median at t_i as η_{t_i} . Given that η_t follows an Ornstein-Uhlenbeck process, the exact discrete-time transition is:

$$\begin{aligned}\eta_{t_i} | \eta_{t_{i-1}} &\sim \mathcal{N}(\mu_i, v_i), \\ \mu_i &\equiv e^{-\Delta_i \kappa_\eta} \eta_{t_{i-1}} + (1 - e^{-\Delta_i \kappa_\eta}) r^*, \\ v_i &\equiv \frac{\sigma_\eta^2}{2\kappa_\eta} (1 - e^{-2\Delta_i \kappa_\eta}).\end{aligned}$$

For a fixed value of κ_η , maximizing the log-likelihood over r^* and σ_η^2 has a closed-form solution. In particular, given a candidate value of κ_η , the log-likelihood of $\{\eta_{t_i}\}_{i=0}^N$ is:

$$\ell(r^*, \sigma_\eta^2; \kappa_\eta) = -\frac{1}{2} \sum_{i=1}^N \left(\ln(2\pi v_i) + \frac{(\eta_{t_i} - \mu_i)^2}{v_i} \right).$$

Taking the first-order condition with respect to r^* :

$$\hat{r}^*(\kappa_\eta) = \frac{\sum_{i=1}^N \frac{(\eta_{t_i} - e^{-\Delta_i \kappa_\eta} \eta_{t_{i-1}})(1 - e^{-\Delta_i \kappa_\eta})}{1 - e^{-2\Delta_i \kappa_\eta}}}{\sum_{i=1}^N \frac{(1 - e^{-\Delta_i \kappa_\eta})^2}{1 - e^{-2\Delta_i \kappa_\eta}}}.$$

Substituting $\hat{r}^*(\kappa_\eta)$ back into the log-likelihood and taking the first-order condition with respect

to σ_η :

$$\hat{\sigma}_\eta^2(\kappa_\eta) = \frac{1}{N} \sum_{i=1}^N \frac{2\kappa_\eta (\eta_{t_i} - e^{-\Delta_i \kappa_\eta} \eta_{t_{i-1}} - (1 - e^{-\Delta_i \kappa_\eta}) r^*(\kappa_\eta))^2}{1 - e^{-2\Delta_i \kappa_\eta}}.$$

I then substitute both $\hat{r}^*(\kappa_\eta)$ and $\hat{\sigma}_\eta^2(\kappa_\eta)$ back into the log-likelihood and I maximize $\ell(\kappa_\eta)$ over $\kappa_\eta > 0$ using a coarse grid in $\log \kappa_\eta$ to bracket the optimum, followed by one-dimensional Brent refinement in $\log \kappa_\eta$. This delivers κ_η .

To discipline the magnitude of monetary policy disagreement, I use the standard deviation of the 1-year-ahead projections. More precisely, let s_{t_i} denote the cross-sectional standard deviation of the participants' 1-year-ahead projections at t_i , and let $\bar{s}^2 \equiv \frac{1}{N} \sum_{i=1}^N s_{t_i}^2$ be its time average. Under the Ornstein-Uhlenbeck dynamics, the conditional variance of η_{t+1} is:

$$\mathbb{V}\text{ar}_t(\eta_{t+1}) = \frac{\sigma_\eta^2}{2\kappa_\eta} (1 - e^{-2\kappa_\eta}).$$

Equating the model-implied forecast dispersion to the average cross-sectional dispersion from the dot plots yields the closed-form estimator:

$$\sigma_\eta = \sqrt{\frac{2\kappa_\eta \bar{s}^2}{1 - e^{-2\kappa_\eta}}}.$$

Note that, while this approach delivers an estimate of r^* , I calibrate the latter in the next section using yield data, as they allow me to use a longer range of data and obtain a more reliable estimate. Table 1.1 reports the calibrated parameters and the empirical time series used to estimate them.

Parameter	Value	Series
κ_η Speed of mean reversion of η_t	0.159	Median value of 1-year-ahead projections
σ_η Diffusion coefficient of η_t	0.005	Standard deviation of 1-year-ahead projections

Table 1.1: Calibration of the parameters of the target short rate.

1.5.2 Current Short Rate and Bond Supply

I set $\alpha_0(\tau) = \alpha_0 (e^{-\tau\delta_\beta} - e^{-\tau\delta_\alpha})$ and $\beta(\tau) = \beta e^{-\tau\delta_\beta}$ for $\tau \leq T$, and $\alpha_0(\tau) = \beta(\tau) = 0$ for $\tau > T$, where I set $T = 30$ years, which is the maximum maturity of U.S. Treasury bonds. The parameter

δ_β captures the front-end rate of decay of the supply slope: a larger value makes the supply elasticity much more concentrated at short maturities, while a smaller value spreads it toward the long end. The parameter δ_α captures the back-end rate of decay of the supply intercept: for $\delta_\alpha > \delta_\beta$, the intercept $\alpha_0(\tau)$ is hump-shaped with a positive bump. The difference $\delta_\alpha - \delta_\beta$ controls how localized the bump is: a larger difference makes the bump narrower and more front-loaded.

I calibrate the parameters $(a, r^*, \kappa_r, \sigma_r, \beta, \delta_\beta, \alpha_0, \delta_\alpha)$ using data on U.S. Treasury yields and trading volume, and estimates of demand elasticity from the literature. Bond yields are from [Gürkaynak et al. \(2007\)](#)'s data set, which reports U.S. Treasury yields at a daily frequency from June 1961 to the present. I use two different samples in the calibration: for (κ_r, σ_r) , I set the starting date in November 1985, because it is the first month in which all maturities from 1 to 30 years are included; for $(a\beta, a\alpha_0, \delta_\alpha, r^*)$, I use the longest period available. In both cases, I re-sample yields at the end of each month. The bond trading volume is from the Federal Reserve 2004 data set, which reports the U.S. Treasury volume by primary dealers, divided into buckets based on the bonds' remaining time to maturity, at a daily frequency from April 2013 to the present. The supply elasticity is from [Krishnamurthy and Vissing-Jorgensen \(2012; KVJ\)](#).

I determine the first seven parameters in the table by equating the seven empirical moments to their model-generated counterparts. This requires solving a non-linear system of seven equations. The moments are related to the volatilities of yields and yield changes, average yields, excess returns, and the fraction of trading volume in a specific maturity segment. Note that the data I use in the calibration does not allow me to identify the arbitrageurs' risk aversion coefficient a separately from the parameters β and α_0 that characterise the slope and intercept of the supply curve; only the products $a\beta$ and $a\alpha_0$ can be identified. Therefore, I determine the eighth parameter β using KVJ's estimates, and I back out a and α_0 from the products $a\beta$ and $a\alpha_0$. Table 1.2 reports the calibrated parameters and the empirical moments used to estimate them.

Parameter	Value	Moment	Value
κ_r	0.161	$\sqrt{\text{Var}(y_t(1))}$	2.564
Speed of mean reversion of r_t		Volatility 1-year yield – levels	
σ_r	0.017	$\sqrt{\text{Var}(y_{t+1}(1) - y_t(1))}$	1.372
Diffusion coefficient of r_t		Volatility 1-year yield – changes	
$a\beta$	44.513	$\mathbb{E}(p_{t+1}(6) - p_t(7) - y_t(1))$	1.358
Risk aversion $\times$ Supply slope		Excess return 7-year bond	
δ_β	0.281	$\frac{\sum_{0 < \tau \leq 2} \mathbb{E}(dS_t(\tau))}{\sum_{0 < \tau \leq 30} \mathbb{E}(dS_t(\tau))}$	0.200
Front-end decay of supply slope		Relative volume for $\tau \in (0, 2]$	
$a\alpha_0$	30.001	$\mathbb{E}(y_t(1))$	4.820
Risk aversion $\times$ Supply intercept		Average yield 1-year bond	
δ_α	0.475	$\mathbb{E}(y_t(5))$	5.426
Back-end decay of supply intercept		Average yield 5-year bond	
r^*	0.046	$\mathbb{E}(y_t(7))$	5.625
Natural rate of interest		Average yield 7-year bond	
β	4.670	Estimate in KVJ (2012)	0.746
Supply slope			

Table 1.2: Calibration of the current short rate and the supply of bonds.

Figure 1.8 shows the model-generated and empirical moments. The two panels on top show the volatility of yields in levels and annual changes, as a function of maturity. By construction, the model-generated moments and the empirical ones coincide for the 1-year maturity. Moreover, both curves are decreasing functions of maturity, albeit the latter decreases more slowly. The two panels in the middle show the excess return over the 1-year yield and the average yield, respectively, as a function of maturity. Again, by construction, the model-generated moments and the empirical ones coincide for the 7-year maturity and the 1-year, 5-year and 7-year maturities, respectively. Moreover, the model fit is remarkably close for all maturities. Finally, the panel at the bottom shows the relative volume traded in different maturity buckets. By construction, the model-generated moments and the empirical ones coincide for maturities below 2 years. Moreover, the model fit is close for all other maturity buckets.

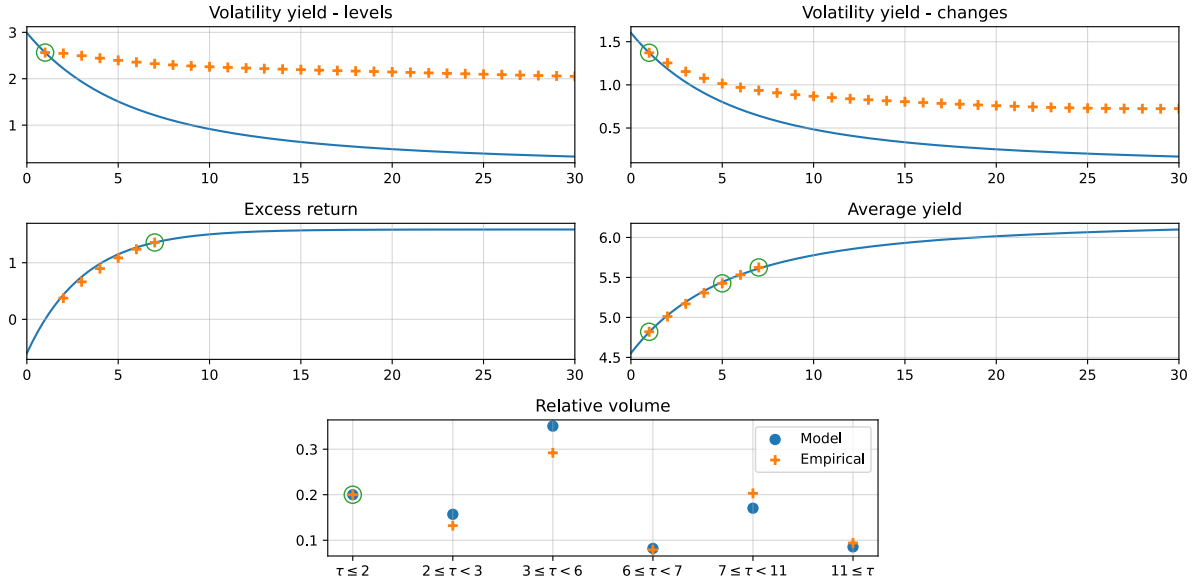


Figure 1.8: Model-generated and empirical moments.

In order to estimate the slope parameter β of the supply function, I use KVJ's estimates of the elasticity of demand of US Treasuries. They regress the yield spread between long-maturity AAA-rated corporate bonds and government bonds on the logarithm of government debt to GDP, and find a coefficient of -0.746. Hence, a 1bp decrease in the yield spread is associated with a $0.01/0.746 = 0.0134$ increase in the logarithm of debt to GDP. Assuming that debt to GDP is initially equal to the average value of 43.9% in KVJ's sample (1919-2008), this implies that the level increase in debt to GDP is $43.9\%(e^{0.0134} - 1) = 0.0059$. I map this estimate to my model by setting it equal to the change in the preferred-habitat supply of government bonds. Further, I assume that the decrease in the yield spread is due to an increase in government bond yields across all maturities. This yields the estimate of β in the table, which in turn implies $a = 9.530$ and $\alpha_0 = 3.148$.

Table 1.3 reports the elasticities of the seven model-generated moments in Table 1.2 with respect to the first seven parameters. In particular, let θ be a parameter of interest, $\bar{\theta}$ be the vector of the remaining parameters, and $m(\theta; \bar{\theta})$ be a model-generated moment. I compute the elasticities as the change $m(1.001\theta; \bar{\theta}) - m(0.999\theta; \bar{\theta})$, divided by $0.002m(\theta; \bar{\theta})$. The parameter κ_r has a strong (negative) impact on the volatility of the 1-year yield in levels, while the parameter σ_r has the strongest (positive) impact on the volatility of the 1-year yield in changes. The other parameters have a much weaker, or even economically insignificant, impact on these moments. Therefore, I conclude that the volatility of the 1-year yield in levels identifies κ_r and the volatility of the 1-year yield in changes identifies σ_r . The parameter $a\beta$ has a strong (negative) impact

on the excess return of the 7-year bond over the 1-year yield. The remaining parameters all have a strong impact on the excess return of the 7-year bond. The parameter δ_β has the second strongest (positive) impact on the relative volume for maturities below 2 years. The parameters $a\alpha_0$, δ_α and r^* have strong (positive) impact on the average yield of the 1-year, 5-year and 7-year bonds. The other parameters have a much weaker impact on these moments. Therefore, I conclude that the excess return identifies $a\beta$, the relative volume identifies δ_β , and the average yield identifies $(a\alpha_0, \delta_\alpha, r^*)$, although the latter identification is less clean.

	κ_r	σ_r	$a\beta$	δ_β	$a\alpha_0$	δ_α	r^*
Vol 1y	-0.512	0.862	-0.047	0.083	0.000	0.000	0.000
dVol 1y	-0.084	0.903	-0.047	0.082	0.000	0.000	0.000
Exc return 7y	-0.257	0.700	-2.375	-1.759	2.706	3.504	-1.728
Vol < 2y	0.103	0.220	0.108	0.921	0.000	0.000	0.000
Avg yield 1y	-0.002	0.063	-0.119	-0.120	0.150	0.195	0.850
Avg yield 5y	-0.020	0.153	-0.381	-0.344	0.462	0.599	0.542
Avg yield 7y	-0.029	0.165	-0.449	-0.391	0.538	0.697	0.468

Table 1.3: Elasticities of model-generated moments with respect to the model parameters.

1.6 Policy Analysis

1.6.1 Unconventional Monetary Policy

The first type of policy I analyse is a forward guidance announcement about the path of the target short-rate. I model this announcement as a change $\Delta g e^{-\kappa_g t}$ in the target short-rate, which represents the level to which the current short-rate mean-reverts. I assume that this shock is unanticipated, takes place at $t = 0$, and mean-reverts deterministically to zero at the rate κ_g . Figure 1.9 shows the effect of the announcement on the term structure at $t = 0$ for $\Delta g = -0.005$ (i.e. a 50 bp reduction) and $\kappa_g \in \{0.1, 0.2\}$, which correspond to half-lives of 6.93 and 3.47 years, respectively. This represents a scenario in which the central bank commits to keeping short-rates lower. The blue line represents the effect of the announcement, and the orange line represents the same effect when the arbitrageurs are assumed to be risk-neutral and the expectations hypothesis holds. First, note that when the speed of mean reversion of the shock is lower (i.e. the central bank commits to lower-for-longer) the impact on the term

structure is larger. This is because a more persistent shock has more time to propagate through the curve. Second, note how the yield curve underreacts compared to the risk-neutral setting, consistently with the result in Proposition 1.2. The magnitude of the underreaction is largest for intermediate maturities. Finally, the model implies that forward guidance is especially effective in changing the yields for intermediate-maturity bonds.

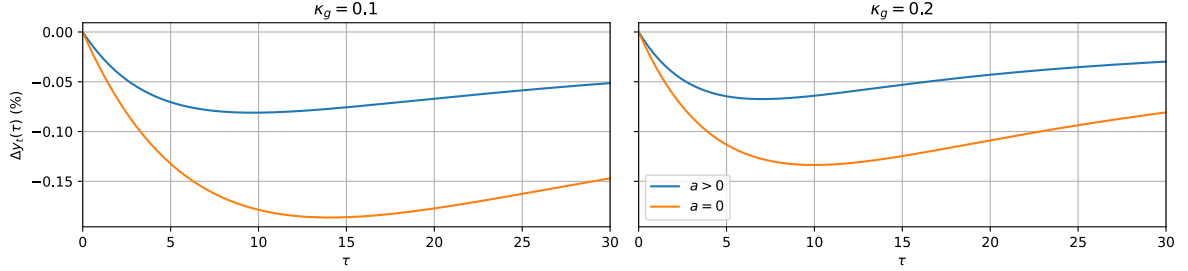


Figure 1.9: Effect of a forward-guidance announcement.

The second policy I analyse is a quantitative easing announcement. I model this announcement as a change $\Delta\alpha_0(\tau)e^{-\kappa_\alpha t}$ in the intercept of the supply by preferred-habitat investors. I assume that the shock is unanticipated, takes place at $t = 0$, and mean-reverts deterministically to zero at the rate κ_α . Figure 1.10 shows the effect of the announcement on the term structure at $t = 0$ for $\Delta\alpha_0(\tau) = -0.12$ (i.e. 12% of GDP) and $\kappa_\alpha \in \{0.1, 0.2\}$. This is approximately the size of the QE implemented by the Federal Reserve during QE1, QE2 and QE3. The blue, orange, green and red lines represent purchases of 2, 5, 10 and 30-year bonds, respectively. First, note that, like for forward guidance, when the speed of mean reversion of the shock is lower (i.e. quantitative easing lasts longer) the impact on the term structure is larger. This is because a bond's yield depends on an average over the bond's life of the risk premia required by arbitrageurs to hold the bond. QE lowers the yield of a bond by lowering these risk premia and, when it lasts longer, the risk premia in the average are impacted more. Second, note how purchases of longer-maturity bonds yield a larger reduction in yields for all maturities. In this sense, the model implies that QE is especially effective when implemented via purchases of long-maturity bonds. At the same time, for each purchase maturity, the largest yield response occurs at intermediate maturities. This is because the QE shock is temporary: long-maturity yields average risk premia over many future years, including years in which the supply shock has substantially mean-reverted, so the effect on very long yields is diluted relative to intermediate yields.

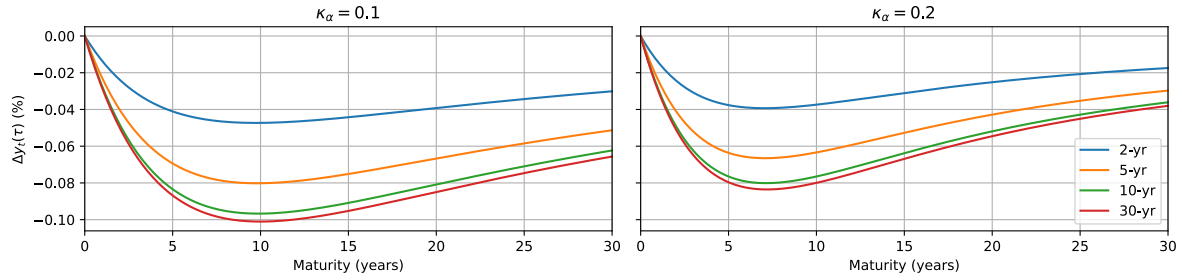


Figure 1.10: Effect of a quantitative easing announcement.

1.6.2 Effect of Monetary Policy Disagreement

Figure 1.11 shows how monetary policy disagreement affects the ability of the central bank to influence the term structure of interest rates using unconventional monetary policy instruments. In particular, I analyse a forward guidance announcement where $\Delta g = -0.005$ and $\kappa_g = 0.05$, and a QE announcement that purchases 20-year bonds, with $\kappa_\alpha = 0.05$. The size of the QE shock is set to $\Delta\alpha_0(20) = -0.096$, so that, in the baseline scenario, the forward guidance and QE announcements have the same largest yield reduction. For both policies, I compare the baseline scenario, where σ_η is equal to the calibrated value from Table 1.1, with a high monetary policy disagreement scenario where σ_η is four times as high. The blue line plots the baseline response, while the orange line plots the high-disagreement response; hence the effect of disagreement is measured by the vertical distance between the two lines. The key result of this paper is that, as monetary policy disagreement increases, forward guidance becomes less effective, while QE becomes more effective. The mechanism is the arbitrageurs' risk-bearing constraint. In the model, a higher σ_η means that the target short-rate is riskier, with the dispersion of the 1-year-ahead dot-plot projections disciplining the magnitude of this risk. A forward-guidance announcement changes the expected path of the target short-rate, but transmitting that change to bond yields requires arbitrageurs to hold positions whose payoffs are exposed to target short-rate risk. When disagreement is high, this exposure is more costly, especially at longer maturities, so arbitrageurs demand larger risk premia, scale back the trades that would otherwise pass the guidance shock through to yields, and may use long bonds partly as hedges. As a result, changes in expected future short rates translate less forcefully into movements in the yield curve. Quantitative easing operates through the same risk-bearing constraint in the opposite direction. By reducing the supply of duration risk that arbitrageurs must absorb, QE lowers the risk premia embedded in bond prices. Since those premia are larger

when monetary policy disagreement is high, the same purchase program removes a larger risk-premium component from yields. The implication is that the higher the level of monetary policy disagreement, the more the central bank should lean toward quantitative easing over forward guidance when aiming to influence the term structure of interest rates.

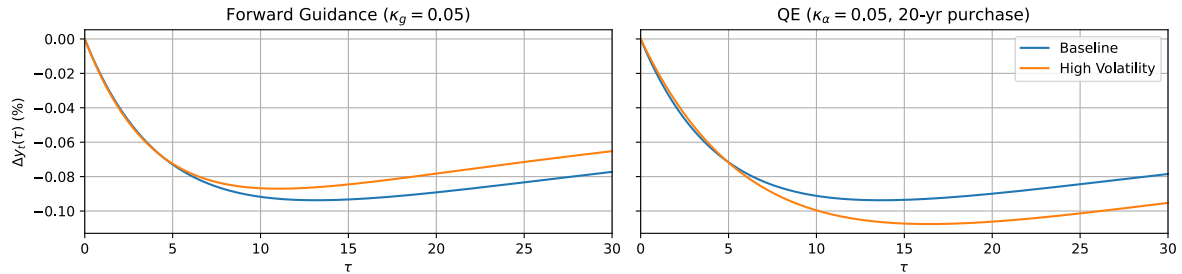


Figure 1.11: Impact of disagreement on the effectiveness of forward-guidance and quantitative easing.

Figure 1.12 shows how the effect of monetary policy disagreement depends on the design of the policy intervention. In this figure, each line plots the difference between the high-disagreement response and the baseline response, rather than the two responses separately. Given the easing policy scenario considered here, a reduction in effectiveness corresponds to a positive value, and an increase in effectiveness corresponds to a negative value. The left panels consider forward guidance. In the top-left panel, I vary the size of the announcement, keeping $\kappa_g = 0.05$. Since the model is linear in the size of the shock, the difference between the high-disagreement and baseline scenarios scales almost one-for-one with Δg : the largest gap is about 0.6 bp for a 25 bp announcement (the effectiveness of the policy is decreased by 15.4% compared to the baseline), 1.2 bp for a 50 bp announcement (15.4% reduction), and 1.8 bp for a 75 bp announcement (15.4% reduction). These maxima occur at long maturities, around 29 years. In the bottom-left panel, I keep $\Delta g = -0.005$ and vary the persistence of the announcement. The disagreement effect is stronger when forward guidance is more persistent: the largest gap is 1.2 bp when $\kappa_g = 0.05$ (15.4% reduction), compared with 0.8 bp when $\kappa_g = 0.1$ (13.6% reduction) and 0.5 bp when $\kappa_g = 0.2$ (11.5% reduction). Thus, monetary policy disagreement matters most for forward guidance when the announcement is intended to move expected short rates far into the future. The right panels consider QE. In the top-right panel, I vary the maturity of the bonds purchased, keeping $\kappa_\alpha = 0.05$. Purchases of short-maturity bonds are barely amplified by higher disagreement: for a 2-year purchase, the most negative gap is only -0.05 bp (1.6% increase). In contrast, the most negative gap is -1.2 bp for a 10-year purchase (14.3% increase), -1.7 bp for a 20-year purchase (20.4% increase), and -1.7 bp for a 30-year purchase

(21.0% increase). In the bottom-right panel, I keep the purchase maturity fixed at 20 years and vary the persistence of QE. The most negative gap is -1.7 bp when $\kappa_\alpha = 0.05$ (20.4% increase), -1.3 bp when $\kappa_\alpha = 0.1$ (19.8% increase), and -1.0 bp when $\kappa_\alpha = 0.2$ (19.2% increase). Therefore, disagreement increases the effectiveness of QE most when purchases remove long-maturity risk for a relatively persistent period.

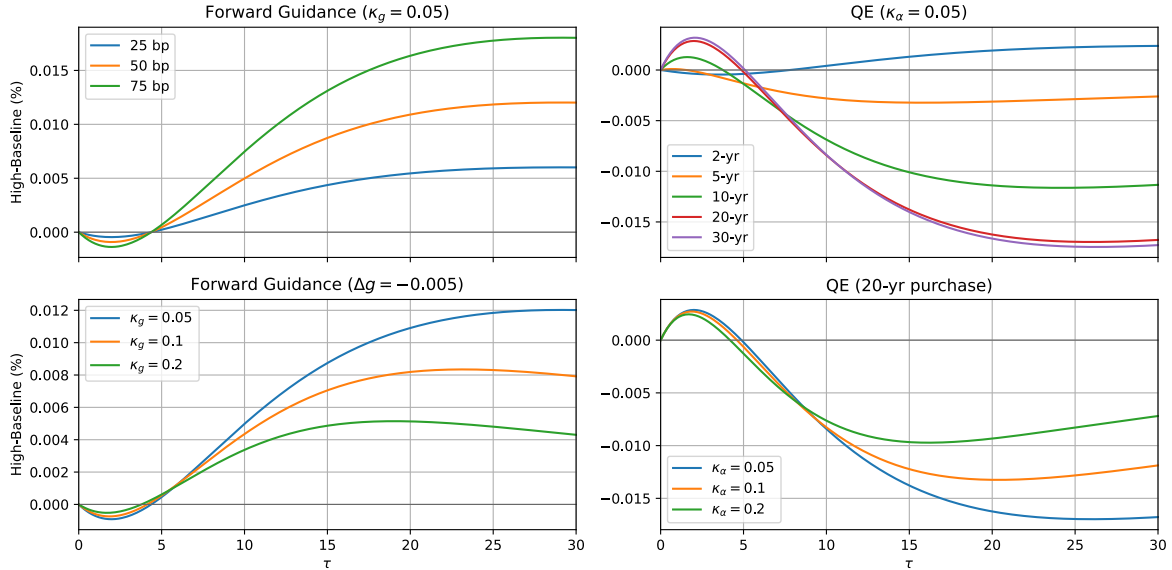


Figure 1.12: Impact of disagreement on the effectiveness of forward-guidance and quantitative easing for different policy implementations.

Figure 1.13 plots the gap between the baseline scenario and the high monetary policy disagreement scenario for selected maturities, as a function of a σ_η multiplier. Positive values in the forward-guidance panel mean that higher disagreement makes the yield reduction smaller, while negative values in the QE panel mean that higher disagreement makes the yield reduction larger. The maturity profile is central. When σ_η is four times the calibrated value, the forward-guidance gap is essentially zero at short maturities, equal to -0.09 bp at 2 years and 0.05 bp at 5 years, but rises to 0.50 bp at 10 years and 1.20 bp at 30 years. In the same high-disagreement scenario, the QE gap is positive and small at 2 years, 0.29 bp, approximately zero at 5 years, -0.02 bp, and then becomes substantially negative at longer maturities, reaching -0.84 bp at 10 years and -1.68 bp at 30 years. The sensitivity also increases with the level of disagreement. At the 30-year maturity, moving from the calibrated σ_η to twice, three times, and four times that value raises the forward-guidance gap from 0 to 0.29 bp, 0.71 bp, and 1.20 bp, respectively. For QE, the corresponding gaps are 0, -0.44 bp, -1.04 bp, and -1.68 bp. The implication is that the tilt in favor of quantitative easing should be even stronger when aiming to influence the long

end of the curve.

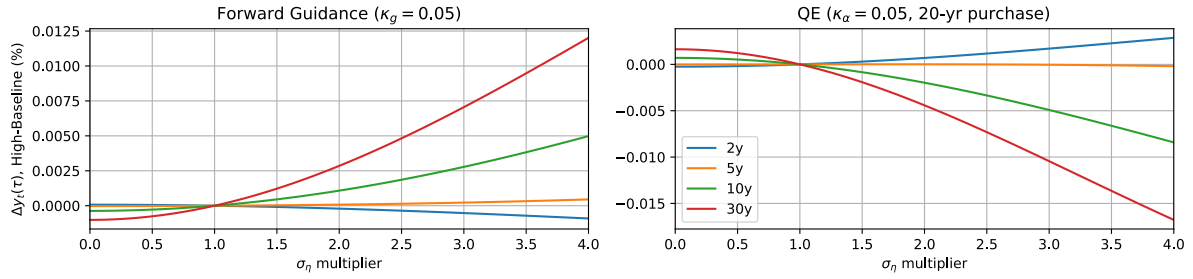


Figure 1.13: Effect of forward-guidance and quantitative easing under different levels of monetary policy disagreement at different maturities.

1.7 Conclusion

In this paper, I introduce a preferred-habitat model of the term structure of interest rates à la [Vayanos and Vila \(2021b\)](#). The model features two types of agents: preferred-habitat investors, who have preferences for bonds with specific maturities, and risk-averse arbitrageurs, who absorb the net supply of bonds. The dynamics of the term structure are driven by two risk factors: the current short-rate, which is the short rate observed in the market and captures the central bank’s interest-rate policy, and the target short-rate, which is the central bank’s intermediate policy target and captures its forward-guidance policy. I impose the supply of bonds exogenously, which captures the central bank’s quantitative-easing policy.

I use the model to study how changes to interest-rate policy are transmitted along the curve, how bond risk premia depend on the shape of the term structure, how alternative monetary policy tools – namely forward guidance and quantitative easing – impact the term structure, and how monetary policy disagreement affects the effectiveness of these instruments. The key result of the paper is that higher monetary policy disagreement, as captured by the dispersion of the 1-year-ahead projections in the dot plot, reduces the effectiveness of forward guidance, while it increases the effectiveness of quantitative easing. This implies that, when disagreement is higher, a central bank seeking to move the term structure should rely relatively more on quantitative easing than on forward guidance, with this shift toward quantitative easing becoming stronger when the objective is to affect the long end of the curve. The mechanism is the risk-bearing constraint faced by arbitrageurs.

A possible extension of this work is to introduce a supply risk factor. In general, supply factors capture the demand of preferred-habitat investors in a particular segment of the yield

curve. For example, one such factor may capture the demand of mutual funds that shifts the supply of short-term bonds, while another may capture the demand of pension funds that shifts the supply of long-term bonds. Introducing one such factor would allow me to model forward guidance about the central bank's balance-sheet policy as in [Greenwood et al. \(2015a\)](#), while maintaining that the supply of bonds is price elastic. The model can be solved using the methodology of [Hayashi \(2018\)](#), who provides numerical algorithms to solve the [Vayanos and Vila \(2021b\)](#) model with an arbitrary number of risk factors. Another possible extension is to study the interaction between forward guidance and the zero lower bound, and the extent to which monetary policy disagreement lessens the effectiveness of forward guidance when it is needed the most.

Chapter 2

Market Failure, the QE Trap and Fiscal Escape

Harjoat Bhamra and Nicolo Ceneda

2.1 Introduction

The 2008 financial crisis depleted the capital of the financial institutions that bear much of the economy's aggregate risk, triggering unprecedented central bank asset purchases.¹ In the United States, the Federal Reserve expanded its balance sheet from under \$1 trillion to over \$4 trillion; the subsequent unwinding took a decade and remained incomplete when the COVID-19 pandemic necessitated further expansion. A second quantitative tightening began in mid-2022 but was halted in 2025 with the balance sheet still above \$6 trillion. Meanwhile, the Troubled Asset Relief Program (TARP) injected capital directly into the financial sector. Similar interventions followed in Europe, though with significant delays in bank recapitalisation (Brunnermeier et al., 2016). Two questions arise: when is QE necessary, and how does one escape the QE trap of permanent asset purchases and depressed output?

This paper departs from the intermediary asset pricing literature—including He and Krishnamurthy (2013, 2019) and Brunnermeier and Sannikov (2014), whose foundational contribution we build directly on—in one critical dimension: in those models, equilibrium exists with or without the intermediary net-worth constraints; the constraint distorts prices and amplifies shocks, but the market always clears. Here, below a threshold, no competitive equilibrium exists at all. The central bank's role is therefore not to improve upon a distorted outcome but to

¹Strictly, *quantitative easing* refers to expanding central bank reserves, *large-scale asset purchases* (LSAPs) to purchasing specific securities, and *credit easing* to targeting particular credit markets. In practice, and in academic usage, these terms are used interchangeably. Following common convention, we use “QE” throughout.

restore the *existence* of an outcome. This resolves the paradox of [Bernanke \(2014\)](#): QE “works in practice but not in theory” only in a literature that treats equilibrium as given; once the leverage constraint binds hard enough to destroy existence, QE is no longer a welfare choice but a necessary condition for any market outcome at all.

We answer both questions in a continuous-time production economy with monopolistically competitive firms, flexible prices, and endogenous labour supply—a real model without money or a zero lower bound. Two types of agents inhabit the economy: financiers, who hold equity and bear all aggregate risk, and households, who save only through risk-free bonds. Financiers face leverage constraints arising from moral hazard and prudential regulation—frictions that long predate any crisis. Because households supply labour, financial frictions affect employment and output directly. The entire economy is characterised by a single state variable: the financier’s share of total wealth, ω . When leverage constraints do not bind, the financier’s wealth share drifts toward the efficient benchmark and the economy functions well—explaining why crises are surprises, not inevitabilities. We establish four sets of results.

First, equilibrium failure. When a large shock depletes financier capital and forces leverage constraints to bind, competitive equilibrium ceases to exist. The household’s bond demand is perfectly inelastic, so once the constrained financier cannot absorb the entire equity supply, excess supply emerges that no price adjustment can eliminate.² This is not a recession or a market crash—the price system itself breaks down. Three ingredients interact: limited participation, leverage constraints, and depleted capital—and all three were present in 2008. In the moderately distressed region ($\omega \in (\underline{\omega}, \omega_c)$), all three are needed: removing any one restores equilibrium. In the severely distressed region ($\omega < \underline{\omega}$), the result is stronger: the leverage constraint forces the financier below full equity investment ($\bar{\phi} < 1$), and even full household equity participation cannot compensate—equilibrium fails regardless of the household’s portfolio.

Second, the central bank as risk-bearer of last resort. A central bank that purchases the equity the constrained financier cannot hold, financing these purchases by issuing bonds, restores market clearing and equilibrium existence. In every existing theory of QE, asset purchases improve upon a distorted but viable competitive outcome—through portfolio balance effects ([Vayanos and Vila, 2021a](#)), credit intermediation ([Gertler and Karadi, 2011](#)), liquidity provision ([Cui and Sterk, 2021](#)), or risk redistribution ([Silva, 2020](#)). Our result is different in kind: without central

²The equilibrium failure builds on the economic logic of [Bernanke \(1983\)](#)’s credit-channel insight: financial intermediaries embody specialised screening, monitoring, and risk-bearing capacity that households cannot replicate, so when intermediary capital is depleted, equity risk cannot simply migrate to households’ balance sheets.

bank purchases no competitive outcome exists at all. The central bank acts not as a lender of last resort or a market-maker of last resort, but as a *risk-bearer of last resort*.

Third, Permanent QE, the QE Trap, and fiscal escape. The restored equilibrium has two steady states. At the good steady state the financier is fully capitalised, no constraints bind, and the economy coincides with the frictionless benchmark. At the bad steady state the financier is minimally capitalised, the central bank permanently holds a large share of the equity market, and output is permanently depressed. If the crisis pushes ω below the bad steady state, QE is literally permanent (*Permanent QE*). Even above the bad steady state, the economy can be trapped: if the deterministic recovery time exceeds a policy-relevant horizon—centuries at typical post-crisis capital ratios—QE is permanent in practice though not in theory. We call this the *QE Trap*, extending a term introduced by Koo (2014); we provide the first general equilibrium characterisation with an analytical formula for the trap boundary. A one-time fiscal recapitalisation—a transfer large enough to push ω past the trap boundary into the self-resolving zone—is the only practical escape. The central bank prevents market failure; the fiscal authority breaks the trap.

Fourth, the leverage paradox. Without the leverage constraints imposed by moral hazard and prudential regulation, the market for aggregate risk always clears. However, equilibrium asset prices exceed fundamental values—a rational bubble in the sense of Hugonnier (2012)—creating an *intermediation gap* (the component of asset prices above fundamental value) that grows convexly as financier capital erodes. Because ours is a production economy with endogenous labour, the intermediation gap *is* the output gap: the gap-to-price ratio maps one-to-one into the shortfall in employment and output. We call the common cause of both phenomena the *risk-bearing gap*—the economy’s shortage of intermediary capacity to absorb aggregate risk. Constraining intermediary leverage eliminates the risk-bearing gap and hence both the intermediation gap in prices and the output gap in quantities. But when leverage constraints bind, they destroy equilibrium. The constraint that corrects prices breaks the market.

This tradeoff is inescapable: we prove that no economy in this class can simultaneously deliver zero intermediation and output gaps and unconditional market clearing. The mechanism is that $Q = 1/\omega$, the inverse financier wealth share, is a local martingale in all versions of the model; its classification as a strict local martingale (positive intermediation gap) or a true martingale (no intermediation gap) is governed by the same boundary behaviour of the intermediary wealth-share volatility that determines whether equilibrium can fail.

2.1.1 Related Literature

QE and unconventional monetary policy. The existing literature identifies several channels: portfolio balance effects on term premia (Gagnon et al., 2011; Krishnamurthy and Vissing-Jørgensen, 2011; Vayanos and Vila, 2021a), credit substitution for impaired intermediation (Gertler and Karadi, 2011; Gertler and Kiyotaki, 2010; Del Negro et al., 2017), liquidity transformation (Cui and Sterk, 2021), the central bank balance sheet as an independent instrument (Cúrdia and Woodford, 2011), government debt maturity management (Greenwood et al., 2015b), and central bank balance sheet sustainability (Reis, 2015, 2016, 2019). Reis (2015, 2019) analyses the fiscal consequences of large central bank balance sheets—solvency risks, seignorage constraints, and limits to fiscal burden alleviation—in nominal monetary frameworks. Our model is real (no money, no inflation), but reaches a complementary conclusion: the central bank can prevent market failure but cannot substitute for fiscal recapitalisation of the private financial sector. The closest paper to ours is Silva (2020), who studies a risk channel of QE in a continuous-time production economy with heterogeneous risk tolerance. In every one of these models, equilibrium exists with or without intervention; QE is a policy choice. Our result is qualitatively different: below a threshold, no equilibrium exists without central bank purchases.

Financial frictions, traps, and recapitalisation. The paper builds on the intermediary asset pricing literature (Brunnermeier and Sannikov, 2014; He and Krishnamurthy, 2013, 2019; Di Tella, 2017), sharing the feature that a single state variable drives all equilibrium quantities but generating the additional result that equilibrium can fail entirely. Our QE Trap differs from existing trap concepts along four dimensions: (i) *non-existence* of competitive equilibrium, not merely distortion; (ii) scarcity of risk-bearing capacity rather than safe assets (Caballero and Farhi, 2018), demand shortfalls from balance-sheet deleveraging (Eggertsson and Krugman, 2012; Guerrieri and Lorenzoni, 2017), or liquidity (Bigio, 2015); (iii) *Permanent QE* at the bad steady state and a *QE Trap* where calibrated recovery can exceed three centuries, requiring fiscal transfer, not self-correcting or escapable through monetary policy alone (Rognlie et al., 2018); and (iv) the leverage paradox linking the trap mechanism to asset pricing. On recapitalisation, Philippon and Schnabl (2013) and Veronesi and Zingales (2010) study the optimal form of intervention, taking its necessity as given; we derive the necessity from the steady-state structure.

Rational bubbles and the intermediation gap. Basak and Cuoco (1998) study the impact of limited household participation on asset prices in an exchange economy, whereas we study

a production economy with endogenous labor and a central bank. [Hugonnier \(2012\)](#) proves that portfolio constraints generate rational bubbles in otherwise bubble-free economies. In our model, leverage constraints can *close* rather than open the intermediation gap—linking the rational bubble literature to the equilibrium existence problem. We derive an exact ODE for the intermediation gap in a production economy, replacing the constant-volatility assumption of [Hugonnier \(2012\)](#)’s exchange setting, and show that the gap-to-price ratio delivers horizon-stable out-of-sample predictive content for the output gap.

The remainder of the paper is organised as follows. Section 2.2 presents the model. Section 2.3 establishes the equilibrium failure result. Section 2.4 introduces QE, derives the two-steady-state equilibrium, and analyses asset pricing implications. Section 2.5 establishes the intermediation gap results. Section 2.6 links the intermediation gap to real outcomes and derives policy implications. Section 2.7 presents empirical evidence. Section 2.8 concludes. All proofs are in the Appendix.

2.2 Model

2.2.1 Environment

Time is continuous with $t \in [0, \infty)$. Uncertainty is driven by a standard Brownian motion $\{Z_t\}_{t \geq 0}$ on a filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{t \geq 0}, \mathbb{P})$ satisfying the usual conditions.

The economy is populated by a representative household, a representative financier, and a continuum of monopolistically competitive firms indexed by $f \in [0, 1]$.

Technology and firms. Each firm f produces a differentiated good using the linear technology $Y_t(f) = A_t N_t(f)$, where A_t is aggregate productivity and $N_t(f)$ is labour input. Productivity evolves as geometric Brownian motion:

$$\frac{dA_t}{A_t} = \mu dt + \sigma dZ_t, \quad \mu \in \mathbb{R}, \sigma > 0. \quad (2.1)$$

Households aggregate differentiated goods via the Dixit–Stiglitz aggregator with elasticity of substitution $\epsilon > 1$. With flexible prices, each firm sets a constant markup over marginal cost. In symmetric equilibrium, factor shares are constant:

$$D_t = \frac{1}{\epsilon} Y_t, \quad \frac{W_t}{P_t} N_t = \frac{\epsilon - 1}{\epsilon} Y_t, \quad (2.2)$$

where D_t denotes aggregate dividends, W_t/P_t is the real wage, $Y_t = A_t N_t$ is aggregate output and $N_t = \int_0^1 N_t(f) df$ is aggregate employment.

Preferences. The household has preferences:

$$U_h = \mathbb{E} \int_0^\infty e^{-\delta t} \left(\ln C_{h,t} - \frac{N_t^{1+\varphi}}{1+\varphi} \right) dt, \quad \delta > 0, \varphi \geq 0. \quad (2.3)$$

The financier has logarithmic preferences over consumption:

$$U_i = \mathbb{E} \int_0^\infty e^{-\delta t} \ln C_{i,t} dt. \quad (2.4)$$

Both agents share the same discount rate δ .³

Financial markets and participation. Two financial assets are traded: a risk-free bond in zero net private supply paying rate r_t , and an equity claim on aggregate dividends in unit net supply. The household can invest only in the risk-free bond. The financier invests in both bonds and equity, with portfolio equity share ϕ_t (the fraction of her wealth allocated to equity). The household's budget constraint is deterministic:

$$dH_{h,t} = \left(r_t H_{h,t} + \frac{W_t}{P_t} N_t - C_{h,t} \right) dt, \quad (2.5)$$

where $H_{h,t}$ is the household's financial wealth. The financier's budget constraint is:

$$dH_{i,t} = (r_t H_{i,t} - C_{i,t}) dt + \phi_t H_{i,t} \left(\frac{dP_{D,t} + D_t dt}{P_{D,t}} - r_t dt \right), \quad (2.6)$$

where $H_{i,t}$ is the financier's financial wealth and $P_{D,t}$ is the equity price. Since households cannot hold equity, market clearing requires $\phi_t H_{i,t} = P_{D,t}$.

Assumption 2.1 (Environment parameters) $\epsilon > 1$, $\mu \in \mathbb{R}$, $\sigma > 0$, $\delta > 0$, and $\varphi \geq 0$.

³Online Appendix [BO.10](#) shows that the equilibrium failure result extends to CRRA preferences with elasticity of intertemporal substitution near unity, and to partial household equity participation. Online Appendix [BO.11](#) endogenises the limited participation restriction: Knightian uncertainty of just 0.7% of the Sharpe ratio makes the household optimally hold zero equity at the bad steady state, and 3.2% suffices throughout the state space. This formalises the sense in which households lack the screening and monitoring capacity that [Bernanke \(1983\)](#) identified as the distinguishing feature of financial intermediation: absent the specialised information and relationship capital embodied in the financial sector, households face ambiguity about aggregate risks that intermediaries, through their informational and screening advantage, do not.

2.2.2 Leverage Constraint

The financier faces a state-dependent leverage limit:

$$\phi_t \leq \bar{\phi}(\omega_t), \quad \bar{\phi}(\omega) \equiv 1 + \lambda(\omega - \underline{\omega}), \quad (2.7)$$

where $\omega_t \equiv H_{i,t}/P_{Y,t}$ is the financier's share of total wealth $P_{Y,t}$, $\underline{\omega} \in (0, 1/\epsilon)$ is the minimum capital threshold, and $\lambda > 0$ is the leverage sensitivity, with the joint restriction $\lambda \underline{\omega} \leq 1$ so that $\bar{\phi}(0) \geq 0$.

Assumption 2.2 (Leverage parameters) $0 < \underline{\omega} < 1/\epsilon$, $\lambda > 0$, and $\lambda \underline{\omega} \leq 1$.

The constraint combines two frictions that predate quantitative easing. *Above the threshold* ($\omega \geq \underline{\omega}$): following [Gertler and Kiyotaki \(2010\)](#) and [Gertler and Karadi \(2011\)](#), the financier can divert a fraction of assets for private benefit. Households lend only against verifiable collateral: the capital buffer above the regulatory minimum, yielding (2.7) with λ parameterising enforcement efficiency. In Online Appendix [BO.6](#), we show that log utility generates a closed-form franchise value of intermediation, making (2.7) a conservative bound. *Below the threshold* ($\omega < \underline{\omega}$): under Prompt Corrective Action (PCA), statutory since 1991 ([Federal Deposit Insurance Corporation, 1991](#)), undercapitalised intermediaries face mandatory restrictions on risky asset holdings, again yielding (2.7).⁴ Critically, in this region $\bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega}) < 1$: PCA forces the financier to hold *less* than 100% of her wealth in equity—she must liquidate risky positions and shift into safe assets. This forced de-equitisation is not a modelling convenience but an institutional fact: PCA mandates progressively harsher corrective actions as capital falls below the minimum, culminating in mandatory receivership. For the entire pre-crisis period, these constraints existed but did not bind; what changed in 2008 was the *state of financier balance sheets*.

In the taxonomy of [He et al. \(2017, Section 4.3\)](#), our constraint is a *leverage constraint*: it directly caps the portfolio equity share ϕ , and the cap tightens as financier wealth ω falls, producing procyclical deleveraging. This contrasts with the *equity constraints* of [He and Krishnamurthy \(2012, 2013\)](#), under which the intermediary must maintain a minimum equity stake

⁴PCA, enacted as Title I of the Federal Deposit Insurance Corporation Improvement Act of 1991, classifies insured depository institutions into five capital categories (well capitalised, adequately capitalised, undercapitalised, significantly undercapitalised, critically undercapitalised) and imposes escalating mandatory supervisory actions as capital falls. Undercapitalised institutions face restrictions on asset growth, acquisitions, and dividends; significantly undercapitalised institutions must divest risky assets and may be required to raise capital or merge; critically undercapitalised institutions face mandatory receivership within 90 days. The constraint $\bar{\phi}(\omega) < 1$ for $\omega < \underline{\omega}$ captures this escalating forced deleveraging in reduced form.

but realised leverage rises countercyclically as equity is depleted. The qualitative implications of this modelling choice for the long-run dynamics are developed in Section 2.4 and Online Appendix BO.13.

2.2.3 Benchmarks and Baseline Dynamics

When both agents can trade all assets (Appendix B.1), the economy admits a complete-markets benchmark in which consumption shares, labour supply, and total wealth are constant, with $P_{Y,t} = Y_t/\delta$.

With limited participation but no leverage constraint (Appendix B.2), the entire equilibrium is characterised by the financier's wealth share $\omega_t \equiv H_{i,t}/P_{Y,t}$. Goods-market clearing under log utility pins consumption by wealth: $C_{h,t} = (1 - \omega_t)Y_t$ and $C_{i,t} = \omega_t Y_t$. The financier's log SDF and the Dixit–Stiglitz markup together pin total wealth $P_{Y,t} = Y_t/\delta$, equity value $P_{D,t} = P_{Y,t}/\epsilon$, and human capital $P_{L,t} = (1 - 1/\epsilon)P_{Y,t}$; the household's financial wealth follows as the accounting residual $H_{h,t} = (1/\epsilon - \omega_t)P_{Y,t}$. Every stock, flow, and price is therefore a function of the single state variable ω_t , and this structure is preserved when the leverage constraint is added in Section 2.3.

The inverse wealth share $Q_t = 1/\omega_t$ is a non-negative local martingale and $\omega_t \rightarrow 1/\epsilon$ almost surely, but with infinite expected hitting time. This “Zeno's paradox” rationalises why the pre-crisis US economy functioned seamlessly for decades: financiers were sufficiently capitalised that the leverage constraint, though present, never bound.

2.3 Capital Depletion Causes Equilibrium Failure

The intermediary asset-pricing literature has studied at length how depleted capital distorts asset prices. We show it can do more than distort: under the leverage constraint (2.7), capital depletion below a threshold ω_c destroys competitive equilibrium altogether—no price vector clears the market for aggregate risk. The mechanism rests on two institutional primitives: limited participation and a state-dependent leverage cap, both salient features of the post-2008 U.S. financial sector, and is triggered by realised capital depletion. Log utility delivers the cleanest argument but is inessential: the non-existence result survives under CRRA and Epstein–Zin preferences, with ω_c itself preference-independent (Section 2.3.2).

The argument rests on three observations, each elementary, that combine into a non-

existence result.

1. *Household bond demand is perfectly inelastic to the interest rate.* Under log utility the household consumes a fraction δ of her wealth at every price vector, so her financial wealth is pinned by ω through the accounting identity $(1 - \omega_t)P_{Y,t} = H_{h,t} + P_{L,t}$. With human capital $P_{L,t} = (1 - 1/\epsilon)P_{Y,t}$ priced by the financier's log SDF, the household's bond demand is

$$B_{h,t}^d = H_{h,t} = \left(\frac{1}{\epsilon} - \omega_t \right) P_{Y,t}. \quad (2.8)$$

Because $P_{Y,t} = Y_t/\delta$ is itself interest-rate-independent under log utility, equation (2.8) is independent of r_t : income and substitution effects of interest-rate changes cancel exactly. No interest-rate movement, however extreme, induces the household to substitute from bonds into equity—her demand schedule is *vertical* at every r . This disables the standard safety-valve of price adjustment, a friction qualitatively sharper than the zero-lower-bound constraint of [Caballero and Farhi \(2018\)](#): there, conventional policy fails at a corner of an existing equilibrium; here no lower bound is needed because the demand curve is rigid at every rate, positive or negative.

2. *Market clearing requires the financier to absorb all equity, giving a hyperbolic required portfolio share.* By limited participation, the financier is the sole potential equity holder. With $P_{D,t} = P_{Y,t}/\epsilon$ (dividend share $D/Y = 1/\epsilon$ priced by the financier's log SDF) and $H_{i,t} = \omega_t P_{Y,t}$ (definition of ω), equity-market clearing requires

$$\phi^*(\omega) \equiv \frac{P_D}{H_i} = \frac{1}{\epsilon \omega}. \quad (2.9)$$

This required portfolio share is hyperbolic in ω and explodes as intermediary capital is depleted. It reflects an accounting fact: the aggregate equity market does not shrink with financier balance sheets, so as ω falls the same fixed claim must be levered onto a shrinking capital base.

3. *Permitted leverage is linear and bounded.* The constraint (2.7) caps portfolio weight at $\bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega})$, a line that equals one at the PCA floor and rises with slope λ . Required leverage is a hyperbola in ω ; permitted leverage is a line in ω ; they cross exactly once in the interior at a point we call ω_c . Figure 2.1 shows the geometry.

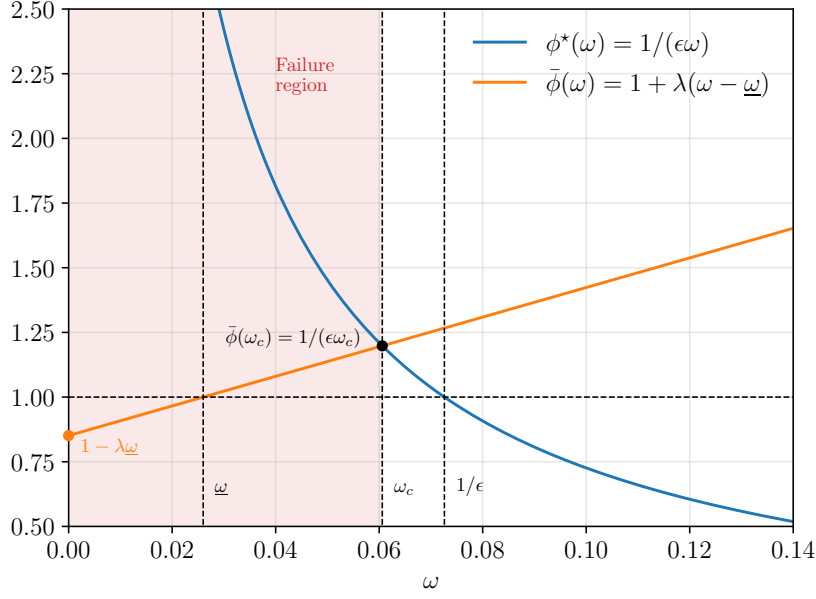


Figure 2.1: The equilibrium-failure mechanism. Required portfolio share $\phi^*(\omega) = 1/(\epsilon\omega)$ (blue hyperbola) is what the financier must hold under limited participation to clear the equity market; it explodes as intermediary capital is depleted. Permitted portfolio share $\bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega})$ (orange line) is the leverage cap, equal to one at the PCA floor $\underline{\omega}$ and rising with slope λ . The two curves cross exactly once, at ω_c ; for $\omega < \omega_c$ (shaded), required exceeds permitted and no price vector clears both markets simultaneously. The line meets the vertical axis at $\bar{\phi}(0) = 1 - \lambda\underline{\omega}$. The state variable ω ranges in $(0, 1)$; the figure zooms into the crisis region, as all interesting structure is concentrated near $\omega = 0$. Calibrated parameters from Table 2.1.

Theorem 2.1 (Equilibrium failure) *With limited household stock-market participation and leverage constrained financiers, competitive equilibrium fails to exist for all $\omega_t < \omega_c$ in the absence of central-bank intervention, where ω_c is the unique positive solution of*

$$\bar{\phi}(\omega_c) = \frac{1}{\epsilon \omega_c} \iff \lambda \omega_c^2 + (1 - \lambda \underline{\omega}) \omega_c - 1/\epsilon = 0, \quad (2.10)$$

i.e.

$$\omega_c = \frac{\underline{\omega}}{2} - \frac{1}{2\lambda} + \sqrt{\left(\frac{\underline{\omega}}{2} - \frac{1}{2\lambda}\right)^2 + \frac{1}{\epsilon\lambda}}. \quad (2.11)$$

Proof sketch. For $\omega_t < \omega_c$, observation 2 gives $\phi^*(\omega_t) > \bar{\phi}(\omega_t)$: the portfolio weight market clearing requires exceeds the cap. The residual excess equity supply

$$[\phi^*(\omega_t) - \bar{\phi}(\omega_t)] \omega_t P_{Y,t} = \left[\frac{1}{\epsilon \omega_t} - \bar{\phi}(\omega_t) \right] \omega_t P_{Y,t} > 0$$

has no buyer. The household's bond demand (2.8) is independent of r (observation 1), so no interest-rate adjustment transfers excess onto her balance sheet, and limited participation rules out her holding equity at any price. No price vector clears both markets simultaneously. The full argument is in Appendix B.5. $\square$

Remark 2.1 (Closed form for ω_c) *The state variable is confined to the interval $(\underline{\omega}, 1/\epsilon)$: the upper endpoint is the frictionless benchmark—at $\omega = 1/\epsilon$ the financier voluntarily holds $\phi = 1$, no leverage is needed, and the economy attains its complete-markets allocation—while the lower endpoint is the PCA floor, where the constraint forces $\bar{\phi}(\underline{\omega}) = 1$, the volatility of ω vanishes, and the state becomes absorbing. A post-crisis economy that has fallen below $1/\epsilon$ therefore sits between the two: drawn upward toward the frictionless benchmark, blocked from below by the absorbing floor.*

A natural measure of the severity of market failure is the fraction of the interval $(\underline{\omega}, 1/\epsilon)$ in which equilibrium fails. Rearranging (2.10):

$$\frac{\omega_c - \underline{\omega}}{1/\epsilon - \underline{\omega}} = \frac{1}{1 + \lambda \omega_c}. \quad (2.12)$$

The factor $1 + \lambda \omega_c$ is the economy's leverage tolerance at the crisis threshold; its reciprocal is the fraction of the good basin that cannot clear markets without central-bank intervention.

Numerical sensitivity. *At the empirical calibration (Table 2.1), holding $\epsilon = 13.7741$ and $\underline{\omega} = 0.0260$ fixed:*

	λ	ω_c	$\omega_c - \underline{\omega}$	$\frac{\omega_c - \underline{\omega}}{1/\epsilon - \underline{\omega}}$
Half of calibrated	2.86	0.065	0.039	84%
Calibrated	5.72	0.061	0.035	74%
Double	11.45	0.055	0.029	62%

At the calibrated λ , for 74% of the interval $(\underline{\omega}, 1/\epsilon)$ markets fail to clear: equilibrium failure is not a knife-edge phenomenon but a broad condition that dominates almost three quarters of the interval. Halving λ pushes this fraction to 84%; doubling it shrinks it to 62%. Comparative statics $\partial \omega_c / \partial \lambda < 0$ and $\partial \omega_c / \partial \underline{\omega} > 0$ follow directly from (2.11). Inverting for λ given $(\epsilon, \underline{\omega}, \omega_c)$ yields $\lambda = (1/\epsilon - \omega_c) / [\omega_c(\omega_c - \underline{\omega})]$: this is how λ is recovered from the calibration.

Theorem 2.1 is a statement about equilibrium *existence*, not distortion: for $\omega < \omega_c$ without a central bank, no price vector simultaneously clears the bond and equity markets at the quantities budget constraints permit. Its three ingredients each correspond to a feature of the 2008 U.S. financial sector: *limited participation* to the inability of the non-bank public to backstop intermediaries in wholesale funding markets; *leverage constraints* to the interaction of risk-weighted capital requirements, moral-hazard enforcement, and Prompt Corrective Action; *depleted capital* to the realised collapse in broker-dealer and bank-holding-company equity. In

the region $(\underline{\omega}, \omega_c)$, all three are needed: relax any one and equilibrium is restored. Below $\underline{\omega}$, the failure is unconditional: Prompt Corrective Action drives $\bar{\phi}(\omega) < 1$, and even with the household at full participation and all her financial wealth $(1/\epsilon - \omega)P_{Y,t}$ in equity, a strictly positive residual $\omega [1 - \bar{\phi}(\omega)] P_{Y,t}$ remains unabsorbed—the non-tradable human-capital component of household wealth cannot be redirected to fill the gap (Theorem [BO.5](#), Online Appendix [BO.11](#)).

2.3.1 Dynamics of ω and the Onset of Failure

Theorem [2.1](#) characterises failure at a fixed ω . To see how the economy *reaches* the failure region, we write the law of motion for ω_t on the domain where equilibrium exists. For $\omega_t \in (\omega_c, 1/\epsilon)$ the leverage constraint [\(2.7\)](#) does not bind: required leverage $\phi^*(\omega_t) = 1/(\epsilon\omega_t)$ is below the cap $\bar{\phi}(\omega_t)$, the financier freely absorbs all equity, and ω_t evolves as the one-dimensional diffusion

$$d\omega_t = \frac{\sigma_\omega(\omega_t)^2}{\omega_t} dt + \sigma_\omega(\omega_t) dZ_t, \quad (2.13)$$

with volatility

$$\sigma_\omega(\omega) = \frac{\sigma(1+\varphi)(1-\omega)(1/\epsilon - \omega)}{(1+\varphi)(1-\omega) - (1/\epsilon - \omega)} > 0 \quad \text{for } \omega \in (\omega_c, 1/\epsilon). \quad (2.14)$$

Equations [\(2.13\)](#)–[\(2.14\)](#) are derived in Theorem [B.1](#) of Appendix [B.2](#) from the financier's budget constraint, goods-market clearing under log utility, and the Dixit–Stiglitz factor shares. The inverse share $Q_t = 1/\omega_t$ is a non-negative local martingale: the drift $\mu_\omega(\omega) = \sigma_\omega(\omega)^2/\omega > 0$ is the financier's expected excess return from bearing equity risk, and ω_t drifts almost surely toward the frictionless steady state $\omega = 1/\epsilon$, where $\phi^*(1/\epsilon) = 1$ and no leverage is needed.

Where failure arises. Note first that the formula [\(2.14\)](#) is smooth at ω_c and below; the boundary of existence is *not* visible in the SDE itself. The reason is that [\(2.13\)](#) is derived under the silent assumption that the financier freely chooses the market-clearing portfolio $\phi_t = \phi^*(\omega_t) = 1/(\epsilon\omega_t)$. Feasibility of that choice — $\phi^*(\omega_t) \leq \bar{\phi}(\omega_t)$ — holds iff $\omega_t \geq \omega_c$. For $\omega_t < \omega_c$, the portfolio the SDE tacitly assumes violates the leverage constraint [\(2.7\)](#): the equilibrium premise behind [\(2.13\)](#) breaks, and Theorem [2.1](#) applies.

Define the first-passage time to the crisis threshold:

$$\tau_c \equiv \inf\{t \geq 0 : \omega_t = \omega_c\}. \quad (2.15)$$

For any starting point $\omega_0 \in (\omega_c, 1/\epsilon)$, $\mathbb{P}(\tau_c < \infty) > 0$: although the drift favours upward motion toward $1/\epsilon$, the diffusion is non-degenerate and two-sided, so a sufficiently adverse run of productivity shocks depletes ω_t to ω_c in finite time. At $t = \tau_c$ the constraint binds with equality, $\bar{\phi}(\omega_c) = \phi^*(\omega_c) = 1/(\epsilon\omega_c)$. The next adverse shock pushes $\omega_t < \omega_c$, the feasibility condition above fails, and no price vector clears both markets. *The equilibrium described by (2.13) terminates at τ_c* —not because the SDE’s formula breaks down, but because its implicit assumption does. The threshold ω_c is a boundary of existence, not of the state space: reached from above, it marks the endpoint of the competitive equilibrium itself.

The picture is not that the economy sits mired in a failure region but that it is always one sufficiently adverse shock-sequence away from one. Decades of pre-2008 U.S. history with $\omega_t \gg \omega_c$ are entirely consistent with (2.13); what the crisis revealed was that the stopping time τ_c was in fact finite.

2.3.2 Which Assumptions Are Essential?

Theorem 2.1 relies on three modelling choices whose robustness merits scrutiny. Online Appendices BO.10 and BO.13 establish that the result is robust along each dimension; we summarise the findings here.

Assumption	Essential?	What happens if relaxed	Reference
Log utility (household)	No	The crisis threshold ω_c is <i>preference-independent</i> : it does not depend on the household’s EIS ψ or risk aversion γ . Non-existence survives under CRRA and Epstein–Zin preferences.	Thm. BO.4 Thm. BO.7
Limited participation	Degree does not matter	Failure survives for any household equity share $\alpha_h \in [0, 1)$. The crisis threshold $\omega_c(\alpha_h)$ falls monotonically in α_h but remains strictly positive.	Thm. BO.2
Bounded intermediary leverage	Yes	Without an upper bound on ϕ , the financier absorbs all equity at any ω and no failure occurs. Under an alternative equity constraint $\bar{\phi} \equiv 1/a$, failure arises for $\omega < a/\epsilon$: boundedness is what matters, not the specific functional form (Online Appendix BO.13).	Thm. 2.1 Prop. BO.13
Exogenous non-participation	No	Ambiguity aversion with $\kappa \approx 0.7\%$ of the Sharpe ratio generates endogenous non-participation at $\omega = \underline{\omega}$; $\kappa \approx 3.2\%$ suffices everywhere.	Thm. BO.5 Cor. BO.4

Household preferences. The crisis threshold ω_c is *completely preference-independent*: it depends on $(\epsilon, \lambda, \underline{\omega})$ but not on the household’s elasticity of intertemporal substitution ψ or risk aversion γ . This is because the failure condition $\bar{\phi}(\omega) \omega < 1/\epsilon$ uses only the leverage constraint and limited participation; the household’s Euler equation does not enter. The structural identity

$P_D/P_Y = 1/\epsilon$ (from the Dixit–Stiglitz factor share and the financier’s log SDF) ensures that the required leverage $\phi^*(\omega) = 1/(\epsilon\omega)$ is preference-free.

We verify this explicitly under two generalisations. Under *CRR*A household preferences (a single parameter, $\gamma = 1/\psi$), the crisis threshold is identical to the log case for any $\psi > 0$ (Theorem BO.4). Under *Epstein–Zin* preferences (risk aversion γ and EIS ψ decoupled), ω_c remains preference-independent (Theorem BO.7). Dynamics above ω_c —in particular the convergence behaviour near the boundaries $\underline{\omega}$ and $1/\epsilon$ —do depend on preferences, but the failure theorem itself does not. We return to the preference-sensitivity of the dynamics in Section 2.4 and Online Appendices BO.10 and BO.12.

Robust failure below $\underline{\omega}$. Perhaps the strongest result is Theorem BO.5: for $\omega < \underline{\omega}$, no competitive equilibrium exists *regardless of the household’s equity share*, including $\alpha_h = 1$. The leverage constraint forces $\bar{\phi}(\omega) < 1$ when $\omega < \underline{\omega}$, so the financier cannot even hold her own wealth entirely in equity. The household’s financial wealth is too small to fill the gap. This result requires no assumption on household preferences, ambiguity, or participation.

2.4 Quantitative Easing, the QE Trap, and Asset Prices

2.4.1 Central Bank as Risk-Bearer of Last Resort

The equilibrium failure of Theorem 2.1 motivates the introduction of a public sector. Once the crisis has depleted financier capital to the point where leverage constraints bind, the central bank purchases the residual equity that the constrained financier cannot absorb.

Assumption 2.3 (Central bank and fiscal authority) *The central bank stands ready to absorb whatever fraction of the equity market the constrained financier cannot hold, financing its purchases by issuing risk-free bonds. Profits and losses are remitted to the fiscal authority (Treasury), which may transfer resources to the financial sector as a recapitalisation (Section 2.6.4).*

2.4.2 Equilibrium with QE

Proposition 2.1 (Equilibrium with QE) *Under Assumptions 2.2 and 2.3, a competitive equilibrium exists for all $\omega \in (0, 1)$. Equity market clearing uniquely pins down the central bank’s equity share as*

the residual:

$$\alpha_t = \max(0, 1 - \epsilon \bar{\phi}(\omega_t) \omega_t) = \begin{cases} 1 - \epsilon \bar{\phi}(\omega_t) \omega_t & \text{if } \omega_t < \omega_c, \\ 0 & \text{if } \omega_t \geq \omega_c. \end{cases} \quad (2.16)$$

At each date t , the equilibrium allocation is given by

$$N_t = \left[\frac{\epsilon - 1}{\epsilon(1 - \omega_t)} \right]^{1/(1+\varphi)}, \quad Y_t = A_t N_t, \quad (2.17)$$

$$C_{h,t} = (1 - \omega_t) Y_t, \quad C_{i,t} = \omega_t Y_t, \quad (2.18)$$

$$P_{Y,t} = Y_t / \delta, \quad P_{D,t} = P_{Y,t} / \epsilon, \quad (2.19)$$

$$H_{i,t} = \omega_t P_{Y,t}, \quad H_{h,t} = (1/\epsilon - \omega_t) P_{Y,t}. \quad (2.20)$$

The financier's portfolio share and the central bank's equity holding are

$$\phi_t = \min\left(\bar{\phi}(\omega_t), \frac{1}{\epsilon \omega_t}\right), \quad \alpha_t = \max(0, 1 - \epsilon \bar{\phi}(\omega_t) \omega_t). \quad (2.21)$$

As functions of the state ω_t , the mappings (2.17)–(2.20) are the same whether or not the leverage constraint binds: the central bank absorbs the residual equity, leaving the static real allocation untouched. The law of motion for ω_t itself is altered when $\omega_t < \omega_c$ (Theorem 2.2), so the realised processes $\{\omega_t, N_t, Y_t, C_{h,t}, C_{i,t}, P_{Y,t}\}$ differ in the crisis regime from what they would be absent the constraint.

The central bank absorbs exactly the equity that the constrained financier cannot hold. The household, whose bond demand is perfectly inelastic, is indifferent between lending to the financier and lending to the central bank, so at each ω the real allocation (2.17)–(2.20) is the same as without the constraint. The formula for α_t in (2.21) is not a policy choice but a market-clearing identity: it holds for *any* bounded leverage function $\bar{\phi}(\omega)$, not just the specific linear form in (2.7). Online Appendix BO.13 verifies this by deriving the corresponding α_t under an equity constraint $\phi \leq 1/a$, where the formula becomes $\alpha_t = \max(0, 1 - \epsilon \omega/a)$ but all qualitative properties—equilibrium existence, intermediation gap elimination, the leverage paradox—are preserved.

2.4.3 Dynamics with the Leverage Constraint

Theorem 2.2 (Dynamics of ω_t with leverage constraint) *The financier's wealth share evolves as*

$$d\omega_t = \frac{\sigma_\omega(\omega_t)^2}{\omega_t} dt + \sigma_\omega(\omega_t) dZ_t, \quad (2.22)$$

where the volatility function is continuous on $(0, 1)$ and given by

$$\sigma_\omega(\omega) = \begin{cases} \frac{\sigma(1+\varphi)\lambda(1-\omega)\omega(\omega-\underline{\omega})}{(1+\varphi)(1-\omega)-\lambda\omega(\omega-\underline{\omega})} & \text{if } \omega < \omega_c, \\ \frac{\sigma(1+\varphi)(1-\omega)(1/\epsilon-\omega)}{(1+\varphi)(1-\omega)-(1/\epsilon-\omega)} & \text{if } \omega \geq \omega_c. \end{cases} \quad (2.23)$$

The drift relation $\mu_\omega = \sigma_\omega^2/\omega$ reflects the financier's excess return from bearing equity risk: the drift is always positive when the diffusion is nonzero, so the financier's capital accumulation is fastest when her wealth share is most volatile.

The sign of σ_ω determines how the economy responds to shocks (Proposition BO.1 in the Online Appendix). Since $\sigma_\omega = \omega(\phi - 1)\sigma_R$, the sign depends on whether the financier is leveraged ($\phi > 1$) or de-leveraged ($\phi < 1$). In the crisis regime ($\omega < \underline{\omega}$), the leverage constraint forces $\phi < 1$: the financier cannot hold even her own wealth entirely in equity. A positive productivity shock then raises total wealth P_Y faster than the financier's equity-underweight portfolio, mechanically reducing $\omega = H_i/P_Y$ —so $\sigma_\omega < 0$. In the recovery regime ($\underline{\omega} < \omega < 1/\epsilon$), $\phi > 1$ (the financier is leveraged) and $\sigma_\omega > 0$: positive shocks improve financier capitalisation, aiding recovery.

Proposition 2.2 (Q-martingale preservation) $Q_t = 1/\omega_t$ is a non-negative local martingale regardless of whether the leverage constraint binds.

This is a consequence of log utility: the financier's consumption-wealth ratio is constant at δ , and the linear and quadratic terms in ϕ cancel exactly.

2.4.4 Two Steady States: Permanent QE and the QE Trap

We call a point ω^* a *steady state* of the wealth-share dynamics if both drift and diffusion vanish there: $\sigma_\omega(\omega^*) = \mu_\omega(\omega^*) = 0$. If the economy starts at ω^* , it stays there forever ($d\omega = 0$). In the Feller classification of one-dimensional diffusions, each steady state is a *natural boundary*:

the economy is drawn toward it from its basin but never reaches it in finite time—like Zeno’s runner, who halves the remaining distance at each step but never arrives.⁵

Theorem 2.3 (Two steady states) *The wealth-share dynamics have exactly two steady states in $(0, 1)$: $\omega = \underline{\omega}$ and $\omega = 1/\epsilon$. Both are natural boundaries: if $\omega_0 = \omega^* \in \{\underline{\omega}, 1/\epsilon\}$, then $\omega_t = \omega^*$ for all $t \geq 0$; if $\omega_0 \neq \omega^*$, then ω_t converges to ω^* almost surely from the corresponding basin but $\omega_t \neq \omega^*$ for all finite t .*

The good steady state ($\omega = 1/\epsilon$). The financier holds exactly 100% of her wealth in equity, no leverage constraint binds, the central bank holds no equity, and the economy coincides with the complete-markets benchmark.

The bad steady state ($\omega = \underline{\omega}$). The central bank permanently holds fraction $\alpha = 1 - \epsilon\underline{\omega}$ of the equity market. Employment is depressed:

$$N(\underline{\omega}) = \left(\frac{1 - 1/\epsilon}{1 - \underline{\omega}} \right)^{\frac{1}{1+\varphi}} < 1 = N(1/\epsilon). \quad (2.24)$$

Both steady states are natural boundaries (Appendix B.9): continuous paths cannot cross them, so no transition between basins is possible. The state space $(0, 1)$ splits into three *invariant basins*—regions the economy cannot leave once entered—each with a determined long-run outcome:

- $(0, \underline{\omega})$: converges to the bad steady state $\underline{\omega}$ (permanent QE, depressed output).
- $(\underline{\omega}, 1/\epsilon)$: converges to the good steady state $1/\epsilon$ (full recovery).
- $(1/\epsilon, 1)$: converges to the good steady state $1/\epsilon$ from above.

The crisis threshold $\omega_c \in (\underline{\omega}, 1/\epsilon)$ is *not* a boundary—continuous paths cross it freely. It marks where the leverage constraint starts to bind, but poses no barrier to the dynamics.

Theorem 2.4 (Convergence) *For any $\omega_0 \in (0, 1) \setminus \{\underline{\omega}, 1/\epsilon\}$:*

$$\omega_t \xrightarrow{a.s.} \begin{cases} \underline{\omega} & \text{if } \omega_0 \in (0, \underline{\omega}), \\ 1/\epsilon & \text{if } \omega_0 \in (\underline{\omega}, 1). \end{cases} \quad (2.25)$$

In all cases, $\mathbb{E}[\tau] = \infty$, where τ is the expected time to reach the relevant steady state.

⁵Formally, a natural boundary is one that is *unattainable*—the first hitting time from any interior point is infinite almost surely—and *non-starting*—it cannot serve as an initial condition for the interior diffusion. The term “absorbing state” is sometimes used in the macro-finance literature for the same object. In stochastic process theory, an absorbing state is one the process *reaches* in finite time and stays at; our steady states are approached only asymptotically, so we avoid this terminology.

The almost sure convergence follows from the supermartingale property of $Q = 1/\omega$ (Appendix B.10). For $\omega_0 \in (0, \underline{\omega})$, the economy converges to the bad steady state: QE is literally permanent — the *Permanent QE zone*. For $\omega_0 \in (\underline{\omega}, 1/\epsilon)$, the economy converges to the good steady state, but the infinite expected arrival time means QE tapers gradually over a timescale that can span centuries. We call the sub-region $(\underline{\omega}, \omega_T)$ — where the deterministic recovery time exceeds a policy-relevant horizon $\bar{T}$ — the *QE Trap*: exit is mathematically certain but practically impossible (Online Appendix BO.14).

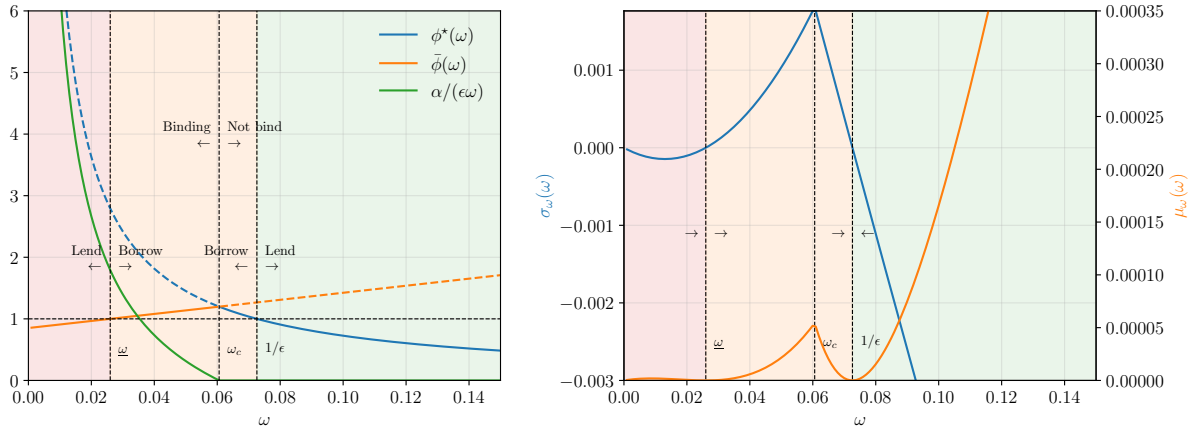


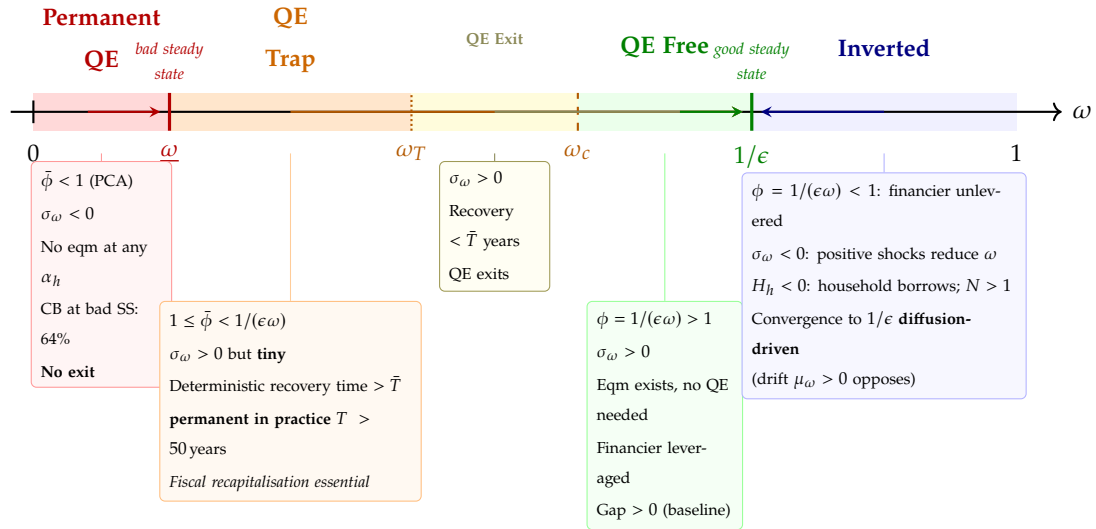
Figure 2.2: Equilibrium dynamics as a function of the state variable ω . (a) Financier and central bank equity holdings and the leverage constraint, with arrows showing the regions where the constraint is (not) binding and where the financier borrows or lends. (b) Wealth-share volatility $\sigma_\omega(\omega)$ and drift $\mu_\omega = \sigma_\omega^2/\omega$, with arrows showing convergence to the absorbing steady states. In both panels the red shading marks the crisis regime, the orange shading marks the recovery regime, and the green shading marks the unconstrained regime. Vertical dashed lines mark $\underline{\omega}$, ω_c , and $1/\epsilon$. Calibrated parameters from Table 2.1.

Remark 2.2 (Permanent QE and the QE Trap are distinct) *The two regimes in which QE persists arise from different mechanisms:*

- **Permanent QE ($0, \underline{\omega}$):** $\underline{\omega}$ is a natural (absorbing) boundary because the leverage constraint forces $\bar{\phi}(\underline{\omega}) = 1$, which implies $\sigma_\omega(\underline{\omega}) = 0$. The economy starting below $\underline{\omega}$ cannot cross this boundary and $\omega_t \rightarrow \underline{\omega}$ almost surely. Exit is impossible even in theory: no arbitrarily long wait delivers $\omega_t > \underline{\omega}$. Only a fiscal transfer Δ with $\omega_0 + \Delta > \underline{\omega}$ escapes.
- **QE Trap ($\underline{\omega}, \omega_T$):** the diffusion σ_ω vanishes at $\underline{\omega}$, so the drift $\mu_\omega = \sigma_\omega^2/\omega$ is tiny just above it. Recovery to ω_c is certain ($\omega_t \xrightarrow{a.s.} 1/\epsilon$) but takes centuries at typical post-crisis capital ratios. Exit is possible in theory but not on any policy-relevant horizon. A fiscal transfer with $\omega_0 + \Delta > \omega_T$ restores self-resolving QE.

The distinction is policy-relevant: a transfer sufficient to escape the QE Trap may be insufficient to escape Permanent QE. We use the umbrella phrase “the trap” only informally; whenever the specific mechanism matters we name it.

Remark 2.3 (Why the leverage form produces Permanent QE) *The interior absorbing steady state at $\underline{\omega}$ is a consequence of the leverage form of (2.7). At $\omega = \underline{\omega}$, the cap forces $\bar{\phi}(\underline{\omega}) = 1$, so $\sigma_\omega(\underline{\omega}) = 0$ (Theorem 2.2) and hence $\mu_\omega(\underline{\omega}) = 0$: both drift and diffusion vanish, pinning the process at $\underline{\omega}$. Under an equity constraint $\bar{\phi} \equiv 1/a > 1$ (He and Krishnamurthy, 2012, 2013), the financier retains leverage strictly above one at every $\omega > 0$; no interior absorbing state arises and Permanent QE disappears as a distinct regime. The QE Trap phenomenon of impractically slow recovery is not eliminated by switching constraint forms, however: $\mu_\omega \rightarrow 0$ as $\omega \rightarrow 0$ under either constraint, and deterministic recovery from severely depleted capital still takes centuries. What the leverage constraint adds is that the slow-drift region is anchored to a strictly positive capital level $\underline{\omega}$ rather than sitting against the boundary at zero. Equilibrium failure also survives under the equity constraint (Online Appendix BO.13, Proposition BO.13), but the failure region $(0, a/\epsilon)$ is typically narrower than the leverage-constraint region $(0, \omega_c)$ —QE is needed for a smaller range of ω while the qualitative necessity survives. Online Appendix BO.13 develops the full comparison.*



Steady states $\underline{\omega}$ and $1/\epsilon$: both $\sigma_\omega = \mu_\omega = 0$. Each is a *natural boundary* — the economy converges toward it but never arrives ($\mathbb{E}[\tau] = \infty$). $Q = 1/\omega$ is a supermartingale $\Rightarrow \omega_t \xrightarrow{\text{a.s.}} \{\underline{\omega}, 1/\epsilon\}$.

ω_T (dotted): the QE Trap boundary. Deterministic recovery time exceeds $\bar{T}$ years for $\omega < \omega_T$. At $\bar{T} = 50$, the exact boundary is $\omega_T \approx 0.0581$; the leading-order formula gives 0.0595. At $\omega = 0.030$: recovery takes **8,547 years**.

ω_c (dashed): crisis threshold where the leverage constraint binds. Not a boundary — freely crossed.

Three invariant basins: $(0, \underline{\omega}) \rightarrow$ bad steady state (Permanent QE); $(\underline{\omega}, 1/\epsilon) \rightarrow$ good steady state; $(1/\epsilon, 1) \rightarrow$ good steady state from above.

Figure 2.3: The five regions of ω . $\omega_t = H_{i,t}/P_{Y,t}$ is the single state variable. The regions partition $(0, 1) \setminus \{\underline{\omega}, 1/\epsilon\}$: Permanent QE $(0, \underline{\omega})$, QE Trap $(\underline{\omega}, \omega_T)$, Self-Resolving QE $[\omega_T, \omega_c)$, QE Free $[\omega_c, 1/\epsilon)$, Inverted $(1/\epsilon, 1)$; the natural boundaries $\underline{\omega}$ and $1/\epsilon$ are unattainable steady states. Solid lines: $\underline{\omega}$ and $1/\epsilon$ (natural boundaries). Dotted: ω_T , the QE Trap boundary — below it, deterministic recovery to ω_c exceeds $\bar{T}$ years (set to 50 in the figure). Dashed: ω_c (crisis threshold, freely crossed). In the Permanent QE zone, no exit is possible even in theory. In the QE Trap, exit is certain but takes centuries — too slow for any government, central bank, or electorate. Only Self-Resolving QE ($\omega > \omega_T$) exits on a human timescale. Fiscal recapitalisation pushes ω from the Trap into the Self-Resolving zone. Horizontal spacing between ω_T and ω_c is exaggerated for label visibility; at the Table 2.1 calibration they are about 0.0025 apart. Calibrated parameters from Table 2.1: $\epsilon = 13.7741$, $\underline{\omega} = 0.0260$, $\lambda = 5.7231$, $\sigma = 0.1505$, and $\bar{T} = 50$ years.

2.4.5 Asset Prices and Risk Premia

Theorem 2.5 (Asset pricing) *The market price of risk, equity return volatility, and risk-free rate are:*

$$\Theta_t = \frac{\sigma_\omega(\omega_t)}{\omega_t} + \sigma + \frac{\sigma_\omega(\omega_t)}{(1 + \varphi)(1 - \omega_t)}, \quad (2.26)$$

$$\sigma_{R,t} = \sigma + \frac{\sigma_\omega(\omega_t)}{(1 + \varphi)(1 - \omega_t)}, \quad (2.27)$$

$$r_t = \delta + \mu_{Y,t} - \sigma_{R,t} \Theta_t, \quad (2.28)$$

where σ_ω is given by (2.23) and $\mu_{Y,t}$ is the drift of output. The equity risk premium is $\mu_{R,t} - r_t = \sigma_{R,t} \Theta_t$.

The risk-free rate (2.28) is depressed in crisis through the $\sigma_{R,t} \Theta_t$ term: as ω falls, leveraged intermediaries amplify return volatility σ_R and the market price of risk Θ , whose product drives r well below the rate of time preference. This balance-sheet channel differs from the habit and long-run risk mechanisms of [Campbell and Cochrane \(1999\)](#) and [Bansal and Yaron \(2004\)](#), where r falls because of increased effective risk aversion or precautionary savings; here the channel is intermediary capital, and it operates even under log utility (constant individual risk aversion).

In the recovery regime ($\underline{\omega} < \omega < 1/\epsilon$), $\sigma_\omega > 0$ and risk premia are above the frictionless benchmark. Risk premia peak at an intermediate ω and decline as $\omega \rightarrow 1/\epsilon$, consistent with the hump shape documented by [He et al. \(2017\)](#). Positive productivity shocks raise ω and compress risk premia, delivering countercyclicality.

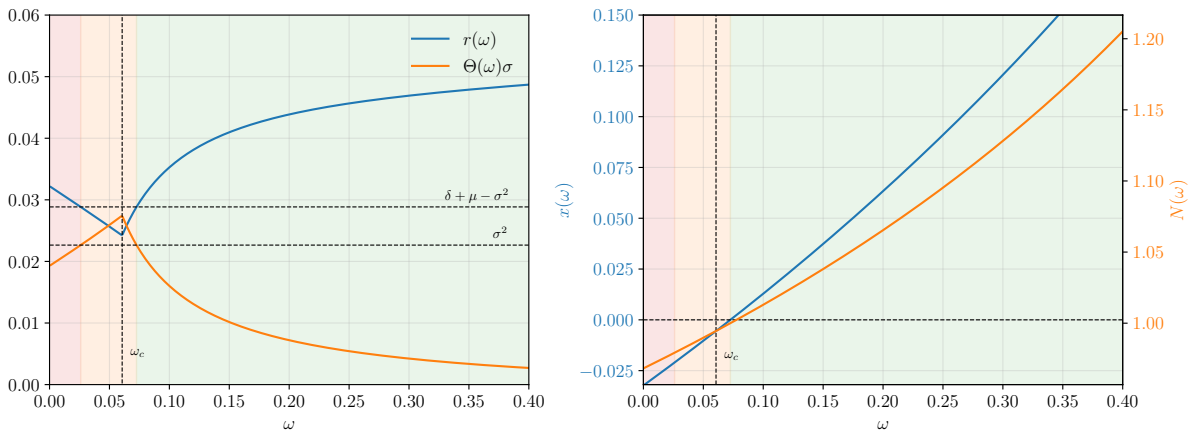


Figure 2.4: Asset prices and real variables as functions of ω . (a) Risk-free rate, and risk premium. (b) Output gap and employment. In both panels the red shading marks the crisis regime, the orange shading marks the recovery regime, and the green shading marks the unconstrained regime. The vertical dashed line marks ω_c . Calibrated parameters from Table 2.1.

Output gap. The log output gap relative to the frictionless benchmark $\omega_n \equiv 1/\epsilon$ is:

$$x_t \equiv \ln Y_t - \ln Y_{n,t} = \frac{1}{1+\varphi} \ln \left(\frac{1-\omega_n}{1-\omega_t} \right). \quad (2.29)$$

2.5 The Intermediation Gap

2.5.1 A Three-Act Narrative

The results in this section are organised around three acts that trace the economy's trajectory through a crisis:

1. **Before the crisis** ($\omega > \omega_c$, no binding constraint): Equilibrium exists. Asset prices exhibit a positive intermediation gap with gap-to-price ratio $b(\omega)$. As intermediary capital erodes, the intermediation gap grows convexly—the exact formula gives $b(\omega_c) \approx 0.02$ at the crisis threshold, rising to $b(0.04) = 0.17$ and $b(\underline{\omega}) \approx 0.38$ in severe distress—partially masking the deterioration of fundamental values: market prices overstate the health of the financial sector (Section 2.6.2).
2. **During the crisis** (ω drops below ω_c): Leverage constraints bind. Without QE, no equilibrium exists. The question of asset pricing is moot.
3. **After QE** (QE active): Equilibrium is restored. The intermediation gap vanishes: the stock price equals its fundamental value. The sudden disappearance of the intermediation gap reveals the true severity of intermediary undercapitalisation. The economy may be trapped at the bad steady state with correct prices but depressed output. Fiscal recapitalisation must be calibrated to these revealed fundamentals, not to pre-crisis prices inflated by the intermediation gap.

The three acts trace an information arc: the intermediation gap obscures fundamentals before the crisis, its disappearance reveals fundamentals during the crisis, and correct post-QE prices enable—but do not substitute for—fiscal intervention. We now formalise each act.

The inverse wealth share $Q_t = 1/\omega_t$ is a non-negative local martingale (Proposition 2.2). Whether Q is a strict or true local martingale determines whether equilibrium asset prices exhibit a positive intermediation gap—the component of asset prices above fundamental value attributable to the concentration of risk-bearing in undercapitalised intermediaries. The inter-

mediation gap is a symptom of intermediary weakness, not monetary accommodation: it is largest when the financial sector is most distressed.

2.5.2 Martingale Criterion

For a one-dimensional diffusion $dX = \sigma_X(X) dZ$ starting at $X_0 > 0$, X is a strict local martingale—meaning $\mathbb{E}[X_t] < X_0$ for some t —if and only if (Engelbert and Schmidt, 1985; Mijatović and Urusov, 2012):

$$\int_{X_0}^{\infty} \frac{x}{\sigma_X(x)^2} dx < \infty. \quad (2.30)$$

When the integral diverges, X is a true martingale and $\mathbb{E}[X_t] = X_0$ for all t .

2.5.3 Connection to the Rational Bubble Literature

Hugonnier (2012) proves that in a two-agent economy with a portfolio constraint, equilibrium asset prices contain rational bubbles if and only if the weighting process—the ratio of marginal utilities—is a strict local martingale. In our model the weighting process reduces to

$$\lambda_t = \frac{1 - \omega_t}{\omega_t} = Q_t - 1, \quad (2.31)$$

so λ inherits the martingale classification of Q . The following equivalence is a direct consequence of Hugonnier (2012, Theorem 1):

$$Q = 1/\omega \text{ is a strict local martingale} \iff \text{intermediation gap} > 0. \quad (2.32)$$

2.5.4 The Exact Gap Formula in the Production Economy

The intermediation gap satisfies a singular Sturm–Liouville boundary-value problem that characterises it exactly, without recourse to the constant-weighting-process approximation of Hugonnier (2012)’s exchange economy.

Proposition 2.3 (Intermediation gap formula) *In the baseline model (limited participation, no leverage constraint), the intermediation gap-to-price ratio $b(\omega_t) \equiv B_t/P_{D,t}$ satisfies the second-order linear ODE*

$$b''(\omega) = \frac{2\delta}{\sigma_\omega^2(\omega)} b(\omega), \quad \omega \in (0, 1/\epsilon), \quad (2.33)$$

with boundary conditions

$$b(0^+) = 1, \quad b(1/\epsilon) = 0. \quad (2.34)$$

The BVP admits a unique C^2 solution, which is strictly positive, strictly decreasing, and strictly convex on $(0, 1/\epsilon)$.

The proof is in Online Appendix [BO.5](#). The derivation proceeds in three steps. First, apply the Feynman–Kac theorem to the discounted resolvent $f(\omega) \equiv \mathbb{E}^\omega[\int_0^\infty e^{-\delta u} (1/\omega_u) du]$, which gives $\frac{1}{2}\sigma_\omega^2 f'' + (\sigma_\omega^2/\omega)f' - \delta f + 1/\omega = 0$. Second, change variables to $b \equiv 1 - \delta\omega f$: the b' and $(1-b)/\omega^3$ coefficients cancel identically because of the Q -martingale identity $\mu_\omega = \sigma_\omega^2/\omega$, leaving only (2.33). Third, establish the boundary conditions: $b(1/\epsilon) = 0$ follows from the absorbing-state structure of the good steady state; $b(0^+) = 1$ follows from the h -transform analysis of the strict local martingale $Q = 1/\omega$ (Pal and Protter, 2010; Delbaen and Schachermayer, 1995), which delivers complete mass loss at infinity. The economic content is the standard decomposition $P_{D,t} = P_{D,t}^{\text{fund}} + B_t$, with the fundamental value $P_{D,t}^{\text{fund}} = \mathbb{E}_t[\int_t^\infty (\Lambda_s/\Lambda_t) D_s ds]$ and the bubble B_t the strict-local-martingale residual in the sense of Hugonnier (2012).

Figure 2.5 plots the exact numerical solution of (2.33)–(2.34) at baseline calibration, with key values marked. The gap rises convexly as ω falls: $b(\omega_c) \approx 0.02$ at the crisis threshold, $b(0.04) = 0.17$, and $b(\underline{\omega}) \approx 0.38$ in severe distress. The full numerical procedure, cross-validation against a second independent solver, and comparison to the exchange-economy closed form of Hugonnier (2012, Proposition 6) are in Online Appendix [BO.5](#).

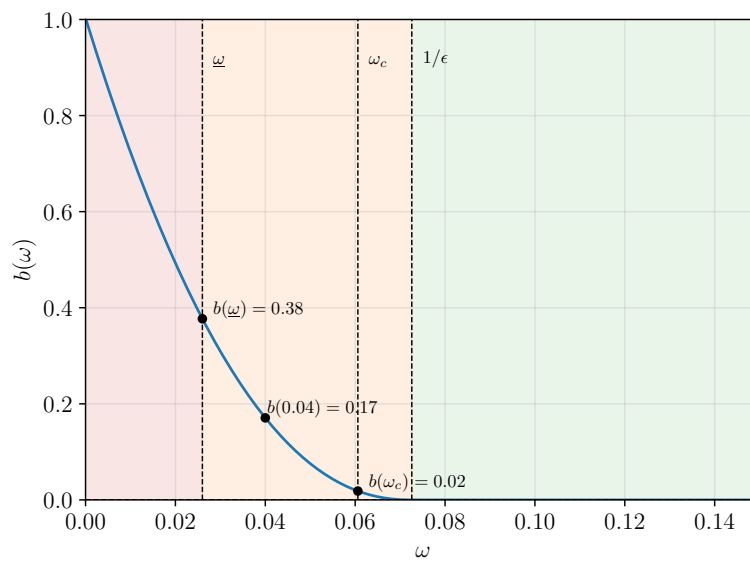


Figure 2.5: Exact intermediation gap-to-price ratio $b(\omega)$, computed by numerical solution of (2.33)–(2.34). Vertical dashed lines mark $\underline{\omega}$, ω_c , and $1/\epsilon$. Calibrated parameters from Table 2.1.

Volatility paradox. The exact ODE (2.33) makes the volatility paradox direct: the right-hand-side coefficient $2\delta/\sigma_\omega^2(\omega)$ is *decreasing* in σ_ω , so lower fundamental volatility produces a *steeper* second derivative of the gap function and hence a larger gap at any given ω . In the Hugonnier closed form this appears as the decreasing-in- v exponent $\frac{\partial \gamma}{\partial v} = \frac{-4\delta}{v^3 \sqrt{1+8\delta/v^2}} < 0$. The economic intuition is common to both: low fundamental volatility means σ_ω is small, making the inverse wealth share Q more prone to explosive excursions (the Mijatović–Urusov integral test in (2.30) converges faster), which is exactly when the strict local martingale property bites hardest.⁶

2.5.5 The Intermediation Gap with and without Leverage Constraints

Theorem 2.6 (Intermediation gap classification)

- (i) **Baseline model** (limited participation, no leverage constraint): In the controlled regime $(0, 1/\epsilon)$, $Q = 1/\omega$ is a **strict local martingale**. Equilibrium asset prices exhibit a positive intermediation gap.
- (ii) **Constrained model** (leverage constraint, QE active): $Q = 1/\omega$ is a **true martingale** in all three regimes, crisis $(0, \underline{\omega})$, recovery $(\underline{\omega}, 1/\epsilon)$, and expansionary $(1/\epsilon, 1)$. Equilibrium asset prices are free of any intermediation gap.

The proof is in Appendix B.11. The entire difference comes from one factor of ω in the volatility function near $\omega = 0$. In the baseline, $\phi = 1/(\epsilon\omega) \rightarrow \infty$, so the ω in $\sigma_\omega = \omega(\phi - 1)\sigma_R$ cancels, leaving σ_ω positive at the boundary. With the constraint, $\phi \leq 1 + \lambda(\omega - \underline{\omega})$ is bounded, so $\sigma_\omega \propto \omega$.

	σ_ω near $\omega = 0$	σ_Q near $Q = \infty$	Integral test	Gap?
Baseline	$\sigma_\omega^0 \rightarrow c_0 > 0$	$c_0 Q^2$	$\int q^{-3} dq < \infty$	Yes
Constrained	$ \sigma_\omega \sim c'_0 \omega$	$c'_0 Q$	$\int q^{-1} dq = \infty$	No

2.5.6 The Leverage Paradox

The intermediation gap classification and the equilibrium failure result are two manifestations of a single mathematical object: the behaviour of σ_ω at the origin.

⁶The model is silent on reach-for-yield effects: the household's bond demand is perfectly inelastic (log utility), so QE has no portfolio rebalancing effect. In richer models, lower safe returns push investors into risky assets, creating a distinct form of overvaluation not connected to the equilibrium existence problem.

Proposition 2.4 (Leverage paradox)

- (i) **Baseline economy** (no leverage constraint): competitive equilibrium exists for all $\omega \in (0, 1/\epsilon)$ and $Q = 1/\omega$ is a strict local martingale—the intermediation gap is positive.
- (ii) **Constrained economy** (leverage constraint (2.7) with QE): $Q = 1/\omega$ is a true martingale in all three regimes—no intermediation gap.
- (iii) **Impossibility.** The combination of gap-free asset prices and unconditional equilibrium existence (without central bank intervention) is impossible: for any leverage function $\bar{\phi} : (0, 1/\epsilon) \rightarrow \mathbb{R}_+$, if competitive equilibrium exists for all ω without intervention, then a positive intermediation gap exists.

The intuition is that σ_ω near $\omega = 0$ carries a double duty. In the unconstrained model, $\phi \sim 1/(\epsilon\omega) \rightarrow \infty$ as $\omega \rightarrow 0$, so $\sigma_\omega = \omega(\phi - 1)\sigma_R$ remains bounded away from zero at the boundary—which, by the integral test, is exactly what makes $Q = 1/\omega$ a strict local martingale and produces the intermediation gap. In the constrained model, $\phi \leq \bar{\phi}(\omega)$ is bounded, so $\sigma_\omega \propto \omega \rightarrow 0$ at the boundary—making Q a true martingale and closing the intermediation gap. The same economic force (bounded leverage) that prevents ϕ from diverging—which would otherwise allow the financier to absorb all equity and prevent market failure—also changes the stochastic classification of Q at the boundary. Correcting prices and clearing markets are two sides of the same coin.

Proof. Parts (i) and (ii) restate Theorem 2.6. For part (iii), suppose competitive equilibrium exists for all $\omega \in (0, 1/\epsilon)$ without central bank intervention. Equity market clearing requires $\phi(\omega) \omega P_Y = P_D = P_Y/\epsilon$, so $\phi(\omega) = 1/(\epsilon\omega)$. The constraint cannot bind at any ω —otherwise $\phi = \bar{\phi}(\omega) < 1/(\epsilon\omega)$ and the equity market does not clear—so $\bar{\phi}(\omega) \geq 1/(\epsilon\omega)$ for all ω . With $\phi = 1/(\epsilon\omega)$, the wealth-share diffusion coefficient is

$$\sigma_\omega(\omega) = \frac{(1/\epsilon - \omega)\sigma_R}{D_u(\omega)}, \quad D_u(\omega) = \left(1 - \frac{1}{\epsilon}\right) + \phi(1 - \omega) > 0,$$

giving $\sigma_\omega(0^+) = \sigma_R/(\epsilon D_u(0)) > 0$. Converting to $Q = 1/\omega$: $\sigma_Q = \sigma_\omega/\omega^2 \sim c_0 Q^2$ with $c_0 = \sigma_\omega(0^+) > 0$. The integral test (2.30) yields $\int_{Q_0}^\infty q \sigma_Q(q)^{-2} dq \sim c_0^{-2} \int q^{-3} dq < \infty$, so Q is a strict local martingale and the intermediation gap is positive. $\square$

Corollary 2.1 (Impossibility) *Within the class of economies in this paper:*

- (a) *Any economy with unconditional equilibrium existence necessarily has a positive intermediation gap.*
- (b) *Any economy with gap-free asset prices is vulnerable to equilibrium failure.*

Proof. Part (a) is Proposition 2.4(iii). Part (b) is its contrapositive: if equilibrium existed for all ω without intervention, Proposition 2.4(iii) would imply a positive intermediation gap, contradicting the hypothesis of gap-free prices. $\square$

The policy package of QE plus fiscal recapitalisation resolves the tradeoff: QE restores equilibrium existence when bounded leverage causes market failure, while the leverage constraint simultaneously closes the intermediation gap. The exact gap ODE provides a constructive proof of closure: on $[\underline{\omega}, 1/\epsilon]$, the only solution to $b'' = (2\delta/\sigma_\omega^2)b$ with $b(\underline{\omega}) = b(1/\epsilon) = 0$ is $b \equiv 0$ (by the maximum principle). Fiscal recapitalisation then moves the economy to the good steady state where the constraint ceases to bind, the central bank exits, and both equilibrium existence and correct pricing are achieved without intervention.

2.5.7 The Intermediation Gap as a Sufficient Statistic for the Output Gap

Both the intermediation gap and the output gap are driven by the same underlying shortage of risk-bearing capacity—the *risk-bearing gap*. Since both are monotone functions of the single state variable ω_t , the intermediation gap-to-price ratio is a sufficient statistic for the output gap:

$$x_t = \frac{1}{1+\varphi} \ln\left(\frac{1-\omega_n}{1-\omega_t}\right) = G(b(\omega_t)), \quad (2.35)$$

for an increasing function G . Larger intermediation gaps signal larger output gaps. The exchange-economy approximation $(1-\omega)^\gamma$ of Hugonnier (2012), being nearly linear over the empirical range, collapses G to an approximately linear relationship, discarding the model's predicted curvature. The exact ODE solution $b(\omega)$ from Proposition 2.3 preserves it.

Since T^* and the benefit-to-cost ratio (2.42) are also functions of ω , the pre-crisis gap-to-price ratio encodes the fiscal cost of future intervention via $\omega_0 = b^{-1}(B/P_D)$. Crucially, this mapping is available only *before* QE. Once the intermediation gap vanishes, the policymaker loses this market-based diagnostic—the same intervention that corrects asset prices destroys their informational content about the output gap.

2.6 Real Effects and Policy

2.6.1 The Unified State Variable

The financier's wealth share ω_t simultaneously governs three equilibrium objects:

Asset prices and the intermediation gap:

$$\frac{B_t}{P_{D,t}} = b(\omega_t), \quad b''(\omega) = \frac{2\delta}{\sigma_\omega(\omega)^2} b(\omega). \quad (2.36)$$

Employment:

$$N_t = \left(\frac{1 - 1/\epsilon}{1 - \omega_t} \right)^{\frac{1}{1+\varphi}}, \quad \text{with } N'(\omega_t) > 0. \quad (2.37)$$

Output:

$$Y_t = A_t N_t = A_t \left(\frac{1 - 1/\epsilon}{1 - \omega_t} \right)^{\frac{1}{1+\varphi}}. \quad (2.38)$$

When intermediary wealth erodes (ω_t falls), three things happen together: the intermediation gap grows (convexly), employment falls, and output contracts. The convexity of the intermediation gap function transmits directly to real variables: when ω_t is low, small further declines produce large increases in the intermediation gap *and* large employment losses, creating a fragility trap that intensifies as conditions deteriorate.

2.6.2 Phantom Wealth

In the baseline economy, the equity price decomposes into a fundamental value and an intermediation gap:

$$P_D = P_D^{\text{fund}} + B, \quad P_D^{\text{fund}} = P_D [1 - b(\omega)], \quad B = P_D b(\omega). \quad (2.39)$$

Proposition 2.5 (Phantom wealth) *In the baseline model:*

- (i) $P_D^{\text{fund}}(\omega) \rightarrow 0$ as $\omega \rightarrow 0$: at the point of maximum crisis, the entire equity market value is an intermediation gap.
- (ii) The fundamental-to-market ratio $P_D^{\text{fund}}/P_D = 1 - b(\omega)$ is strictly increasing in ω .
- (iii) The gap ratio $B/P_D = b(\omega)$ overstates the financier's equity position by a factor that grows convexly as the crisis deepens.

Proof. (i) $P_D(0) = AN(0)/(\epsilon\delta) > 0$ and $b(0) = 1$, so $P_D^{\text{fund}}(0) = 0$. (ii) $b' < 0$ follows from $b > 0$, $b'' > 0$, $b(0) = 1$, $b(1/\epsilon) = 0$. (iii) Convexity: $b'' = (2\delta/\sigma_\omega^2)b > 0$ whenever $b > 0$. $\square$

Part (i) is the key result for policy. A policymaker observing pre-crisis market prices sees positive equity values and may conclude the financial sector retains substantial wealth. In fact, as ω approaches zero, fundamental value vanishes and the entire observed price is phantom wealth—value that would disappear if the economy transitioned to a regime with correct prices. The exact gap formula quantifies this: at $\omega = 0.04$, the gap-to-price ratio is $b \approx 0.17$; at $\omega = \underline{\omega}$, $b \approx 0.38$. Because the required recapitalisation (2.41) is calibrated to the *fundamental* state variable ω , not to market prices that include the gap, policymakers who rely on market-based measures systematically underestimate the severity of undercapitalisation. The sufficient-statistic property extends to the fiscal transfer:

$$T^*(\omega_0) = \Omega(\underline{\omega}) - \Omega(\omega_0), \quad \omega_0 = b^{-1}\left(\frac{B}{P_D}\right), \quad (2.40)$$

where b^{-1} is the inverse of the (strictly decreasing) exact gap function. Crucially, this mapping is available only before QE: once the gap vanishes, the policymaker loses this market-based diagnostic.

2.6.3 Large-Scale Asset Purchases

Under Assumption 2.3, the central bank absorbs the residual equity supply that the constrained financier cannot hold, preventing equilibrium failure but not altering the economy's long-run dynamics.

Proposition 2.6 (Central bank balance sheet at the bad steady state) *At $\omega = \underline{\omega}$, the central bank permanently holds equity worth $(1 - \epsilon\underline{\omega}) Y_t/(\epsilon\delta)$ and earns the equity premium σ^2 on these holdings.*

At the bad steady state, the central bank holds a potentially large share of the equity market: with $\underline{\omega} = 0.0260$ and $\epsilon = 13.7741$, it holds about 64%. The equity premium equals its frictionless value σ^2 because $\sigma_\omega(\underline{\omega}) = 0$ eliminates all wealth redistribution dynamics.

2.6.4 Fiscal Recapitalisation

Proposition 2.7 (Recapitalisation) *Suppose $\omega_0 < \underline{\omega}$ (crisis regime). A one-time capital injection that raises the financier's wealth share to $\omega_{0+} > \underline{\omega}$ places the economy in the recovery regime, which then converges almost surely to the good steady state $1/\epsilon$ without further intervention.*

Remark 2.4 (Transfer size with general equilibrium feedback) *Because total wealth $P_Y(\omega) = AN(\omega)/\delta$ is increasing in ω , a dollar transfer raises ω by less than $T/P_Y(\omega_0)$. The transfer required to reach target ω^* is:*

$$T^* = \frac{A}{\delta} [\omega^* N(\omega^*) - \omega_0 N(\omega_0)], \quad (2.41)$$

which exceeds the naïve estimate by $\omega^*[P_Y(\omega^*) - P_Y(\omega_0)]$, though the correction is small under baseline calibration (approximately 1.1% for a transfer from zero to $\underline{\omega}$). The correction is small because log utility pins total wealth at $P_Y = Y/\delta$: the general equilibrium feedback operates only through the employment response $N(\omega)$, which is close to linear over the relevant range.

Employment multiplier. Recapitalisation from ω_0 to $\omega_0 + \Delta$ produces an immediate employment gain of $[(1 - \omega_0)/(1 - \omega_0 - \Delta)]^{1/(1+\varphi)} - 1$. The employment function $N(\omega)$ is convex, so the output gap accelerates as intermediary capital falls: each unit decline in ω produces a larger employment loss than the previous one, making the cumulative benefit of escaping Permanent QE or the QE Trap larger for deeper crises. Under the baseline calibration, the permanent long-run gain from escaping either regime is about 2.1% of efficient employment. The formal statement and proof are in Appendix [BO.7](#) (Corollary [BO.1](#)).

Cost–benefit analysis. The steady-state benefit-to-cost ratio of recapitalisation is:

$$\frac{\text{Steady-state output gain}}{\text{Cost}} = \frac{N(1/\epsilon) - N(\underline{\omega})}{\underline{\omega} N(\underline{\omega}) - \omega_0 N(\omega_0)}. \quad (2.42)$$

Under baseline calibration, the same worst-case comparison gives a ratio of about 0.8: the government spends approximately 2.6% of total wealth to permanently raise output by about 2.1%. Since convergence to $1/\epsilon$ has infinite expected hitting time (Theorem [2.4](#)), the discounted present value of the output gain is lower; the ratio above should be interpreted as a long-run comparison of steady states. The fiscal cost of delay is not a larger transfer but a larger cumulative output loss; each period in the crisis regime forgoes output on the steep portion of the employment curve, so delays at the onset are most costly.

Welfare. The household weakly prefers the bad steady state (higher consumption share $1 - \underline{\omega}$), so recapitalisation is not a Pareto improvement; it involves redistribution from household to financier. The detailed welfare calculation and moral hazard analysis are in Appendix [BO.7.2](#).

2.6.5 Mapping to Policy

The model rationalises three distinct phases of crisis intervention:

1. *Large-scale asset purchases (central bank)*: when $\omega < \omega_c$ and leverage constraints bind, the central bank purchases equity to prevent market failure—corresponding to the Fed’s emergency facilities in 2008–2009.
2. *Recapitalisation (fiscal authority)*: the Treasury injects capital to push ω above $\underline{\omega}$, as in TARP (October 2008); European recapitalisations followed in 2010–2014, with the delay producing the persistent stagnation the model predicts.
3. *Taper and exit*: once $\omega > \underline{\omega}$ and recovery begins, the leverage constraint eventually stops binding and the central bank’s equity holdings decline to zero.

Phase 1 alone is insufficient: QE maintains equilibrium existence but cannot alter long-run dynamics. Because employment and the intermediation gap share a common driver, policies that stabilise ω_t simultaneously stabilise financial valuations and employment—financial stability and full employment are complementary objectives.

2.6.6 Empirical Support

The model’s predictions are supported by the empirical record. In the United States, TARP injected \$245 billion within months of the crisis; [Berger and Roman \(2017\)](#) find it significantly increased net job creation through the credit-supply channel that raises ω in our model. In Europe, bank recapitalisations were delayed by 2–4 years ([Brunnermeier et al., 2016](#)), and the euro area experienced precisely the persistent stagnation the model predicts from Phase 1 without Phase 2. Across 69 systemic banking crises since 1980, [Homar and van Wijnbergen \(2017\)](#) find that the *speed* of recapitalisation—not merely its occurrence—determines recovery duration, consistent with the absorbing steady-state structure and the convex employment losses from delay.

2.7 Empirical Evidence

This section tests empirically the main implications of the model. The objective is not to provide a fully identified causal evaluation of each aspect of the model, but to assess whether the

data display the mappings, state dependence, and policy substitution patterns implied by the theory. We proceed in two steps. We first describe the data and calibration that map the model's objects into observables. We then present five sets of tests: Test 1 studies whether intermediary capital is linked to current and future real activity; Test 2 examines dealer intermediation and central-bank substitution in Treasury and agency MBS markets; Test 3 asks whether the exact pre-QE intermediation gap is a sufficient statistic for real slack; Test 4 uses cross-country annual data to evaluate the role of fiscal recapitalization in financial and macroeconomic recovery; and Test 5 studies whether the post-2007 period is associated with a persistent shift to a lower-capital regime.

2.7.1 Data

The empirical analysis is based mainly on U.S. quarterly data. The model's central state variable is the intermediary's wealth share, ω_t , which we proxy with the intermediary capital ratio of [He et al. \(2017\)](#). Their measure is constructed for New York Fed primary-dealer holding companies as aggregate market equity divided by aggregate market equity plus book debt. It is therefore a market-based measure of intermediary balance-sheet strength, with lower values indicating weaker risk-bearing capacity.

For the calibration, we use productivity growth constructed from nonfarm business labor productivity (OPHNFB), S&P 500 excess returns, and the S&P 500 price and dividend series from Robert Shiller's official long-run stock-market data. These quarterly calibration moments use the pre-crisis sample 1970Q1–2007Q3. To discipline the goods-market parameter ϵ , we use the annual finance-and-insurance value-added share of GDP from the U.S. Bureau of Economic Analysis industry accounts; the calibration uses the 1997–2006 annual sample.

All data used in Tests 1–3 and Test 5 are quarterly. The common empirical proxy for the model's state variable is again the HKM intermediary capital ratio. For real activity, we use the U.S. output gap, defined as real GDP (GDPC1) relative to CBO potential GDP (GDPPOT), the unemployment gap, defined as the unemployment rate (UNRATE) minus the CBO natural rate (NROU), and growth rates of payroll employment (PAYEMS) and industrial production (INDPRO). The baseline quarterly controls are productivity growth and CPI inflation (CPIAUCSL); robustness exercises also use the federal funds rate (DFF), the term spread (DGS10 minus DTB3), and the credit spread (BAA minus AAA). Test 3 additionally uses the cyclically adjusted price–earnings ratio from Robert Shiller's official data page as the benchmark valuation predictor. The under-

lying non-dealer U.S. quarterly sample runs from 1970Q1 to 2024Q4 when available. Test 1 uses the post-1995 baseline sample; the pre-QE Test 3 specifications end in 2007Q4; and Test 5 uses the full 1970Q1–2024Q4 HKM sample.

Test 2 combines these macro series with measures of private and public risk absorption in Treasuries and agency MBS. On the private side, dealer financing to clients and dealer gross positions are taken from the Federal Reserve’s dealer-capacity dataset and scaled by the stock of securities held by the public excluding SOMA (Cochran et al., 2024). The dealer-financing ratios are available from 2003Q1 through 2024Q1 and are used in Test 2A and in the longer-sample Test 2B robustness exercise. The dealer gross-position ratios are available from 2012Q1 through 2024Q1 and are used in baseline Test 2B. On the public side, we construct $RB_t^{CB,j}$ as a duration-weighted stock of Fed holdings in asset class j . Specifically, we start from weekly SOMA CUSIP-level holdings, assign each security a duration, multiply holdings by duration, aggregate across CUSIPs within Treasuries and agency MBS, take quarter-end observations, and normalize by nominal GDP (GDP). The resulting central-bank risk-bearing series are available from 2003Q3 through 2024Q4. Appendix B.12.1 reports the detailed calculations used to compute the duration and the resulting $RB_t^{CB,j}$ measure.

Test 4 is the only exercise implemented in annual panels. The estimation sample is a balanced annual panel of ten countries over 2003–2020: the United States, the United Kingdom, Germany, France, Italy, Spain, Ireland, Greece, Portugal, and the Netherlands. Intermediary capital is measured by bank equity-to-assets ratios: for the United States we use the FRED EQTA series, while for the United Kingdom and euro-area countries we use ECB balance-sheet data, preferring the BSI capital-and-reserves-to-assets measure and using CBD2 balance-sheet data when coverage is better in the crisis window. Annual macro outcomes—the output gap, unemployment rate, and real GDP growth—come from the *OECD Economic Outlook* (Organisation for Economic Co-operation and Development, 2025). QE is measured as central-bank total assets relative to nominal GDP, using BIS balance-sheet data and OECD GDP for the relevant currency areas; for euro-area countries, the QE measure is the ECB balance sheet scaled by euro-area GDP. Recapitalization flows and stocks are assembled from official program-level sources: TARP transactions for the United States, cleaned HM Treasury and UKFI records for the United Kingdom, and Eurostat supplementary tables on government interventions to support financial institutions for euro-area countries.

Unless otherwise noted, the U.S. macroeconomic and interest-rate series used both in the

calibration and in the quarterly tests are taken from the Federal Reserve Economic Data database, and the specific series are identified by their mnemonics.

2.7.2 Calibration

Before turning to the empirical tests, we calibrate the model parameters to a parsimonious set of macro-financial moments, and then hold the resulting parameters fixed throughout the rest of the section.

First, we anchor the crisis region to the U.S. experience around the Great Financial Crisis. We set the crisis threshold ω_c to the average value of the HKM intermediary capital ratio between 2007Q3 and 2007Q4. This choice places the threshold at the transition between the last pre-crisis observation and the first quarter of acute crisis stress, so that ω_c captures the point at which the leverage constraint begins to bind. We set the PCA threshold $\underline{\omega}$ to the minimum value of the HKM intermediary capital ratio observed between 2008Q1 and 2009Q2, corresponding to the trough of intermediary capitalization during the Great Financial Crisis.

The remaining parameters are disciplined directly by the model's equilibrium restrictions. We use the pre-crisis sample 1970Q1–2007Q3 for the quarterly series, and 1997–2006 for the annual series. These are the longest available samples ending before the onset of the financial crisis. Indeed, observations after 2007Q3 are outcomes of the crisis and policy regime whose effects the model seeks to explain. We set the drift μ equal to the annualized average growth rate of labor productivity, and set the volatility parameter σ equal to the annualized volatility of S&P 500 excess returns. Since the model implies $P_{D,t} = Y_t/(\epsilon\delta)$ and $D_t = Y_t/\epsilon$, the dividend-price ratio satisfies $D_t/P_{D,t} = \delta$; we therefore set δ equal to the average dividend yield on the S&P 500. Similarly, because $D_t/Y_t = 1/\epsilon$, we set ϵ equal to the inverse of the average finance-sector value-added share of GDP. Given $(\epsilon, \omega_c, \underline{\omega})$, the leverage-sensitivity parameter λ is implied by the threshold condition $\lambda\omega_c^2 + (1 - \lambda\underline{\omega})\omega_c - 1/\epsilon = 0$. Similarly, given (μ, σ, δ) , the benchmark risk-free rate is pinned down by $r = \delta + \mu - \sigma^2$. Finally, we set the inverse Frisch elasticity φ equal to the inverse of the aggregate-hours Frisch elasticity of 0.75 recommended by [Chetty et al. \(2011\)](#). Table 2.1 reports the resulting parameter values.

Table 2.1: Calibration

Parameter	Value	Empirical target
ω_c	0.0606	Average HKM intermediary capital ratio, 2007Q3 and 2007Q4
$\underline{\omega}$	0.0260	Minimum HKM intermediary capital ratio, 2008Q1–2009Q2
μ	0.0201	Annualized average labor-productivity growth, 1970Q1–2007Q3
σ	0.1505	Annualized volatility of S&P 500 excess returns, 1970Q1–2007Q3
δ	0.0314	Mean S&P 500 dividend yield, 1970Q1–2007Q3
ϵ	13.7741	Inverse average finance-sector value-added share of GDP, 1997–2006
λ	5.7231	Implied by the crisis-threshold condition (2.10)
r	0.0289	Implied by the benchmark risk-free rate equation (B.1)
φ	1.3333	Inverse aggregate-hours Frisch elasticity from Chetty et al. (2011)

Notes: Quarterly moments use the pre-crisis sample from 1970Q1 to 2007Q3. Productivity growth and S&P 500 excess-return volatility are annualized from quarterly data. The crisis threshold ω_c is the average of the HKM intermediary capital ratio in 2007Q3 and 2007Q4. The lower threshold $\underline{\omega}$ is the minimum HKM ratio observed from 2008Q1 to 2009Q2. The parameter ϵ is disciplined by annual finance-sector value added over 1997–2006. The benchmark risk-free rate is not estimated separately; it is implied by the calibration identity. The parameter φ is the inverse of the aggregate-hours Frisch elasticity.

Figure 2.6 plots the HKM intermediary capital ratio together with the three calibrated state values. The upper dashed line, $\omega = 1/\epsilon$, is the good steady state at which the financier can hold 100% of her wealth in equity, without leverage. The middle dashed line, $\omega = \omega_c$, is the crisis threshold at which required leverage equals the leverage cap. The lower dashed line, $\omega = \underline{\omega}$, is the bad steady state. The series falls sharply through ω_c during the Great Financial Crisis, remains below the good-state benchmark for much of the post-2007 period, and declines again around the COVID-19 shock.

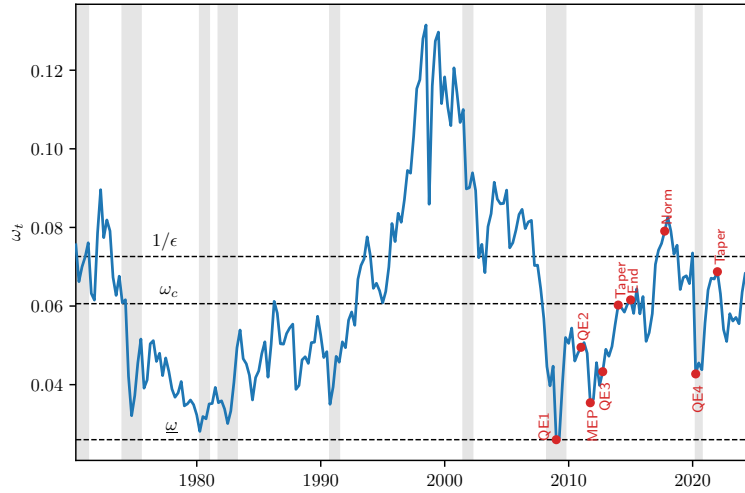


Figure 2.6: HKM intermediary capital ratio. The blue line is the quarterly HKM intermediary capital ratio, gray shading denotes NBER recessions, and the dashed horizontal lines mark $\underline{\omega}$, ω_c , and $1/\epsilon$. Red dots mark salient points in the history of QE policy.⁷

⁷LSAP1, announced on November 25, 2008 for agency debt and agency MBS purchases and expanded on March 18, 2009 to include longer-term Treasuries; LSAP2, announced on November 3, 2010 for additional longer-term Treasury purchases; the Maturity Extension Program, announced on September 21, 2011 and extended on June 20, 2012; LSAP3, announced on September 13, 2012 as open-ended agency MBS purchases and broadened on December 12, 2012 to continue longer-term Treasury purchases; tapering, announced on December 18, 2013; the end of net

2.7.3 Test 1: Real Economic Activity

Test 1A The first model implication we test is that intermediary capital maps nonlinearly into real activity (see (2.29)). This mapping holds in any regime in which equilibrium exists; with QE, it holds even when the leverage constraint binds. If the model is informative about real activity, the model-implied output gap should line up with the real-activity gap observed in the data. Therefore, we regress empirical proxies for real activity on the model-implied output gap:

$$x_t^{data} = \alpha + \beta x_t^{model} + \varepsilon_t,$$

where

$$x_t^{model} = \frac{1}{1 + \varphi} \ln \left(\frac{1 - \omega_n}{1 - \omega_t} \right).$$

We implement this test on the modern quarterly sample 1995Q1–2024Q4 for two measures of real economic activity, namely the output gap and the unemployment gap, with the latter signed so that larger values correspond to stronger real activity. We use the post-1995 sample as the main specification because it covers the modern macro-financial regime most relevant for the paper: it includes the pre-GFC expansion, the GFC, the QE period, and the post-QE observations, while avoiding the need to impose a single level mapping across earlier financial regimes. The coefficient of interest is β . A positive estimate implies that lower intermediary capital is associated with weaker real activity in the data, that is, a lower output gap and a larger unemployment gap. Moreover, an estimate close to one indicates that the model not only gets the sign right, but also captures the magnitude of the mapping reasonably well. We report robustness checks using alternative sample windows in Online Appendix BO.15.1. Table 2.2 shows the baseline results.

The results support the model’s mapping from intermediary capital to real activity. In both specifications, the coefficient on the model-implied gap is positive and statistically significant. For the output gap, the estimate is 1.040, while for the sign-reversed unemployment gap, the estimate is 0.954. Both coefficients are close to one, indicating that the theoretical mapping

purchases, announced on October 29, 2014; balance-sheet normalization, announced on September 20, 2017; COVID-era LSAPs, announced on March 15, 2020 and expanded on March 23, 2020; and tapering of COVID-era purchases, announced on November 3 and December 15, 2021.

Table 2.2: Test 1A – theory-implied level fit, output gap

	Output gap	- Unemployment gap
α	-0.004* (0.003)	-0.007*** (0.002)
β	1.040*** (0.191)	0.954*** (0.210)
Adj. R^2	0.348	0.334
N	120	120

Notes: HAC standard errors with four lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

delivers the right sign and captures the magnitude of the relation reasonably well. This is consistent with the prediction that lower intermediary capital is associated with weaker real activity. The model-implied gap explains about one third of the variation in both slack measures, with adjusted R^2 equal to 0.348 for the output gap and 0.334 for the unemployment gap. Taken together, these results suggest that the model's state variable contains meaningful information about the level of real activity.

Test 1B While the model's primitive implication is a contemporaneous mapping between the state variable and real activity, and the model does not by itself imply predictability, we now ask whether the same state variable also contains dynamic information about future real activity. Moreover, the model suggests that the state variable follows a regime-dependent law of motion (see Theorem 2.2), and the state space is partitioned into two invariant regions. Therefore, we also ask ourselves whether the predictive relation between the state variable and future real activity differs above and below the crisis threshold. We thus complement the level-fit evidence in Test 1A with h -step predictive regressions of future real activity on the model-implied output gap and its low-capital interaction.⁸

$$y_{t+h} = \alpha_h + \rho_h y_t + \beta_h z_t + \kappa_h z_t \mathbf{1}\{\hat{\omega}_t < \hat{\omega}_c\} + \Gamma'_h X_t + \varepsilon_{t+h},$$

where

$$z_t = \hat{x}_t^{model} - \hat{x}_c^{model}, \quad \hat{x}_t^{model} = \frac{1}{1 + \varphi} \log \left(\frac{1 - \omega_n}{1 - \hat{\omega}_t} \right), \quad \hat{x}_c^{model} = \hat{x}^{model}(\hat{\omega}_c).$$

We implement this test for horizons $h \in \{1, 2, 4\}$ quarters in the same post-1995 sample for

⁸This is not a causal impulse-response exercise. It is an h -step predictive specification motivated by the persistence of intermediary capital and by the regime-dependent law of motion for ω_t .

four measures of real economic activity, namely the output gap, the negative unemployment gap, cumulative payroll growth, and cumulative industrial production growth.

Including y_t controls for the strong persistence in real activity. The coefficient β_h is the high-capital slope, while $\beta_h + \kappa_h$ is the low-capital slope. The coefficient κ_h therefore tests whether the predictive slope differs below the crisis threshold (i.e. we test for the equality of low- and high-capital slopes, $H_0 : \kappa_h = 0$). For all four outcomes, which are coded so that larger values correspond to stronger real activity, the model predicts a positive low-capital slope, $\beta_h + \kappa_h > 0$: when the model-implied output gap is lower because intermediary capital is lower, future real activity should also be weaker. If the low-capital state is quantitatively important, the model also suggests a positive interaction term, $\kappa_h > 0$, so that the predictive relation is steeper below the crisis threshold. The high-capital slope β_h is expected to be weaker, because intermediary balance sheets are not constrained above the threshold.

We keep the controls deliberately parsimonious because many macro-financial variables are themselves endogenous responses to intermediary distress and potential transmission channels from intermediary capital to real activity; controlling for them in the baseline specification would risk partialling out the mechanism the model is designed to capture. For the output gap and negative unemployment gap, we control for productivity growth and inflation. For payroll and industrial production growth, we additionally control for the current output gap.

Online Appendix [BO.15.1](#) reports robustness checks using alternative sample windows, COVID exclusion, moving-block bootstrap inference, and a horse-race. Tables [2.3–2.6](#) report the baseline results.

The estimates provide suggestive dynamic evidence that the model's state variable contains forward-looking information. Across all four outcomes, the high-capital slope β_h is positive and increasing with the forecast horizon. It ranges from 0.074 to 0.306 for the output gap, from 0.094 to 0.319 for the negative unemployment gap, from 0.071 to 0.561 for cumulative payroll growth, and from 0.449 to 2.177 for cumulative industrial production growth. These estimates are not statistically different from zero for the output gap, negative unemployment gap, or payroll growth, but they are more precisely estimated for industrial production throughout. The low-capital interaction κ_h is positive in all specifications, indicating steeper predictive slopes below the crisis threshold. The implied low-capital slope $\beta_h + \kappa_h$ is positive and statistically significant for all outcomes: it ranges from 1.610 to 2.373 for the output gap, from 1.668 to 2.567 for the negative unemployment gap, from 1.778 to 3.211 for cumulative payroll growth,

Table 2.3: Test 1B – predictive regressions, output gap

	$h = 1$	$h = 2$	$h = 4$
α_h	-0.005 (0.004)	-0.005 (0.005)	-0.007 (0.006)
ρ_h	0.601*** (0.176)	0.429** (0.203)	0.139 (0.224)
β_h (high-capital slope)	0.074 (0.132)	0.171 (0.249)	0.306 (0.369)
κ_h (low-capital interaction)	1.536** (0.626)	1.632*** (0.479)	2.067*** (0.704)
$\beta_h + \kappa_h$ (low-capital slope)	1.610** (0.674)	1.803*** (0.482)	2.373*** (0.693)
Controls	Yes	Yes	Yes
Adj. R^2	0.744	0.564	0.380
N	119	118	116

Notes: Regressors include the current dependent variable, the centered model-implied output gap, its low-capital interaction, and the listed controls. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. The low-capital slope is $\beta_h + \kappa_h$. The p -value on κ_h tests equality of the low- and high-capital slopes. Controls: productivity growth and inflation. ***, **, * denote significance at 1%, 5%, 10%.

Table 2.4: Test 1B – predictive regressions, negative unemployment gap

	$h = 1$	$h = 2$	$h = 4$
α_h	-0.006 (0.004)	-0.007 (0.005)	-0.008 (0.006)
ρ_h	0.592*** (0.185)	0.428** (0.207)	0.149 (0.232)
β_h (high-capital slope)	0.094 (0.086)	0.156 (0.145)	0.319 (0.247)
κ_h (low-capital interaction)	1.573** (0.731)	1.788*** (0.344)	2.248*** (0.529)
$\beta_h + \kappa_h$ (low-capital slope)	1.668** (0.778)	1.944*** (0.403)	2.567*** (0.502)
Controls	Yes	Yes	Yes
Adj. R^2	0.788	0.644	0.502
N	119	118	116

Notes: Regressors include the current dependent variable, the centered model-implied output gap, its low-capital interaction, and the listed controls. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. The low-capital slope is $\beta_h + \kappa_h$. The p -value on κ_h tests equality of the low- and high-capital slopes. Controls: productivity growth and inflation. ***, **, * denote significance at 1%, 5%, 10%.

and between 3.080 and 4.399 for cumulative industrial production growth. Thus, the dynamic relation between intermediary capital and subsequent real activity is concentrated primarily in low-capital states. Overall, Test 1B supports the view that intermediary capital contains dynamic information about subsequent real activity in parsimonious predictive specifications, with stronger evidence in low-capital states. We interpret this evidence as suggestive dynamic validation. Test 1A remains the cleaner contemporaneous test of the model's real-activity mapping.

Table 2.5: Test 1B – predictive regressions, cumulative payroll growth

	$h = 1$	$h = 2$	$h = 4$
α_h	-0.003 (0.006)	-0.003 (0.009)	-0.005 (0.014)
ρ_h	-0.026 (0.109)	0.015 (0.116)	0.274 (0.211)
β_h (high-capital slope)	0.071 (0.152)	0.219 (0.312)	0.561 (0.595)
κ_h (low-capital interaction)	1.707* (0.900)	1.969** (0.780)	2.650** (1.236)
$\beta_h + \kappa_h$ (low-capital slope)	1.778* (0.968)	2.188** (0.877)	3.211** (1.256)
Controls	Yes	Yes	Yes
Adj. R^2	0.139	0.100	0.130
N	119	118	116

Notes: Regressors include the current dependent variable, the centered model-implied output gap, its low-capital interaction, and the listed controls. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. The low-capital slope is $\beta_h + \kappa_h$. The p -value on κ_h tests equality of the low- and high-capital slopes. Controls: output gap, productivity growth and inflation. ***, **, * denote significance at 1%, 5%, 10%.

Table 2.6: Test 1B – predictive regressions, cumulative industrial production growth

	$h = 1$	$h = 2$	$h = 4$
α_h	-0.012* (0.007)	-0.016 (0.010)	-0.023* (0.014)
ρ_h	0.223** (0.090)	0.408** (0.189)	0.619*** (0.236)
β_h (high-capital slope)	0.449* (0.252)	0.979** (0.484)	2.177*** (0.826)
κ_h (low-capital interaction)	2.631** (1.061)	2.762** (1.260)	2.222 (2.106)
$\beta_h + \kappa_h$ (low-capital slope)	3.080*** (1.142)	3.742*** (1.301)	4.399** (2.163)
Controls	Yes	Yes	Yes
Adj. R^2	0.397	0.361	0.379
N	119	118	116

Notes: Regressors include the current dependent variable, the centered model-implied output gap, its low-capital interaction, and the listed controls. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. The low-capital slope is $\beta_h + \kappa_h$. The p -value on κ_h tests equality of the low- and high-capital slopes. Controls: output gap, productivity growth and inflation. ***, **, * denote significance at 1%, 5%, 10%.

2.7.4 Test 2: Dealer Intermediation and Central Bank Substitution

Test 2A The second model implication we test is that private dealer intermediation should weaken when intermediary capital is scarce. In the model, private intermediaries must absorb the risky-asset supply, so the required portfolio share rises as intermediary capital falls (see (2.9)); however, once the leverage constraint binds, private balance sheets can no longer intermediate the required quantity of risk. As in Test 1B, the direct theoretical implication is a state-dependent relation rather than a fully identified causal impulse response. We therefore ask whether current intermediary capital, viewed as a persistent state variable, predicts future dealer intermediation

differently above and below the crisis threshold:

$$y_{t+h}^{dealer} = \alpha_h + \rho_h y_t^{dealer} + \beta_h m_t + \kappa_h m_t \mathbf{1}\{\hat{\omega}_t < \hat{\omega}_c\} + \Gamma'_h X_t + \varepsilon_{t+h},$$

where $m_t = \hat{\omega}_t - \hat{\omega}_c$.

We implement this test for horizons $h \in \{1, 2, 4\}$ quarters. The dependent variables are dealer financing ratios: Treasury dealer financing to clients scaled by Treasury securities held by the public excluding SOMA, and agency MBS dealer financing to clients scaled by agency MBS held by the public excluding SOMA. Agency MBS are closer to the model's literal risky-security interpretation. Treasuries, however, provide an important complementary test of dealer balance-sheet absorption: Treasury intermediation is not risky in the credit-risk sense, but it requires dealers to warehouse duration, liquidity, funding, and basis risk in the largest market directly targeted by Federal Reserve asset purchases. We therefore interpret the Treasury specification as a test of dealer risk-bearing capacity, while the MBS specification is the closer asset-class analogue to the model's risky-security supply.

Including y_t^{dealer} controls for the strong persistence in dealer intermediation. The coefficient β_h is the high-capital slope, $\beta_h + \kappa_h$ is the implied low-capital slope, and κ_h tests whether the low-capital and high-capital slopes differ. Since m_t is lower when intermediary capital is lower, the model predicts a positive low-capital slope, $\beta_h + \kappa_h > 0$: lower intermediary capital should map into lower future dealer intermediation. If the low-capital state is quantitatively important, the model also suggests a positive interaction term, $\kappa_h > 0$, although this state-dependence is a sharper requirement than the sign of the low-capital slope itself. We keep the control vector parsimonious and include the output gap, productivity growth, and inflation.

Online Appendix [BO.15.2](#) reports sample-exclusion, wider-control, and moving-block bootstrap robustness checks. Tables [2.7](#) and [2.8](#) report the baseline results.

The results are mixed but informative. Treasury dealer financing provides the clearest evidence that intermediary capital has predictive content for private dealer intermediation. The implied low-capital slope, $\beta_h + \kappa_h$, is positive and statistically significant at all reported horizons, with estimates of 1.511, 2.896, and 3.484 at horizons of one, two, and four quarters. Thus, conditional on current dealer financing and the macro controls, lower intermediary capital predicts weaker subsequent Treasury dealer financing. However, the low-capital interaction κ_h is imprecisely estimated, so the evidence is better interpreted as support for the capital-

Table 2.7: Test 2A – predictive regressions, treasury dealer financing

	$h = 1$	$h = 2$	$h = 4$
α_h	0.013 (0.013)	0.024 (0.015)	0.007 (0.018)
ρ_h	0.962*** (0.035)	0.895*** (0.055)	0.787*** (0.095)
β_h (high-capital slope)	1.093 (0.877)	2.106 (1.566)	6.094** (3.101)
κ_h (low-capital interaction)	0.418 (1.172)	0.790 (2.091)	-2.610 (3.197)
$\beta_h + \kappa_h$ (low-capital slope)	1.511** (0.674)	2.896** (1.334)	3.484*** (1.182)
Controls	Yes	Yes	Yes
Adj. R^2	0.968	0.939	0.890
N	78	76	75

Notes: Regressors include current dealer financing, centered intermediary capital, its low-capital interaction, and the listed controls. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. The low-capital slope is $\beta_h + \kappa_h$. The p -value on κ_h tests equality of the low- and high-capital slopes. Controls are output gap, productivity growth, and inflation. ***, **, * denote significance at 1%, 5%, 10%.

Table 2.8: Test 2A – predictive regressions, agency MBS dealer financing

	$h = 1$	$h = 2$	$h = 4$
α_h	-0.002 (0.001)	-0.005** (0.002)	-0.009*** (0.003)
ρ_h	0.972*** (0.022)	0.953*** (0.036)	0.862*** (0.064)
β_h (high-capital slope)	0.184*** (0.069)	0.300*** (0.116)	0.892*** (0.214)
κ_h (low-capital interaction)	-0.167 (0.160)	-0.129 (0.261)	-0.894*** (0.314)
$\beta_h + \kappa_h$ (low-capital slope)	0.016 (0.116)	0.171 (0.196)	-0.002 (0.187)
Controls	Yes	Yes	Yes
Adj. R^2	0.957	0.922	0.858
N	78	76	75

Notes: Regressors include current dealer financing, centered intermediary capital, its low-capital interaction, and the listed controls. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. The low-capital slope is $\beta_h + \kappa_h$. The p -value on κ_h tests equality of the low- and high-capital slopes. Controls are output gap, productivity growth, and inflation. ***, **, * denote significance at 1%, 5%, 10%.

intermediation link than as definitive evidence of a steeper low-capital slope. The agency MBS financing results are weaker. Although the high-capital slope is positive and statistically significant at every horizon, the implied low-capital slope is small and statistically insignificant throughout. At the four-quarter horizon, the low-capital interaction is negative and statistically significant, which cuts against the model's low-capital amplification prediction. Overall, Test 2A provides supportive evidence for the private balance-sheet absorption channel in Treasury dealer financing, but not in agency MBS dealer financing.

Test 2B The next model implication we test is a substitution result. In the model, once intermediary capital falls below the crisis threshold, private balance sheets can no longer absorb the full risky-asset supply, so public risk-bearing must rise to fill the gap. We therefore regress changes in central-bank risk-bearing on lagged changes in dealer intermediation and on the lagged capital shortfall:

$$\Delta RB_t^{CB} = \alpha + \rho \Delta RB_{t-1}^{CB} + \phi s_{t-1} + \beta \Delta y_{t-1}^{dealer} + \Gamma' X_{t-1} + \varepsilon_t,$$

where the intermediary-capital shortfall is:

$$s_t = (\hat{\omega}_c - \hat{\omega}_t) \mathbf{1}\{\hat{\omega}_t < \hat{\omega}_c\}.$$

The dependent variable is the quarterly change in central-bank risk-bearing in Treasuries or agency MBS, normalized by GDP. Baseline Test 2B uses dealer gross positions because they are closer to the balance-sheet absorption margin in the model than dealer financing. These gross-position data are available only in the post-2012 sample, so the baseline evidence should be interpreted as suggestive post-crisis central-bank/dealer substitution evidence, not as evidence about the 2008 GFC itself.

The model predicts $\beta < 0$: when dealer intermediation contracts, the central bank should subsequently increase its own risk-bearing in the same market. The coefficient ϕ should be positive if public risk-bearing rises more when the intermediary-capital shortfall is larger. As in Test 2A, the Treasury and MBS specifications play complementary roles. Agency MBS are closer to the model's risky-security interpretation, while Treasuries provide a clean empirical test of balance-sheet substitution in a market where dealers intermediate duration, liquidity, funding, and basis risk and where Federal Reserve purchases are quantitatively central. We keep the control vector parsimonious and include the output gap, productivity growth, and inflation.

Online Appendix [BO.15.2](#) reports COVID-exclusion, wider-control, bootstrap, and financing-based longer-sample robustness checks. Tables [2.9](#) and [2.10](#) report the baseline results.

The substitution evidence is strongest in Treasuries. The coefficient on lagged dealer gross-position growth is negative and statistically significant, with an estimate of -2.731 . Thus, when Treasury dealer intermediation contracts, the Fed subsequently increases its Treasury risk-bearing. The shortfall coefficient is also positive and statistically significant, with an estimate

Table 2.9: Test 2B – central bank substitution, treasury dealer gross positions

Treasury gross positions	
α	0.063*** (0.024)
ρ	1.050*** (0.255)
ϕ	6.497** (2.645)
β	-2.731** (1.301)
Controls	Yes
Adj. R^2	0.601
N	48

Notes: The dependent variable is the quarterly change in Treasury Fed risk-bearing to GDP. Dealer intermediation, capital shortfall, and controls are lagged one quarter. Controls are output gap, productivity growth, and inflation. The gross-position specification uses the shorter sample available from 2012 onward. The baseline evidence should be read as post-crisis substitution evidence because gross-position data are not available for the GFC period. HAC standard errors with four lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Table 2.10: Test 2B – central bank substitution, agency MBS dealer gross positions

Agency MBS gross positions	
α	0.022* (0.012)
ρ	0.730** (0.290)
ϕ	3.886** (1.979)
β	-0.137 (0.563)
Controls	Yes
Adj. R^2	0.357
N	48

Notes: The dependent variable is the quarterly change in agency MBS Fed risk-bearing to GDP. Dealer intermediation, capital shortfall, and controls are lagged one quarter. Controls are output gap, productivity growth, and inflation. The gross-position specification uses the shorter sample available from 2012 onward. The baseline evidence should be read as post-crisis substitution evidence because gross-position data are not available for the GFC period. HAC standard errors with four lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

of 6.497, indicating that Treasury interventions are larger when intermediary capital is further below the crisis threshold. The agency MBS evidence is more limited. The coefficient on lagged dealer gross-position growth is negative, as predicted, but small and statistically insignificant. By contrast, the shortfall coefficient remains positive and statistically significant, with an estimate of 3.886, suggesting that MBS risk-bearing by the Fed also rises when intermediary capital is scarce, even though the direct private-to-public substitution margin is imprecisely measured. Overall, Test 2B supports the central-bank substitution mechanism most clearly in the Treasury market. The MBS results are not contradictory, but they mainly support the state-contingent intervention margin rather than the direct dealer-substitution coefficient.

2.7.5 Test 3: Pre-QE Intermediation Gap as a Sufficient Statistic

Test 3A The third model implication we test is that the intermediation gap is a sufficient statistic for real activity (see (2.35)). This mapping holds in the unconstrained pre-QE regime. If the model is informative about real activity, the model-implied intermediation gap should line up with the real-activity gap observed in the data. Therefore, we regress empirical proxies for real activity on the model-implied (pre-QE) intermediation gap. We estimate both a linear specification,

$$x_t^{data} = \alpha + \beta_1 \widetilde{b}_t^{pre} + \varepsilon_t,$$

and a quadratic specification,

$$x_t^{data} = \alpha + \beta_1 \widetilde{b}_t^{pre} + \beta_2 \left(\widetilde{b}_t^{pre} \right)^2 + \varepsilon_t,$$

where

$$\widetilde{b}_t^{pre} = \hat{b}_t^{pre} - \bar{\hat{b}}^{pre}, \quad \hat{b}_t^{pre} = b_u(\hat{\omega}_t), \quad b_u''(\omega) = \frac{2\delta}{[\sigma_\omega^u(\omega)]^2} b_u(\omega), \quad b_u(0) = 1, \quad b_u(1/\epsilon) = 0.$$

We implement this test in the pre-QE sample from 1970Q1 to 2007Q4 for two measures of real economic activity, namely the output gap and the sign-reversed unemployment gap. All parameters are fixed at their calibrated values. The linear specification provides the direct level-fit analogue of Test 1A. The quadratic specification is motivated by the theory's sufficient-statistic relation: the model implies that real activity is a nonlinear function of the exact gap, $s_t^{data} = G(\hat{b}_t^{pre})$, rather than a linear function of $\hat{b}_t^{pre}$. The quadratic term therefore gives a parsimonious second-order approximation to this nonlinear mapping over the observed pre-QE state space. Centering the exact gap makes β_1 the marginal relation at the average pre-QE gap and reduces collinearity between the linear and quadratic terms.

Since a larger exact gap corresponds to lower intermediary capital and weaker real activity, the model predicts a negative marginal relation between real activity and the exact gap. In the linear specification, this prediction is $\beta_1 < 0$. In the quadratic specification, the relevant object is the implied marginal effect, $\frac{\partial s_t^{data}}{\partial \hat{b}_t^{pre}} = \beta_1 + 2\beta_2 \widetilde{b}_t^{pre}$. We therefore report the marginal effect evaluated in 2007Q4, near the end of the pre-QE period, together with the joint test of whether the intermediation-gap terms are zero. We report robustness checks using alternative pre-QE

sample starts in Online Appendix [BO.15.3](#). Table 2.11 shows the baseline results.

Table 2.11: Test 3A – theory-implied level fit, exact intermediation gap

	Output gap		Negative unemployment gap	
	Linear	Quadratic	Linear	Quadratic
α	-0.004 (0.003)	-0.006 (0.004)	-0.004** (0.002)	-0.007*** (0.003)
β_1	-0.082** (0.035)	-0.112** (0.047)	-0.045** (0.020)	-0.082*** (0.027)
β_2	—	0.307 (0.328)	—	0.378* (0.196)
$\beta_1 + 2\beta_2\widehat{b}_{2007Q4}^{pre}$	—	-0.134** (0.065)	—	-0.110*** (0.038)
p -value of $H_0 : \beta_1 = \beta_2 = 0$	—	0.044	—	0.011
Adj. R^2	0.130	0.137	0.113	0.161
N	152	152	152	152

Notes: The sample ends in 2007Q4. The table reports contemporaneous level-fit regressions. Both specifications use $\widehat{b}_t^{pre} = b_t^{pre} - \overline{b_t^{pre}}$, where $b_t^{pre} = b_u(\hat{\omega}_t)$ and the mean is computed within the corresponding pre-QE estimation sample. The quadratic columns also include $(\widehat{b}_t^{pre})^2$. The marginal-effect row evaluates $\beta_1 + 2\beta_2\widehat{b}_t^{pre}$ in 2007Q4 for the quadratic columns. The reported p -value tests $H_0 : \beta_1 = \beta_2 = 0$ for the quadratic columns. HAC standard errors with four lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

The results support the model's pre-QE mapping from the intermediation gap to real activity. In the linear specifications, the coefficient on the centered exact gap is negative and statistically significant for both the output gap and the sign-reversed unemployment gap, with estimates of -0.082 and -0.045 , respectively. The quadratic specifications preserve the same negative relation. Evaluated in 2007Q4, the implied marginal effect is -0.134 for the output gap and -0.110 for the sign-reversed unemployment gap, and the intermediation-gap terms are jointly significant in both quadratic specifications. The quadratic term itself is positive and statistically significant only in the unemployment specification, so we do not interpret the estimates as sharp evidence on the precise curvature of G . The main evidence is instead that the exact pre-QE pricing object has the predicted negative relation with observed slack, both in the linear projection and in the parsimonious nonlinear approximation.

Test 3B We next ask whether the exact pre-QE intermediation-gap proxy has incremental predictive content beyond raw intermediary capital. The purpose of this test is narrower than a general forecasting horse race: the model's sufficient-statistic result says that, in the unconstrained pre-QE regime, the exact nonlinear transformation $\hat{b}_t^{pre} = b_u(\hat{\omega}_t)$ is the structural pricing object linked to real activity. Once QE is active, however, the actual intermediation gap is no longer present. The counterfactual object $\hat{b}_t^{pre}$ may still forecast output because it remains

a nonlinear transformation of the persistent state variable $\hat{\omega}_t$, but its separate structural content should be weaker. We therefore estimate a fully saturated pooled interaction specification:

$$x_{t+h}^{data} = \alpha_h + \alpha_h^\Delta D_t^{QE} + \left(\rho_h + \rho_h^\Delta D_t^{QE} \right) x_t^{data} + \left(\phi_h + \phi_h^\Delta D_t^{QE} \right) \hat{\omega}_t + \left(\beta_h + \beta_h^\Delta D_t^{QE} \right) \hat{b}_t^{pre} + \varepsilon_{t+h}.$$

We implement this test for horizons $h \in \{1, 2, 4\}$ for the output gap. We use the whole 1970Q1–2024Q4 sample, omitting the 2008Q1–2009Q1 transition quarters and dropping observations whose forecast target crosses regimes, so that pre-QE observations forecast only pre-QE outcomes and QE-active observations forecast only QE-active outcomes.

Including x_t^{data} controls for the strong persistence in the output gap. The variable D_t^{QE} is an indicator equal to one in QE-active quarters. The pre-QE exact-gap coefficient is $\beta_h^{pre} = \beta_h$, the QE-active coefficient is $\beta_h^{QE} = \beta_h + \beta_h^\Delta$, and β_h^Δ is the regime-difference coefficient. The coefficient of primary interest is β_h^{pre} . The model predicts $\beta_h^{pre} < 0$: a larger exact pre-QE gap corresponds to weaker future real activity. The regime-difference test is more demanding. The reported Wald test evaluates whether the predictive slope differs in the two regimes (i.e. we test for the equality of pre- and post-QE slopes, $H_0 : \beta_h^\Delta = 0$). A weaker or imprecisely estimated β_h^{QE} is consistent with the disappearance of the actual intermediation gap under QE, but the model does not require $\hat{b}_t^{pre}$ to become uncorrelated with future output.

The results support the pre-QE sufficient-statistic implication. The exact-gap coefficient is negative and statistically significant at all reported horizons, with estimates of -0.037 , -0.062 , and -0.085 for $h = 1, 2, 4$, respectively. By contrast, the raw intermediary-capital coefficient is small and statistically insignificant in the pre-QE regime once the exact nonlinear transformation is included. This indicates that the exact pre-QE pricing object contains incremental information about future output gaps beyond a linear control for raw intermediary capital. The QE-active exact-gap coefficients are also negative, but imprecisely estimated. Moreover, the Wald tests do not reject equality of the pre-QE and QE-active exact-gap coefficients, with p -values of 0.546, 0.716, and 0.815. We therefore interpret Test 3B as evidence for the pre-QE structural mapping, rather than as a sharp formal rejection of incremental content under QE. Online Appendix [BO.15.3](#) reports sample, control, bootstrap, and benchmark-predictor robustness checks.

Table 2.12: Test 3B – pooled regression, exact intermediation gap

	$h = 1$	$h = 2$	$h = 4$
ρ_h	0.880*** (0.039)	0.728*** (0.071)	0.380*** (0.112)
ρ_h^{QE}	0.519* (0.280)	0.494** (0.232)	0.333 (0.247)
ϕ_h^{pre}	-0.038 (0.041)	-0.040 (0.080)	-0.000 (0.141)
ϕ_h^{QE}	0.118 (0.289)	-0.198 (0.485)	-0.042 (0.375)
β_h^{pre} (exact gap, pre-QE)	-0.037*** (0.013)	-0.062** (0.025)	-0.085** (0.037)
β_h^{QE} (exact gap, QE-active)	-0.101 (0.104)	-0.096 (0.093)	-0.070 (0.055)
$\beta_h^{QE} - \beta_h^{pre}$ (regime difference)	-0.063 (0.105)	-0.035 (0.096)	0.015 (0.066)
Wald p-value: $\beta_h^{QE} = \beta_h^{pre}$	0.546	0.716	0.815
Adj. R^2	0.791	0.608	0.338
N	213	211	207
Pre-QE obs.	151	150	148
QE obs.	62	61	59

Notes: Fully saturated pooled interaction regression comparing the incremental content of the exact pre-QE gap proxy before QE and during QE-active quarters. The sample includes pre-QE observations through 2007Q4 and QE-active observations with $QeOn = 1$. Transition quarters 2008Q1–2009Q1 are excluded, and observations whose forecast target crosses regimes are omitted. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

2.7.6 Test 4: Fiscal Recapitalization and Recovery

Test 4A The fourth model implication we test is that fiscal recapitalization, rather than QE itself, should rebuild intermediary capital after a systemic banking crisis. In the model, QE restores equilibrium existence when intermediary balance sheets are too weak to absorb the risky-asset supply, but it does not by itself shift the law of motion of intermediary net worth toward the good steady state. By contrast, a fiscal capital injection directly raises intermediary net worth and should therefore predict subsequent recovery in intermediary capital. Thus, we regress future changes in intermediary capital on current recapitalization flows and QE:

$$y_{i,t+h} - y_{i,t} = \alpha_i + \tau_t + \beta_h RecapFlow_{i,t} + \gamma_h QE_{i,t} + \psi_h x_{i,t-1} + \varepsilon_{i,t+h}, \quad (2.43)$$

for horizons $h \in \{1, 2, 4\}$ years. We implement this test in an annual panel of 10 countries over 2003–2020, using a balanced panel for which the equity-to-assets ratio is observed for all countries. The dependent variable is the equity-to-assets ratio for country i , and the regressors are the annual public recapitalization flow scaled by GDP, and central-bank assets scaled by GDP. We use the annual recapitalization flow rather than the cumulative stock because the

year-end capital ratio already embeds much of the contemporaneous accounting effect of capital injections, so the flow specification provides a cleaner test of subsequent capital rebuilding. The coefficient of primary interest is β_h . The model predicts $\beta_h > 0$: larger recapitalization flows should be associated with faster subsequent recovery in intermediary capital. By contrast, γ_h should be small if QE stabilizes markets without directly recapitalizing the intermediary sector. We keep the control vector parsimonious and include only the lagged output gap, which captures differences in initial crisis severity across countries while avoiding conditioning on a contemporaneous macro outcome that may itself already reflect the effects of recapitalization and QE in year t . We also include country and time fixed effects to absorb, respectively, persistent cross-country differences in financial structure and crisis exposure, and global shocks common to all countries in a given year. Table 2.13 reports the results.

Table 2.13: Test 4A – predictive regressions, intermediary capital recovery

	$h = 1$	$h = 2$	$h = 4$
β_h	0.169*** (0.055)	0.274*** (0.068)	0.201* (0.113)
γ_h	-0.033 (0.034)	-0.070 (0.052)	-0.057 (0.076)
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Adj. R^2	0.130	0.123	0.090
N	160	150	130

Dependent variable is $EqAssets_{i,t+h} - EqAssets_{i,t}$ (change in the equity-to-assets ratio, decimal). Annual panel of 10 countries over 2003–2020 with country and time fixed effects. Main regressor is the annual recapitalization flow to GDP rather than the cumulative recap stock. Control: lagged output gap. Horizons are measured in years. Country-clustered standard errors are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

The results provide clear support for the model’s recapitalization mechanism. The coefficient on recapitalization flows is positive and statistically significant at every reported horizon, with estimates of 0.169, 0.274, and 0.201 at one, two, and four years, respectively. Economically, these estimates imply that countries undertaking larger bank recapitalizations experience systematically faster subsequent rebuilding of intermediary capital. The effect is largest at the two-year horizon and remains positive and statistically significant at four years, which is consistent with recapitalization having persistent effects on financial-sector balance sheets rather than only a short-lived accounting effect. By contrast, the coefficient on QE is negative throughout and never statistically distinguishable from zero. Taken together, these results suggest that fiscal recapitalization, not QE alone, is the policy instrument that directly restores intermediary capital in the data.

Test 4B Fiscal recapitalization should not only rebuild intermediary capital, but also translate into stronger subsequent real recovery. In the model, recapitalization raises intermediary net worth, moves the economy away from the bad steady state, and thereby improves the subsequent path of output and employment. Therefore, we regress future real outcomes on current recapitalization and QE:

$$y_{i,t+h} - y_{i,t} = \alpha_i + \tau_t + \beta_h \text{Recap}_{i,t} + \gamma_h \text{QE}_{i,t} + \psi_h x_{i,t-1} + \varepsilon_{i,t+h}, \quad (2.44)$$

for horizons $h \in \{1, 2, 4\}$ years. We implement this test in the same annual panel of 10 countries over 2003–2020, using three dependent variables: the cumulative change in the output gap, the sign-reversed cumulative change in the unemployment rate, and cumulative real GDP growth. The main regressor is the cumulative recapitalization stock scaled by GDP, rather than the recapitalization flow, because the model’s prediction for real recovery is a state-variable prediction: what matters for subsequent macroeconomic dynamics is the amount of intermediary capital that has been restored, not only the contemporaneous injection flow. The coefficient of primary interest is β_h . For all three outcomes, the model predicts $\beta_h > 0$: countries with larger recapitalizations should subsequently experience faster output gap closure, stronger labor-market recovery, and higher GDP growth. By contrast, if QE mainly stabilizes markets without shifting long-run real dynamics, γ_h should be smaller and less robust. As in Test 4A, we keep the control vector parsimonious, with only the lagged output gap as a control, and include country and time fixed effects. Tables 2.14–2.16 report the results.

Table 2.14: Test 4B – predictive regressions, output gap

	$h = 1$	$h = 2$	$h = 4$
β_h (Recap)	0.062*** (0.023)	0.127*** (0.040)	0.128*** (0.036)
γ_h (QE)	-0.053* (0.030)	-0.095* (0.049)	-0.094 (0.093)
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Adj. R^2	0.292	0.451	0.495
N	160	150	130

Dependent variable is $x_{i,t+h} - x_{i,t}$ (cumulative change in output gap, decimal). Annual panel of 10 countries over 2003–2020 with country and time fixed effects. Main regressor is the cumulative recapitalization stock to GDP. Control: lagged output gap. Horizons are measured in years. Country-clustered standard errors are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Table 2.15: Test 4B – predictive regressions, negative unemployment gap

	$h = 1$	$h = 2$	$h = 4$
β_h (Recap)	0.105*** (0.015)	0.213*** (0.029)	0.339*** (0.056)
γ_h (QE)	0.082*** (0.016)	0.118*** (0.035)	0.000 (0.096)
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Adj. R^2	0.295	0.453	0.598
N	144	135	117

Dependent variable is $-(u_{i,t+h} - u_{i,t})$ (sign-reversed cumulative change in unemployment rate, decimal). Annual panel over 2003–2020 with country and time fixed effects. The unemployment specification can use fewer than 10 countries because the annual series inherits missing values from the quarterly OECD source. Main regressor is the cumulative recapitalization stock to GDP. Control: lagged output gap. Horizons are measured in years. Country-clustered standard errors are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Table 2.16: Test 4B – predictive regressions, cumulative GDP growth

	$h = 1$	$h = 2$	$h = 4$
β_h (Recap)	0.345*** (0.052)	0.758*** (0.093)	1.659*** (0.141)
γ_h (QE)	-0.120 (0.090)	-0.244 (0.169)	-0.224 (0.150)
Country FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Controls	Yes	Yes	Yes
Adj. R^2	0.144	0.293	0.655
N	160	150	130

Dependent variable is $\sum_{j=1}^h \Delta y_{i,t+j}$ (cumulative annual real GDP growth, decimal). Annual panel of 10 countries over 2003–2020 with country and time fixed effects. Main regressor is the cumulative recapitalization stock to GDP. Control: lagged output gap. Horizons are measured in years. Country-clustered standard errors are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

The results are supportive overall. For the output gap, the coefficient on recapitalization is positive and statistically significant at every reported horizon, with estimates of 0.062, 0.127, and 0.128 at one, two, and four years, respectively. Thus, countries with larger recapitalizations subsequently close their output gaps faster. For cumulative GDP growth, the evidence is even stronger: the recapitalization coefficient is positive and statistically significant throughout, rising from 0.345 at one year to 0.758 at two years and 1.659 at four years. This pattern is consistent with recapitalization having persistent medium-run real effects rather than only short-run financial effects. The unemployment results point in the same direction. The coefficient on recapitalization is positive and statistically significant at every horizon, with estimates increasing from 0.105 to 0.213 and 0.339, indicating that recapitalized banking systems are followed by stronger labor-market recoveries. The QE coefficients do not display the same robust recovery

pattern: they are negative at short horizons in the output-gap specification, statistically insignificant in the GDP-growth specification, and positive only at the one- and two-year horizons in the unemployment specification. Taken together, these three specifications suggest that fiscal recapitalization is systematically associated with stronger subsequent macroeconomic recovery, whereas QE by itself does not deliver the same robust pattern across outcomes.

2.7.7 Test 5: QE Trap

The model’s most distinctive dynamic prediction is permanent QE: once ω falls below $\underline{\omega}$, the economy converges to the bad steady state (Theorem 2.4); even above $\underline{\omega}$, the economy can lie in the QE Trap—the region where convergence to the good steady state is mathematically certain but deterministic recovery takes centuries—so intermediary capital remains durably below its pre-crisis level on any policy-relevant horizon (Appendix BO.14.1). To assess whether the data display a corresponding persistent post-2007 shift in intermediary capital, we estimate multiple structural breaks in the mean of the HKM intermediary capital ratio using the Bai and Perron (1998) procedure on the full 1970Q1–2024Q4 sample. To ensure that the selected number of breaks is not mechanically bounded, the break-selection exercise uses a trimming parameter of 0.10 and allows up to nine breaks. The Bayesian information criterion identifies seven breaks, many of them pre-2007. Yet, for the model, the key question is not the exact number of breaks in the sample, but whether the data show a persistent downward shift in intermediary capital after 2007. Table 2.17 reports the results. Appendix B.12.2 reports the full break-count disclosure and a modern-sample robustness exercise.

Table 2.17: Test 5 – regime summary, HKM intermediary capital ratio

Regime	Start	End	$\bar{\omega}$	Δ vs. pre-2007
Pre-2007 benchmark	2002-06	2007-09	0.0798	0.0000
Post-crisis low-capital regime	2007-12	2013-06	0.0445	-0.0353
Later post-crisis regime	2013-09	2024-12	0.0632	-0.0166

Notes: This table reports the immediate pre-crisis benchmark regime together with the post-2007 regimes from the BIC-selected Bai–Perron mean-shift specification. Full break-count disclosure is reported separately in Table B.3. The calibrated threshold is $\omega_c = 0.061$ and should be interpreted descriptively rather than as an endogenous threshold estimate.

The structural-break evidence is consistent with a durable post-2007 downward shift in intermediary capital. The immediate pre-crisis benchmark regime, spanning 2002Q2–2007Q3, has mean $\bar{\omega} = 0.0798$. The first post-crisis regime, 2007Q4–2013Q2, has mean $\bar{\omega} = 0.0445$, a decline of 0.0353, and lies below the calibrated threshold used in the paper. The later post-

crisis regime, 2013Q3–2024Q4, recovers only partially, with mean $\bar{\omega} = 0.0632$, which is only slightly above the crisis threshold and still 0.0166 below the modern pre-crisis benchmark. Thus, the evidence is not that intermediary capital remained permanently at crisis levels, but that the post-2007 period is characterized by a lower-capital regime than the one that prevailed immediately before the crisis. We interpret this result as descriptive support for the QE-trap narrative, not as an endogenous estimate of the model threshold or as a causal decomposition of why capitalization remained lower.

2.8 Conclusion

A financial crisis that depletes intermediary capital can destroy competitive equilibrium—not merely distort prices. In the moderately distressed region, three ingredients interact: limited participation, leverage constraints, and depleted capital. In the severely distressed region, the failure is unconditional: Prompt Corrective Action forces the financier below full equity investment, and even full household participation cannot compensate. The crisis threshold is preference-independent—identical under CRRA with any elasticity of intertemporal substitution and under Epstein–Zin preferences (Online Appendices [BO.10](#) and [BO.12](#)). Large-scale asset purchases are a necessary condition for market clearing: the central bank acts as the residual risk-bearer whose intervention restores the existence of a competitive outcome. But QE, once started, may not be reversible: if the crisis pushes intermediary capital below the lower natural boundary $\underline{\omega}$, the economy enters *Permanent QE* where asset purchases persist without exit even in theory; if capital lands just above $\underline{\omega}$, the economy enters the *QE Trap* where recovery is certain but takes centuries (Remark [2.2](#)). A one-time fiscal recapitalisation of the financial sector is the only escape from both regimes, with an employment multiplier that is convex in crisis severity—rationalising the urgency of early intervention documented by [Homar and van Wijnbergen \(2017\)](#).

A leverage paradox links these results: the same leverage constraint that destroys equilibrium simultaneously eliminates the intermediation gap between market prices and fundamental values. The intermediation gap satisfies a production-economy ODE (Proposition [2.3](#)) whose exact solution is substantially smaller than the exchange-economy closed form of [Hugonnier \(2012\)](#): under the calibration in Table [2.1](#), the closed form overstates the exact gap by about 2.5× at $\underline{\omega}$, 5.3× at $\omega = 0.04$, and more than 30× at ω_c . No economy in this class can simultaneously

achieve gap-free prices and unconditional market clearing—a fundamental impossibility that the policy package of QE plus fiscal recapitalisation resolves.

Several extensions are developed in the Online Appendices: endogenous participation via ambiguity aversion (Online Appendix [BO.11](#)), CRRA and Epstein–Zin robustness with full characterisations of how preferences shape both Permanent QE and the QE Trap (Online Appendices [BO.10](#) and [BO.12](#)), leverage versus equity constraints (Online Appendix [BO.13](#)), and cross-country tests of the fiscal escape mechanism (Section [2.7.6](#)).

Chapter 3

The Econometrics of Portfolio Sorts

Nicolo Ceneda, Valentina Corradi and Walter Distaso

3.1 Introduction

Portfolio sorting according to some stock characteristics is a standard practice in empirical asset pricing. The objective is generally to uncover some systematic cross-sectional relationship between expected returns and stock characteristics. A special role is played by sorting according to betas (Fama and MacBeth, 1973). Other well known examples of sorting variables are size and book-to-market value (Fama and French, 1993), momentum (Jegadeesh and Titman, 1993), idiosyncratic volatility (Ang et al., 2006), coskewness (Harvey and Siddique, 2000), and realized skewness and kurtosis (Amaya et al., 2015).

The null hypothesis of interest is that the expected return over low and high quantiles is the same, against its negation. For example, the capital asset pricing model (Sharpe, 1964; Lintner, 1965) predicts a positive relation between expected returns and betas. Yet, there is some empirical evidence about a negative relation, so that a strategy based on selling (buying) high (low) beta stocks gives positive returns (Frazzini and Pedersen, 2014). The literature has acknowledged problems arising with these tests. For example, Fama and MacBeth (1973), when testing the CAPM, find over-rejection of the null hypothesis due to a positive bias on high percentiles and a negative on low percentiles. Lo and MacKinlay (1988) point out over-rejection due to possible correlation between the sorting variable and the regression residuals. As a remedy to this snoopy effect, authors such as Liang (2000) have suggested to randomly pick a subset of the sample for sorting variable estimation and use the rest of the sample for

test statistic construction. Cattaneo et al. (2022) study the statistical properties of beta-sorted portfolios and provide conditions ensuring consistency and asymptotic normality, along with uniform inference procedures allowing for uncertainty quantification and general hypothesis testing. Lopez de Prado and Zoonekynd (2024) study the problem of specification errors — namely, omitted variable bias and colliders — on the performance of factor investing strategies, and propose ways to address those errors.

Sorting can be done according to observable or unobservable characteristics. In the latter case the sorting variable needs to be estimated, and the measurement error leads to two types of classification errors when constructing the portfolios. First, we may include in a portfolio stocks that should be included in a lower or higher portfolio if we could observe the true sorting variable. Second, we may discard from a given portfolio stocks that should instead be included. The former can be seen as a contamination error, and the latter as an information loss error. Heuristically, if the expected return does not display a systematic pattern with the sorting variable, these two forms of misclassification should not affect the outcome of the test, while they should under the alternative hypothesis of a systematic pattern. In this paper, we formally investigate the effect of sorting errors on the test outcome.

Moreover, quantiles are not observable and are approximated by their corresponding order statistics, e.g. the j -th decile is approximated by the $jN/10$ smallest observations of the sorting variable. Hence, even when sorting according to an observable characteristic, a source of error may be given by the discrepancy between order statistics and quantile. We establish that such error is asymptotically negligible only when the sorting variable does not display strong cross-sectional correlation. Otherwise, it contributes to the asymptotic variance. Hence, inference based on "naive" standard errors may be incorrect even in the no measurement error case.

For the case of estimated sorting variables, we suggest an adaptive trimming rule to adjust the test statistic by discarding a fraction of stocks just above and below each order statistic, and correct the measurement error. Needless to say, the larger is the fraction we discard, the smaller the chance of contamination error but the larger the information loss. We establish sufficient conditions for performing tests for equalities of returns across percentiles which have correct asymptotic size and unit power. The extra cost is that we also need to take into account the contribution to the asymptotic variance due to the difference between percentiles based on the estimated and the true sorting variable. Importantly, we show that the difference between the infeasible statistics, based on the true sorting variable, and the trimmed statistics

is asymptotically zero mean normal under the null but diverges to plus or minus infinity under the alternative.

Throughout the paper we consider a realistic environment in which returns are observed at a daily frequency, and new stocks can enter and others exit on any given day. This gives rise to an unbalanced panel. At every rebalancing date, we use a rolling window of past returns to construct an estimate of the sorting variable for each stock and compute the relative order statistics. We use end-of-period returns over different percentiles to construct the test. Hence, our testing procedure is genuinely out of sample.

As a leading example, we consider sorting stocks according to their (estimated) market beta. We calibrate the simulation design to the historical distribution of estimated betas and study both the classification-error problem and the finite-sample behavior of the testing procedure. First, treating rolling beta estimates as the latent true sorting variable, we quantify contamination and information-loss errors in the extreme portfolios and show that the proposed trimming rule substantially reduces contamination. Second, using simulated return panels generated under the null and the alternative, and under both weak and strong cross-sectional correlation of the sorting variable, we compare the untrimmed and trimmed test statistics. The simulations show that contamination makes the untrimmed statistic oversized under the null and underpowered under the alternative, whereas trimming combined with subsampling moves the rejection frequencies in the direction predicted by the theory, delivering tests with size closer to the nominal level and power closer to one.

The rest of the paper is organized as follows. Section 3.2 describes the set-up. Section 3.3 considers the case of sorting without measurement error. Section 3.4 considers the case of sorting with measurement error, introduces our trimming rule and establishes the asymptotic properties of the trimmed statistics. In particular, we show how percentile measurement error contributes to the asymptotic variance. As the latter is not always known in closed form, in Section 3.5 we suggest how to construct asymptotically valid subsampling critical values, which properly mimic estimation error. Section 3.7 studies the case of sorting according to betas. All proofs are gathered in the Appendix.

3.2 Set-Up

We denote the time span by T , expressed in days, and the universe of stocks by N . Over time, stocks may cease to be traded because of defaults, mergers and acquisitions or de-listings, and in the meantime new stocks may start to be traded because of IPOs or re-listings. Therefore, we need to consider an unbalanced panel, with a different number of stocks present at each day. Let $T_i = \bar{T}_i - \underline{T}_i + 1$ denote the lifespan of stock i , so that at any day t we have $N_t = \sum_{i=1}^N \mathbb{1} \left\{ \underline{T}_i \leq t \leq \bar{T}_i \right\}$ stocks.

The practice of sorting involves forming portfolios at the end of each rebalancing period, according to a sorting variable which can be either observable or unobservable. In the remainder of the paper, the number N of portfolios is fixed a priori, and does not change with the sample size. Popular choices in the literature are deciles, quintiles or terciles¹. Moreover, for simplicity and without loss of generality, we take the rebalancing period to be a month, which we assume to be made up of $\mathcal{T}$ days. At rebalancing day $\tau(t) = \mathcal{T} \lceil t/\mathcal{T} \rceil$ there are $N_{\tau(t)} = \sum_{i=1}^N \mathbb{1} \left\{ \underline{T}_i \leq \tau(t) \leq \bar{T}_i \right\}$ stocks actively traded. Moreover, when the sorting variable is unobservable, we use a window of observations of R days to obtain an estimator. For simplicity, we assume R to be a multiple of $\mathcal{T}$ ². In addition, we need to define the following variables:

- For $i = 1, \dots, N_{\tau(t)}$, $S_{i,\tau(t)}$ (resp. $\widehat{S}_{i,\tau(t),R}$) is the observable (resp. estimated) sorting variable at the rebalancing day $\tau(t)$.
- For $j = 1, \dots, N$, $S_{\tau(t)}^{p(j)}$ (resp. $\widehat{S}_{\tau(t),R}^{p(j)}$) is the $jN_{\tau(t)}/N$ -th order statistic of $S_{i,\tau(t)}$ (resp. $\widehat{S}_{i,\tau(t),R}$) approximating the true unobservable upper bound of the j -th portfolio.
- For $j = 1, \dots, N$, $S_{\tau(t)}^{q(j)}$ is the $j100/N$ -th quantile of the sorting variable $S_{i,\tau(t)}$ (or $\widehat{S}_{i,\tau(t),R}$) determining the true unobservable upper bound of the j -th portfolio.
- For $j = 1, \dots, N$,

$$\mu^{q(j)} = \lim_{T, N \rightarrow \infty} \frac{N}{T - \mathcal{T}} \sum_{t=\mathcal{T}+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \mathbb{E} (r_{i,t}) \mathbb{1} \left\{ S_{\tau(t)}^{q(j-1)} < S_{i,\tau(t)} \leq S_{\tau(t)}^{q(j)} \right\}$$

is the expected return of the j -th quantile portfolio, where $r_{i,t}$ denotes the simple return

¹Cattaneo et al. (2020) develop a rule for finding the optimal number of portfolios, in terms of bias-variance tradeoff of the estimator of the mean return of a percentile portfolio. The higher the number of percentiles, the smaller the bias but the higher the variance (because of a smaller number of stocks present in the portfolio).

²This simplifying assumption allows to write the first rebalancing date as R as opposed to $\tau(R)$.

of stock i over the interval $[t, t + 1]^3$.

The objective is to test the null hypothesis that returns are independent of the sorting variable, against its negation. Focusing on the bottom and top portfolios, the null and alternative hypotheses are:

$$H_0 : \mu^{q(1)} - \mu^{q(N)} = 0, \quad (3.1)$$

$$H_A : \mu^{q(1)} - \mu^{q(N)} \neq 0. \quad (3.2)$$

3.3 Sorting Without Measurement Error

We begin by considering the case where the sorting variable is observable and the only source of error is the discrepancy between the order statistic and the unobservable quantile, i.e. there is no measurement error. This occurs, for example, when stocks are sorted according to size, or book to market ratio. At each rebalancing date $\tau(t) \geq 0$, we use the available $N_{\tau(t)}$ stocks to compute the order statistics $S_{\tau(t)}^{p(j)}$ for $j = 1, \dots, N$.

3.3.1 Asymptotic Results Under Weak Correlation of the Sorting Variable

In the sequel, we need the following definition and assumptions.

Definition 3.1 *The test statistic based on the observable sorting variable is given by:*

$$Z_{T,N} = \frac{N}{\sqrt{T - \mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(1)} \right\} - \mathbb{1} \left\{ S_{\tau(t)}^{p(N-1)} < S_{i,\tau(t)} \right\} \right).$$

Notice that the test statistic is defined as the average return over time of the 1-st quantile portfolio minus the average return over time of the N -th quantile portfolio (where the composition of the portfolios changes at each rebalancing day), multiplied by the square root of the number of observations. This formulation allows us to apply the central limit theorem on the time dimension to derive our results.

Assumption 3.1

(i) For $i = 1, \dots, N$ and $t \in T_i$, $r_{i,t}$ and $S_{i,\tau(t)}$ are strictly stationary and strong mixing with size

$$\alpha(k) = C\lambda^{-k}, \lambda > \frac{4s}{s-2}, s > 2.$$

³Note that if the sorting variable is estimated, we replace $\mathcal{T}$ with R . Moreover, the notation $\sum_{i=1}^{N_{\tau(t)}}$ indicates the sum over the $N_{\tau(t)}$ stocks traded at time $\tau(t)$, which are not necessarily the first $N_{\tau(t)}$ stocks in the panel.

(ii) For $i = 1, \dots, N$ and $t \in T_i$, $E \left(r_{i,t}^{2\kappa} \right) < \infty$ and $E \left(S_{i,\tau(t)}^{2\kappa} \right) < \infty$, $\kappa > 2$.

(iii) For $i, j = 1, \dots, N$ and $t \in T_i \cap T_j$, $\inf_{i,j} \text{cov} \left(r_{i,t}, r_{j,t} \right) \geq \varepsilon > 0$.

(iv) For $t \in T$, $\frac{1}{N^\rho} \sum_{i=1}^N \left(S_{i,\tau(t)} - E \left(S_{i,\tau(t)} \right) \right) = O_p(1)$, $1/2 \leq \rho < 1$.

Assumption 3.1.(i) is a distribution and memory condition. In particular, time stationarity of returns is assumed for simplicity and some forms of time heterogeneity could be introduced at the expense of extra conditions controlling its degree. The same conditions on the sorting variables imply that the quantile does not depend on time. The mixing condition on returns instead implies that the degree of serial correlation decreases exponentially. The same conditions on the sorting variable ensure that the order statistics converge to the unobservable quantile. Assumption 3.1.(ii) is a moment condition. In particular, it imposes uniform bounds on the moments of order higher than four, and rules out time trends. Assumption 3.1.(iii) implies strong cross-sectional correlation among stock returns. In particular, stock returns are assumed to be positively correlated as they are exposed to common systematic factors (such as the market factor)⁴. Assumption 3.1.(iv) requires weak cross-sectional correlation of the sorting variable. In particular, we assume that the cross-sectional sum of the sorting variables satisfies the central limit theorem when adequately standardised by $N_{\tau(t)}^\rho$. When the sorting variables are cross-sectionally independent (or weakly dependent), the appropriate value is $\rho = 1/2$ and the central limit theorem applies. This is equivalent to stating that none (or a small number) of covariances are non-zero.

Assumption 3.2

(i) $\inf_{t \in T} 0 < N_{\tau(t)}/N \leq 1$.

(ii) $\inf_{i \leq N} R < T_i$.

Assumption 3.2.(i) requires that a positive fraction of all stocks is traded at any time. Assumption 3.2.(ii) imposes a lower bound on the lifetime of each stock.

Theorem 3.1 studies the asymptotic behavior of the test statistic under assumption 3.1.(iv) of weak cross-sectional correlation of the sorting variable.

⁴Notice that if the returns were uncorrelated (or correlated with just a finite number of stocks), then the test statistic would need to be multiplied also by $\sqrt{N_{\tau(t)}/N}$.

Theorem 3.1 *Let assumptions 3.1 and 3.2 hold.*

(i) Under H_0 :

$$Z_{T,N} \xrightarrow{d} N(0, V),$$

where:

$$V = \lim_{N,T \rightarrow \infty} \text{var} \left(\frac{N}{\sqrt{T - \mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq S^{q(1)} \right\} - \mathbb{1} \left\{ S^{q(N-1)} < S_{i,\tau(t)} \right\} \right) \right).$$

(ii) Under H_A , there exists some $\eta > 0$ such that:

$$\lim_{N,T \rightarrow \infty} \Pr \left(\frac{1}{\sqrt{T - \mathcal{T}}} |Z_{N,T}| > \eta \right) = 1.$$

The theorem states that under the assumption of weak cross-sectional correlation of the sorting variable the quantile estimation error $S_{\tau(t)}^{p(j)} - S^{q(j)}$ does not contribute to the limiting distribution. Indeed, V is the variance obtained by replacing $S_{\tau(t)}^{p(1)}$ and $S_{\tau(t)}^{p(N-1)}$ with $S^{q(1)}$ and $S^{q(N-1)}$ in $Z_{T,N}$. Under the alternative hypothesis, the test statistic diverges at the rate $\sqrt{T}$.

3.3.2 Asymptotic Results Under Strong Correlation of the Sorting Variable

In the sequel, we need the following alternative assumption.

Assumption 3.3 *For $i, j = 1, \dots, N$ and all $t \in T_i \cap T_j$, $\inf_{i,j} \text{cov}(S_{i,\tau(t)}, S_{j,\tau(t)}) \geq \varepsilon > 0$.*

Assumption 3.3 requires strong cross-sectional correlation of the sorting variable. This would occur if the sorting variables have a common factor component. The appropriate value is $\rho = 1$ in assumption 3.1.(iv) or, equivalently, we can state that all covariances are non-zero.

Theorem 3.2 studies the asymptotic behavior of the test statistic under assumption 3.3 of strong cross-sectional correlation of the sorting variable.

Theorem 3.2 *Let assumptions 3.1.(i)–(iii), 3.2 and 3.3 hold.*

(i) Under H_0 :

$$Z_{T,N} \xrightarrow{d} N(0, V^*),$$

where:

$$\begin{aligned}
V^* = \lim_{N, T \rightarrow \infty} \text{var} & \left(\frac{\mathcal{N}}{\sqrt{T - \mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq S^{q(1)} \right\} - \mathbb{1} \left\{ S^{q(N-1)} < S_{i,\tau(t)} \right\} \right) \right. \\
& + \mathbb{E}(r_{i,t}) f \left(S^{q(1)} \right) \frac{\mathcal{N}}{\sqrt{T - \mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \left(S_{\tau(t)}^{p(1)} - S^{q(1)} \right) \\
& \left. - \mathbb{E}(r_{i,t}) f \left(S^{q(N-1)} \right) \frac{\mathcal{N}}{\sqrt{T - \mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \left(S_{\tau(t)}^{p(N-1)} - S^{q(N-1)} \right) \right),
\end{aligned}$$

where $f(x)$ denotes the density of the sorting variable evaluated at x .

(ii) Under H_A , there exists some $\eta > 0$ such that:

$$\lim_{N, T \rightarrow \infty} \Pr \left(\frac{1}{\sqrt{T - \mathcal{T}}} |Z_{N,T}| > \eta \right) = 1.$$

The theorem states that under the assumption of strong cross-sectional correlation of the sorting variable there is an additional term in the asymptotic covariance matrix. Intuitively, under weak cross-sectional dependence, the cross-section helps estimate the cutoff: the order-statistic error shrinks with N , so its normalized time-summed contribution is $o_p(1)$. Under strong cross-sectional dependence, a common component makes the cutoff errors move together across stocks, so the cross-section no longer provides an additional rate of convergence. The term $(T - \mathcal{T})^{-1/2} \sum_t (S_{\tau(t)}^{p(j)} - S^{q(j)})$ is no longer negligible and must enter the asymptotic variance.

3.4 Sorting With Measurement Error

We now turn to the case where the sorting variable is not observable and needs to be estimated. In this case, the source of error is given both by the measurement error and the discrepancy between the order statistic and the unobservable quantile. This occurs, for example, when stocks are sorted according to beta, momentum, idiosyncratic volatility, coskewness, or realized skewness and kurtosis. At each rebalancing date $\tau(t) \geq R$, we use rolling windows of the previous R observations to estimate $\widehat{S}_{i,\tau(t),R}$, and use the available $N_{\tau(t)}$ stocks to compute the estimated order statistics $\widehat{S}_{\tau(t),R}^{p(j)}$ for $j = 1, \dots, \mathcal{N}$.

3.4.1 Classification Errors

When forming the j -th quantile portfolio, we can make two possible classification errors. First, we can include in the j -th portfolio a stock which belongs to a different portfolio. This happens when for $j = 2, \dots, \mathcal{N} - 1$,

$$1 = \left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(j-1)} \right\} + \mathbb{1} \left\{ S_{\tau(t)}^{p(j)} < S_{i,\tau(t)} \right\} \right) \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(j-1)} < \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(j)} \right\}.$$

Similarly, for $j = 1, \mathcal{N}$,

$$1 = \mathbb{1} \left\{ S_{\tau(t)}^{p(1)} < S_{i,\tau(t)} \right\} \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1)} \right\},$$

$$1 = \mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(\mathcal{N}-1)} \right\} \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(\mathcal{N}-1)} < \widehat{S}_{i,\tau(t),R} \right\}.$$

This classification error can be seen as a contamination error, in the sense that stocks belonging to different portfolios contaminate the j -th one.

Second, we can fail to include in the j -th portfolio a stock which belongs to it. This happens when for $j = 2, \dots, \mathcal{N} - 1$,

$$1 = \left(\mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(j-1)} \right\} + \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(j)} < \widehat{S}_{i,\tau(t),R} \right\} \right) \mathbb{1} \left\{ S_{\tau(t)}^{p(j-1)} < S_{i,\tau(t)} \leq S_{\tau(t)}^{p(j)} \right\}.$$

Similarly, for $j = 1, \mathcal{N}$,

$$1 = \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(1)} < \widehat{S}_{i,\tau(t),R} \right\} \mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(1)} \right\},$$

$$1 = \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(\mathcal{N}-1)} \right\} \mathbb{1} \left\{ S_{\tau(t)}^{p(\mathcal{N}-1)} < S_{i,\tau(t)} \right\}.$$

This classification error can be seen as an information loss error, in the sense that stocks belonging to the j -th portfolio are discarded from it.

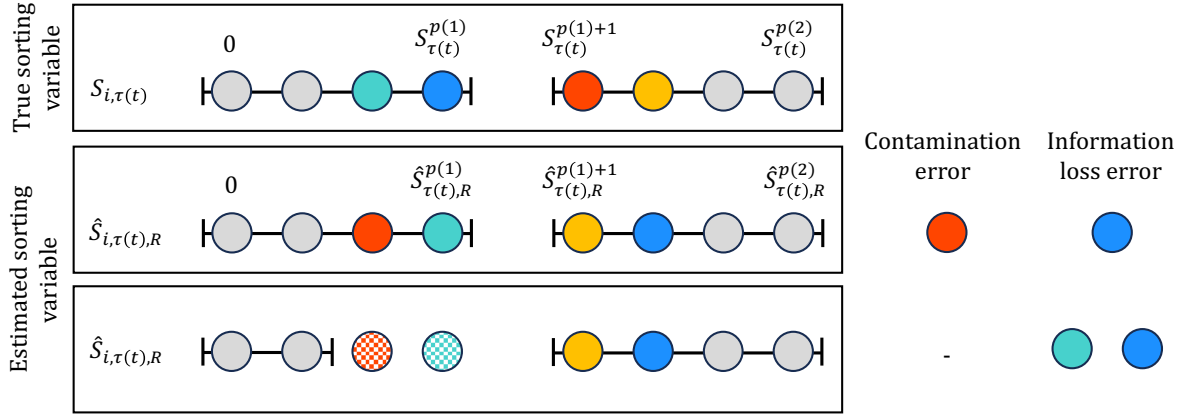


Figure 3.1: Contamination error and information loss error for portfolio 1.

We cannot jointly control both types of errors. Consider the example in figure 3.1, which shows how eight stocks (represented by bubbles) are allocated to two portfolios according to their true but unobservable sorting variable (top plot) and the estimated one (center plot). Four of these stocks have been coloured to highlight how the estimation of the sorting variable can lead to classification errors. The sorting variable for uncoloured stocks is assumed to be perfectly estimated. Focusing on the first portfolio, we can see how we incorrectly included the red stock (contamination error), while we failed to include the blue one (information loss error). We then trim away the stocks that are close to the boundary with the second portfolio, i.e. those marked with a chequered pattern (bottom plot). We can see how we can completely eliminate the contamination error (there are no stocks incorrectly included in the first portfolio), although this comes at the expense of a higher information loss error (as we now fail to include the green stock). If the objective is to test the hypotheses in equations (3.1) and (3.2), and in general to compare mean returns across portfolios, then it is more important to control the probability of contamination error.

3.4.2 Trimming Rule

Our aim is to establish an adaptive trimming rule which eliminates the contamination error while minimizing the information loss error. In doing so, we need to account for the fact that, if we do not trim away enough stocks, then we form contaminated portfolios; but if we trim away too many, then we entail a big loss of information. We might reduce the contamination error if, for a given portfolio j , we set the lower threshold above $\hat{S}_{\tau(t),R}^{p(j-1)}$ and the upper threshold below $\hat{S}_{\tau(t),R}^{p(j)}$. We look for almost sure lower and upper bounds for the errors, $c_{i,\tau(t),R}^L$ and $c_{i,\tau(t),R}^U$, such

that for each $i = 1, \dots, N_{\tau(t)}$ and each $\tau(t)$:

$$\Pr \left(\liminf_{R \rightarrow \infty} \left(\widehat{S}_{i,\tau(t),R} + c_{i,\tau(t),R}^L \right) \leq S_{i,\tau(t)} \leq \limsup_{R \rightarrow \infty} \left(\widehat{S}_{i,\tau(t),R} + c_{i,\tau(t),R}^U \right) \right) = 1, \quad (3.3)$$

where $c_{i,\tau(t),R}^L < 0$ and $c_{i,\tau(t),R}^U > 0$ ⁵. Hereafter, we set $-c_{i,\tau(t),R}^L = c_{i,\tau(t),R}^U = c_{i,\tau(t),R}$. Now, at each rebalancing day $\tau(t) \geq R$, let $I_{i,\tau(t),R}^j = \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(j-1)} < \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(j)} \right\}$, and define for $j = 2, \dots, \mathcal{N} - 1$,

$$\widehat{S}_{i,\tau(t),R}^{j,U} = \widehat{S}_{i,\tau(t),R} + I_{i,\tau(t),R}^j \sup_{i \text{ s.t. } I_{i,\tau(t),R}^j=1} c_{i,\tau(t),R},$$

$$\widehat{S}_{i,\tau(t),R}^{j,L} = \widehat{S}_{i,\tau(t),R} - I_{i,\tau(t),R}^j \sup_{i \text{ s.t. } I_{i,\tau(t),R}^j=1} c_{i,\tau(t),R}.$$

Similarly, let $I_{i,\tau(t),R}^1 = \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1)} \right\}$ and $I_{i,\tau(t),R}^{\mathcal{N}} = \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(\mathcal{N}-1)} < \widehat{S}_{i,\tau(t),R} \right\}$, and define for $j = 1, \mathcal{N}$,

$$\widehat{S}_{i,\tau(t),R}^{1,L} = \widehat{S}_{i,\tau(t),R} - I_{i,\tau(t),R}^1 \sup_{i \text{ s.t. } I_{i,\tau(t),R}^1=1} c_{i,\tau(t),R},$$

$$\widehat{S}_{i,\tau(t),R}^{\mathcal{N},U} = \widehat{S}_{i,\tau(t),R} + I_{i,\tau(t),R}^{\mathcal{N}} \sup_{i \text{ s.t. } I_{i,\tau(t),R}^{\mathcal{N}}=1} c_{i,\tau(t),R}.$$

Hereafter, we also set $c_{j,\tau(t),R} = \sup_{i \text{ s.t. } I_{i,\tau(t),R}^j=1} c_{i,\tau(t),R}$. Now, let $\widehat{S}_{\tau(t),R}^{p(j),U}$ and $\widehat{S}_{\tau(t),R}^{p(j),L}$ be the $jN_{\tau(t)}/\mathcal{N}$ -th order statistics of $\widehat{S}_{i,\tau(t),R}^{j,U}$ and $\widehat{S}_{i,\tau(t),R}^{j,L}$ respectively. When constructing the j -th portfolio, our trimming rule requires that we replace $\widehat{S}_{\tau(t),R}^{p(j-1)}$ with $\widehat{S}_{\tau(t),R}^{p(j-1),U}$ and $\widehat{S}_{\tau(t),R}^{p(j)}$ with $\widehat{S}_{\tau(t),R}^{p(j),L}$. By doing so, at any rebalancing time $\tau(t)$, we trim from the j -th portfolio all stocks $i = 1, \dots, N_{\tau(t)}$ for which either:

$$\widehat{S}_{\tau(t),R}^{p(j-1)} < \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(j-1),U},$$

⁵If the sorting variable depends on a common factor $F_{\tau(t)}$ (i.e. when the sorting variable displays strong cross-sectional correlation), the upper and lower bounds would also be $F_{\tau(t)}$ -measurable functions, and they need to hold almost surely conditional on $F_{\tau(t)}$. In the sequel, for brevity, we suppress the possible dependence on $F_{\tau(t)}$ in the notation.

or,

$$\widehat{S}_{\tau(t),R}^{p(j),L} < \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(j)}.$$

Note that trimming varies across quantiles. Figure 3.2 illustrates the trimming procedure for portfolio $j = 2, N - 1$. An analogous reasoning holds for portfolios $j = 1, N$, except that in this case the trimming is one-sided.

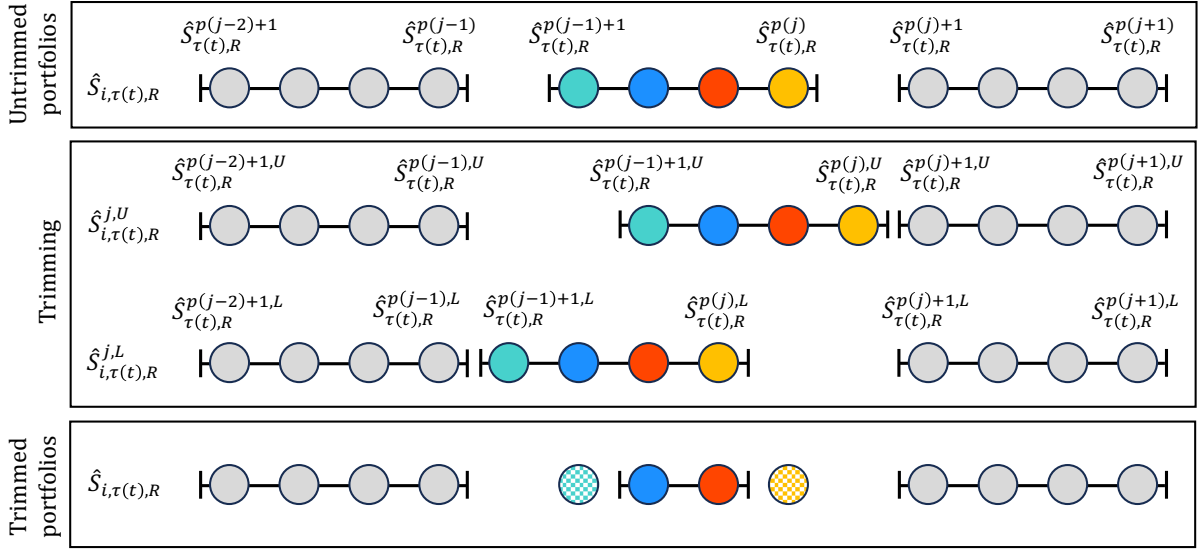


Figure 3.2: Trimming procedure.

3.4.3 Asymptotic Results

We now define the feasible and trimmed counterpart of $Z_{T,N}$. Moreover, in order to properly control for sorting errors, we need three additional assumptions.

Definition 3.2 *The feasible and trimmed counterpart of $Z_{T,N}$ is given by:*

$$\widehat{Z}_{T,N,R} = \frac{N}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} - \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(N-1),U} < \widehat{S}_{i,\tau(t),R} \right\} \right).$$

Assumption 3.4 *For $\varepsilon \geq 0$ and all $i, i' \in N_{\tau(t)}$, let $S_{\tau(t)}^i$ (resp. $S_{\tau(t)}^{i'}$) denote the $i/N_{\tau(t)}$ -th (resp. $i'/N_{\tau(t)}$ -th) cross-sectional order statistic such that:*

$$0 < \lim_{N \rightarrow \infty} i'/N_{\tau(t)} \text{ and } \lim_{N \rightarrow \infty} i/N_{\tau(t)} < 1.$$

We require that, almost surely there is a constant c such that:

$$\inf_{t \leq T} \inf_{i, i'} |S_{\tau(t)}^i - S_{\tau(t)}^{i'}| \geq c \frac{|i - i'|}{N^{1+\varepsilon}}.$$

Assumption 3.4 requires that order statistics are sufficiently spread out. This is equivalent to requiring that the density of the sorting variable is bounded above zero everywhere in the interior of its support. In other words, the order statistics should not be excessively correlated (such as when they have a common factor and very small variance). Intuitively, if the order statistics are too close, the trimming procedure would need to exclude a lot of stocks to have some effect. In particular, if $S_{\tau(t)}^i$ were independently and identically distributed, the assumption would hold with $\varepsilon = 0$, as the distance between the order statistics is of order $1/N$. In making this assumption, we do not consider extreme order statistics, as the latter would cause problems in the estimation. See chapter 1 in [Reiss \(2012\)](#) for further details.

Assumption 3.5 For $j = 1, \dots, N$, there exists $c_{j, \tau(t), R}^L$ and $c_{j, \tau(t), R}^U$ such that equation (3.3) holds. Moreover $c_{j, \tau(t), R}^L = -c_{j, \tau(t), R}^U = c_{j, \tau(t), R}$.

Assumption 3.5 requires the existence of almost sure bounds. This is a quite high level assumption. However, whenever the estimated sorting variable is a (function of a) sample moment, the existence of an almost sure bound $c_{i, \tau(t), R}$ is guaranteed by the law of iterated logarithm (LIL), which gives bounds for the lim inf and lim sup of the estimation error. Note that the LIL applies under mild additional memory and moment conditions (see for example [Eberlein \(1986\)](#)).

Assumption 3.6 Let

$$C_{1, \tau(t), R} = \left\{ i : \mathbb{1} \left\{ \widehat{S}_{\tau(t), R}^{p(1), L} \leq \widehat{S}_{i, \tau(t), R} \right\} \mathbb{1} \left\{ S_{i, t} \leq S_{\tau(t)}^{p(1)} \right\} = 1 \right\},$$

$$C_{N, \tau(t), R} = \left\{ i : \mathbb{1} \left\{ \widehat{S}_{i, \tau(t), R} \leq \widehat{S}_{\tau(t), R}^{p(N-1), U} \right\} \mathbb{1} \left\{ S_{\tau(t)}^{p(N-1)} \leq S_{i, t} \right\} = 1 \right\}.$$

We require that:

$$\frac{N}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \left(\sum_{i \in C_{1, \tau(t), R}} (r_{i, t} - \mu_i) - \sum_{i \in C_{N, \tau(t), R}} (r_{i, t} - \mu_i) \right) = o_p(1)$$

Assumption 3.6 requires that the information loss induced by trimming be asymptotically

balanced across the two extreme portfolios. The sets $C_{1,\tau(t),R}$ and $C_{N,\tau(t),R}$ collect the stocks that truly belong to the first and last portfolios, respectively, but are discarded by the trimming rule. The assumption then imposes that, after averaging over time and normalizing by $\sqrt{T-R}$, the difference between the demeaned return contributions of these discarded stocks is asymptotically negligible. Intuitively, trimming may remove some correctly classified stocks from both tails, but this loss of information must not generate a first-order distortion in the bottom-minus-top spread.

We proceed to study the effect of the measurement error in the sorting variable. Theorem 3.3 quantifies the difference between the trimmed feasible and the infeasible statistics, when the sorting variables are either weakly or strongly cross-sectionally correlated.

Theorem 3.3 *Let either assumptions 3.1,3.2,3.4-3.6 or 3.1.(i)–(iii), 3.2-3.6 hold:*

(i) *Under H_0 , as $T, N, R \rightarrow \infty$:*

$$\widehat{Z}_{T,N,R} - Z_{T,N} = \frac{Nm_1}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{p(1)} - S_{\tau(t)}^{p(1)} \right) - \frac{Nm_N}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{p(N-1)} - S_{\tau(t)}^{p(N-1)} \right) + o_p(1).$$

(ii) *Under H_A , as $T, N, R \rightarrow \infty$:*

$$\begin{aligned} \widehat{Z}_{T,N,R} - Z_{T,N} &= \frac{Nm_1}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{p(1)} - S_{\tau(t)}^{p(1)} \right) - \frac{Nm_N}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{p(N-1)} - S_{\tau(t)}^{p(N-1)} \right) + o_p(1) \\ &\quad + \frac{N}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \left(\sum_{i \in C_{1,\tau(t),R}} \mu_i - \sum_{i \in C_{N,\tau(t),R}} \mu_i \right). \end{aligned}$$

The theorem states that, under the null hypothesis, assumption 3.6 makes the centered information loss component asymptotically negligible, while the corresponding mean term vanishes because all stocks have the same expected return under the null. Therefore, the only first-order difference between the trimmed feasible and infeasible statistics is due to measurement error in the sorting variable, namely the discrepancies $\widehat{S}_{\tau(t),R}^{p(1)} - S_{\tau(t)}^{p(1)}$ and $\widehat{S}_{\tau(t),R}^{p(N-1)} - S_{\tau(t)}^{p(N-1)}$. Under the alternative hypothesis, instead, the difference between the two statistics contains an additional term capturing the difference in the average expected returns of the stocks discarded from the first and last portfolios by trimming. Given that the estimator of the unobservable sorting variable is based on R observations, heuristically the contamination error vanishes if $(T-R)/R \rightarrow 0$. This is formally established in corollary 3.1.

Corollary 3.1

(i) Under H_0 , as $T, N, R \rightarrow \infty$:

(a) If $\frac{T-R}{R} \rightarrow 0$ and either assumptions 3.1, 3.2, 3.4, 3.5, or 3.1.(i)–(iii), 3.2–3.5 hold:

$$\widehat{Z}_{T,N,R} = Z_{T,N} + o_p(1). \quad (3.4)$$

(b) If $\frac{T-R}{R} \rightarrow \pi$ where $0 < \pi < \infty$ and assumptions 3.1, 3.2, 3.4, 3.5 hold:

$$\widehat{Z}_{T,N,R} \xrightarrow{d} N(0, \Omega),$$

where

$$\Omega = V + V_{Sq(1)} + V_{Sq(N)} + covs$$

where V is defined as in theorem 3.1 and,

$$V_{Sq(1)} = \lim_{N,T,R \rightarrow \infty} \text{var} \left(\frac{Nm_1}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{q(1)} - S_{\tau(t)}^{q(1)} \right) \right),$$

$V_{Sq(N)}$ is defined analogously, and $covs$ denotes the covariance between the statistic in absence of measurement error and the measurement error component.

(c) If $\frac{T-R}{R} \rightarrow \pi$ where $0 < \pi < \infty$ and assumptions 3.1.(i)–(iii), 3.2–3.5 hold:

$$\widehat{Z}_{T,N,R} \xrightarrow{d} N(0, \Omega^*),$$

where

$$\Omega^* = V^* + V_{Sq(1)} + V_{Sq(N)} + covs$$

where V^* is defined as in theorem 3.2.

(ii) Under H_A , there exists $\eta > 0$ such that:

$$\lim_{T,N,R \rightarrow \infty} \Pr \left(\frac{1}{\sqrt{T-R}} |\widehat{Z}_{T,N,R}| \geq \eta \right) = 1$$

The corollary states that, under the null hypothesis, if $(T - R)/R \rightarrow 0$ then the measurement error (and hence the classification errors) is asymptotically negligible, and the trimmed statistic $\widehat{Z}_{T,N,R}$ is equivalent to the infeasible one $Z_{N,T}$. If instead $(T - R)$ and R grow at the same rate, then we need to take into account the quantile estimation error, due to the difference between quantiles computed using estimated or true sorting variables. More precisely, the variance is equal to the sum of the variance without the measurement error, the variance due to the measurement error of the first and last quantiles (from theorem 3.3), and their covariance. It follows that the trimmed feasible statistic is unbiased, but has higher variance than the infeasible statistic.

In practice, it may be the case that $T - R$ is much larger than the rolling window R , so that $(T - R)/R \rightarrow \infty$ seems to be a more appropriate description of the relative rate of growth. In this case we need to use a different scaling for the feasible statistic. Let $\widetilde{Z}_{T,N,R} = \sqrt{\frac{R}{T-R}} \widehat{Z}_{T,N,R}$. From the proofs of theorem 3.3 and corollary 3.1 it follows that:

$$\widetilde{Z}_{T,N,R} = \frac{Nm_1\sqrt{R}}{T-R} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{p(1)} - S_{\tau(t)}^{p(1)} \right) - \frac{Nm_N\sqrt{R}}{T-R} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{p(N-1)} - S_{\tau(t)}^{p(N-1)} \right) + o_p(1).$$

Therefore:

$$\widetilde{Z}_{T,N,R} \xrightarrow{d} N(0, V_{Sq(1)} + V_{Sq(N)} + covs),$$

where covs here denotes the covariance between lowest and highest percentile estimation error.

3.5 Subsampling

As established in corollary 3.1, unless $(T - R)/R \rightarrow 0$, we need to take into account the contribution to the asymptotic variance due to the difference between quantiles formed by sorting on the estimated and the true (but unobservable) sorting variable.

In general, we do not have a closed form expression for the covariance term in Ω and furthermore computation of $V_{Sq(1)}$ and $V_{Sq(N)}$ is not trivial. Hence, the natural solution is to rely on resampling techniques to get asymptotically valid critical values. Panel data with strong cross-sectional dependence can be dealt with using the block bootstrap proposed by Gonçalves (2011). However, her setup does not easily extend to unbalanced panels. Moreover, while we have observations for the returns at a daily frequency, sorting variables can be available at lower

frequencies (whether they are observed or estimated). Therefore we resort to subsampling to obtain asymptotically valid critical values.

Notice that Ω in corollary 3.1 depends explicitly on π , the speed of growth of $T - R$ relative to R . Therefore, it is crucial to keep the same ratio in the subsampling scheme. Accordingly, for $0 < \delta < 1$, we consider overlapping subsamples of $T_\delta = \lfloor T^\delta \rfloor$ days and we define $R_\delta = \lfloor RT_\delta/T \rfloor$ ⁶. Then, for all $\tau(t) \in [\underline{T}_i, \bar{T}_i]$ where $\tau(t) > \underline{T}_i + R_\delta$, we estimate $S_{i,\tau(t)}$ using rolling windows of R_δ observations in order to obtain $\widehat{S}_{i,\tau(t),R_\delta}$. Thus, we construct $K = T - T_\delta + 1$ statistics, which are defined as follows:

Definition 3.3 For $k = 1, \dots, K$, the subsampled counterpart of $\widehat{Z}_{T,N,R}$ is given by:

$$\widehat{Z}_{T_\delta,N,R_\delta}^{(k)} = \frac{N}{\sqrt{T_\delta - R_\delta}} \sum_{t=1+R_\delta+k}^{1+R_\delta+k+T_\delta} \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R_\delta} \leq \widehat{S}_{\tau(t),R_\delta}^{p(1),L_\delta} \right\} - \mathbb{1} \left\{ \widehat{S}_{\tau(t),R_\delta}^{p(N-1),U_\delta} < \widehat{S}_{i,\tau(t),R_\delta} \right\} \right).$$

Note that $\widehat{S}_{\tau(t),R_\delta}^{p(1),L_\delta}$ and $\widehat{S}_{\tau(t),R_\delta}^{p(N-1),U_\delta}$ are defined analogously to $\widehat{S}_{\tau(t),R}^{p(1),L}$ and $\widehat{S}_{\tau(t),R}^{p(N-1),U}$, respectively, except that R_δ is used in place of R . Finally, we construct the empirical distribution of $\widehat{Z}_{T_\delta,N,R_\delta}^{(1)}, \dots, \widehat{Z}_{T_\delta,N,R_\delta}^{(K)}$ and, for a given significance level $\alpha \in (0, 1)$, we denote its $\alpha/2$ and $1 - \alpha/2$ empirical quantiles by $c_{T_\delta,K,\alpha/2}$ and $c_{T_\delta,K,1-\alpha/2}$. Before stating our next result, we need an additional assumption.

Assumption 3.7 For each $k = 1, \dots, K$:

$$\lim_{T_\delta, N, R_\delta \rightarrow \infty} \text{var} \left(\widehat{Z}_{T_\delta, N, R_\delta}^{(k)} \right) = \lim_{T, N, R \rightarrow \infty} \text{var} \left(\widehat{Z}_{T, N, R} \right),$$

and for each $j = 1, \dots, N$:

$$\lim_{T_\delta, N, R_\delta \rightarrow \infty} \text{var} \left(\frac{N}{\sqrt{T_\delta}} \sum_{t=1+R_\delta+k}^{1+R_\delta+k+T_\delta} \left(\widehat{S}_{\tau(t),R_\delta}^{q(j)} - S^{q(j)} \right) \right) = \lim_{T, N, R \rightarrow \infty} \text{var} \left(\frac{N}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{q(j)} - S^{q(j)} \right) \right).$$

Assumption 3.7 mirrors condition (8) in Politis et al. (1997) for the validity of subsampling in the case of heteroskedastic observations.

⁶Notice that to keep the same ratio π , the following relation must hold: $\frac{T_\delta - R_\delta}{R_\delta} = \frac{T - R}{R}$. Solving the equality we obtain the expression for R_δ .

Theorem 3.4 establishes the first order validity of inference based on subsampling.

Theorem 3.4 *Let either assumptions 3.1,3.2,3.4-3.7, or 3.1.(i)–(iii), 3.2-3.7 hold.*

(i) *Under H_0 , if $\frac{T_\delta}{T} \rightarrow 0$, and $\frac{T-R}{R} \rightarrow \pi$ and $\frac{T_\delta-R_\delta}{R_\delta} \rightarrow \pi$ for $0 < \pi < \infty$ as $T, N, R, T_\delta, K, R_\delta \rightarrow \infty$:*

$$\lim_{T,N,R \rightarrow \infty} \Pr \left(c_{T_\delta, K, \alpha/2} \leq \widehat{Z}_{T,N,R} \leq c_{T_\delta, K, 1-\alpha/2} \right) = 1 - \alpha.$$

(ii) *Under H_A , under the same rate conditions as in part (i),*

$$\lim_{T,N,R \rightarrow \infty} \Pr \left(\left(\widehat{Z}_{T,N,R} < c_{T_\delta, K, \alpha/2} \right) \cup \left(c_{T_\delta, K, 1-\alpha/2} < \widehat{Z}_{T,N,R} \right) \right) = 1.$$

The theorem states that, under the null hypothesis, the subsampled statistics properly mimic the distribution of the (sample) statistic, regardless of whether the contribution of sorting error vanishes or not, as long as $\frac{T-R}{R} \rightarrow \pi$ with $0 < \pi < \infty$. Under the alternative hypothesis, the statistic diverges to plus or minus infinity at a faster rate than its subsampled counterpart, thus ensuring unit asymptotic power.

Finally, if instead $\frac{T-R}{R} \rightarrow \infty$, then inference should be based on the rescaled statistic $\widetilde{Z}_{T,N,R}$ and on critical values computed from the empirical distribution of $\sqrt{\frac{R_\delta}{T_\delta-R_\delta}} \widehat{Z}_{T_\delta, N, R_\delta}^{(k)}$.

3.6 A Motivating Example: Sorting on Stocks' Betas

The relationship between betas and expected returns has a long history in the asset pricing literature. The Sharpe-Lintner CAPM model (Sharpe, 1964; Lintner, 1965) predicts a positive association between betas and expected returns. Empirically there is widespread evidence of violations of the CAPM prediction (see, e.g. Black et al. (1972), Fama and French (1992) and Fama and French (2006)). More recently, Frazzini and Pedersen (2014) show that agents constrained in their leverage capability tend to overinvest in risky high-beta stocks and underinvest in low-beta stocks. This leads to a reduction of returns on high-beta stocks and determines the abnormal returns of the Betting Against Beta strategy. Moreover, the empirical evidence against the unconditional CAPM has sparked interest in considering conditional betas. Following, e.g., Shanken (1990) and Cederburg and O'Doherty (2016), we model betas as affine functions of stock characteristics and macroeconomic variables.

In this section, we show how sorting according to unconditional or conditional betas satisfies the existence of almost sure bounds under mild memory and moment conditions.

3.6.1 Unconditional Betas

We begin with the standard CAPM model, where:

$$r_{i,t} = \alpha_{i,\tau(t)} + \beta_{i,\tau(t)} r_{M,t} + \epsilon_{i,t}.$$

The reason why, even in the unconditional case, we let the true alpha and beta of stock i depend on $\tau(t)$ is that the pool of stocks alive varies over time, and so, with an abuse of notation, we allow stock i at time $\tau(t)$ to be different from stock i at time $\tau(s)$, with $s \neq t$. We estimate $\beta_{i,\tau(t)}$ by OLS for $i = 1, \dots, N_{\tau(t)}$, and all $\tau(t)$, using a window of R observations:

$$\begin{aligned} \hat{\beta}_{i,\tau(t),R} &= \left(\frac{1}{R} \sum_{s=\tau(t)-R+1}^{\tau(t)} (r_{M,s} - \bar{r}_{M,\tau(t),R})^2 \right)^{-1} \left(\frac{1}{R} \sum_{s=\tau(t)-R+1}^{\tau(t)} (r_{M,s} - \bar{r}_{M,\tau(t),R}) (r_{i,s} - \bar{r}_{i,\tau(t),R}) \right) \\ &= \hat{A}_{\tau(t),R}^{-1} \frac{1}{R} \sum_{s=\tau(t)-R+1}^{\tau(t)} (r_{M,s} - \bar{r}_{M,\tau(t),R}) (r_{i,s} - \bar{r}_{i,\tau(t),R}), \end{aligned} \quad (3.5)$$

where $\hat{A}_{\tau(t),R}$ is defined accordingly and

$$\bar{r}_{M,\tau(t),R} = \frac{1}{R} \sum_{s=\tau(t)-R+1}^{\tau(t)} r_{M,s} \quad \text{and} \quad \bar{r}_{i,\tau(t),R} = \frac{1}{R} \sum_{s=\tau(t)-R+1}^{\tau(t)} r_{i,s}.$$

We estimate the asymptotic variance of the rolling OLS estimator, $\text{avar}(\sqrt{R}(\hat{\beta}_{i,\tau(t),R} - \beta_{i,\tau(t)}))$, by a Newey–West HAC estimator computed over the same R -day rolling window:

$$\hat{\sigma}_{i,\tau(t),R}^2 = \hat{A}_{\tau(t),R}^{-1} \hat{v}_{i,\tau(t),R}^2 \hat{A}_{\tau(t),R}^{-1}, \quad (3.6)$$

where:

$$\begin{aligned} \hat{v}_{i,\tau(t),R}^2 &= \hat{\Gamma}_{i,\tau(t),R}(0) + 2 \sum_{h=1}^{l_R} \left(1 - \frac{h}{l_R + 1} \right) \hat{\Gamma}_{i,\tau(t),R}(h), \\ \hat{\Gamma}_{i,\tau(t),R}(h) &= \frac{1}{R} \sum_{u=\tau(t)-R+1+h}^{\tau(t)} \hat{\psi}_{i,u,\tau(t),R} \hat{\psi}_{i,u-h,\tau(t),R}, \\ \hat{\psi}_{i,u,\tau(t),R} &= (r_{M,u} - \bar{r}_{M,\tau(t),R}) \hat{\epsilon}_{i,u,\tau(t),R}, \end{aligned}$$

$$\widehat{\epsilon}_{i,u,\tau(t),R} = r_{i,u} - \bar{r}_{i,\tau(t),R} - \widehat{\beta}_{i,\tau(t),R} (r_{M,u} - \bar{r}_{M,\tau(t),R}).$$

Next, we define:

$$\widehat{c}_{LIL,i,\tau(t),R} \equiv -\widehat{c}_{LIL,i,\tau(t),R}^L = \widehat{c}_{LIL,i,\tau(t),R}^U = \widehat{\sigma}_{i,\tau(t),R} \sqrt{\frac{2 \ln(\ln(R))}{R}}.$$

Finally, let $I_{i,\tau(t),R}^j = \mathbb{1} \left\{ \widehat{\beta}_{\tau(t),R}^{p(j-1)} < \widehat{\beta}_{i,\tau(t),R} \leq \widehat{\beta}_{\tau(t),R}^{p(j)} \right\}$, and define for $j = 2, \dots, N-1$,

$$\widehat{\beta}_{i,\tau(t),R}^{j,U} = \widehat{\beta}_{i,\tau(t),R} + I_{i,\tau(t),R}^j \sup_{i \text{ s.t. } I_{i,\tau(t),R}^j=1} \widehat{c}_{LIL,i,\tau(t),R},$$

$$\widehat{\beta}_{i,\tau(t),R}^{j,L} = \widehat{\beta}_{i,\tau(t),R} - I_{i,\tau(t),R}^j \sup_{i \text{ s.t. } I_{i,\tau(t),R}^j=1} \widehat{c}_{LIL,i,\tau(t),R}.$$

Similarly, let $I_{i,\tau(t),R}^1$ and $I_{i,\tau(t),R}^N$ be defined according to the same logic, and define for $j = 1, N$,

$$\widehat{\beta}_{i,\tau(t),R}^{1,L} = \widehat{\beta}_{i,\tau(t),R} - I_{i,\tau(t),R}^1 \sup_{i \text{ s.t. } I_{i,\tau(t),R}^1=1} \widehat{c}_{LIL,i,\tau(t),R},$$

$$\widehat{\beta}_{i,\tau(t),R}^{N,U} = \widehat{\beta}_{i,\tau(t),R} + I_{i,\tau(t),R}^N \sup_{i \text{ s.t. } I_{i,\tau(t),R}^N=1} \widehat{c}_{LIL,i,\tau(t),R},$$

and let $\widehat{\beta}_{\tau(t),R}^{p(j),U}$ and $\widehat{\beta}_{\tau(t),R}^{p(j),L}$ be the $jN_{\tau(t)}/N$ -th order statistics of $\widehat{\beta}_{i,\tau(t),R}^{j,U}$ and $\widehat{\beta}_{i,\tau(t),R}^{j,L}$ respectively.

Then, the trimmed statistic for the unconditional betas reads as:

$$\widehat{Z}\widehat{\beta}_{N,T,R} = \frac{N}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ \widehat{\beta}_{i,\tau(t),R} \leq \widehat{\beta}_{\tau(t),R}^{p(1),L} \right\} - \mathbb{1} \left\{ \widehat{\beta}_{\tau(t),R}^{p(N-1),U} \leq \widehat{\beta}_{i,\tau(t),R} \right\} \right). \quad (3.7)$$

In the sequel, we need the following assumptions.

Assumption B1 For $i = 1, \dots, N_{\tau(t)}$, $E |r_{M,t} \epsilon_{i,t}|^{2(4+\delta)} < C$, for some $\delta > 0$.

Assumption B2 For $i = 1, \dots, N_{\tau(t)}$, $r_{i,t}$ and $r_{M,t}$ are strictly stationary and strong mixing with size $-4(4+\delta)/\delta$, for the same δ as in Assumption B1.

Proposition B If Assumptions B1 and B2 hold:

(i)

$$\Pr\left(\liminf_{R \rightarrow \infty} \left(\widehat{\beta}_{i,\tau(t),R} + \widehat{c}_{LIL,i,\tau(t),R}^L\right) \leq \beta_{i,\tau(t)} \leq \limsup_{R \rightarrow \infty} \left(\widehat{\beta}_{i,\tau(t),R} + \widehat{c}_{LIL,i,\tau(t),R}^U\right)\right) = 1.$$

(ii)

$$\begin{aligned} \frac{1}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} & \left(\mathbb{1} \left\{ \widehat{\beta}_{i,\tau(t),R} \leq \widehat{\beta}_{\tau(t),R}^{p(1),L} \right\} \mathbb{1} \left\{ \widehat{\beta}_{\tau(t),R}^{p(1)} < \beta_{i,\tau(t)} \right\} \right. \\ & \left. - \mathbb{1} \left\{ \widehat{\beta}_{\tau(t),R}^{p(N-1),U} < \widehat{\beta}_{i,\tau(t),R} \right\} \mathbb{1} \left\{ \beta_{i,\tau(t)} \leq \widehat{\beta}_{\tau(t),R}^{p(N-1)} \right\} \right) = o_{a.s.}(1). \end{aligned}$$

3.6.2 Conditional Betas

We now establish the conditional beta counterpart of Proposition B. Stock returns are generated by:

$$r_{i,t} = \alpha_{i,\tau(t)} + \left(\beta_{0,i,\tau(t)} + \beta_{1,i,\tau(t)}^\top W_{i,t} \right) r_{M,t} + \epsilon_{i,t},$$

where $W_{i,t} \equiv W_{i,\tau(t)-\mathcal{T}}$, for $\tau(t)-\mathcal{T} < t \leq \tau(t)$, contains both common macroeconomic and stock specific variables. As before, we let $\alpha_{i,\tau(t)}$, $\beta_{0,i,\tau(t)}$ and $\beta_{1,i,\tau(t)}$ depend on $\tau(t)$, since what we denote as stock i may vary across time. Because most macroeconomic variables are not available on a daily basis, we allow the variables in the conditioning set to evolve at a lower frequency than stock returns. For each stock, we use rolling windows of R observations to construct:

$$\begin{aligned} \widehat{\theta}_{i,\tau(t),R} &= \left(\widehat{\beta}_{0,i,\tau(t),R}, \widehat{\beta}_{1,i,\tau(t),R}^\top \right) \\ &= \left(\frac{1}{R} \sum_{s=\tau(t)-R+1}^{\tau(t)} \left(X_{i,s} - \overline{X}_{i,\tau(t),R} \right) \left(X_{i,s} - \overline{X}_{i,\tau(t),R} \right)^\top \right)^{-1} \\ & \quad \left(\frac{1}{R} \sum_{s=\tau(t)-R+1}^{\tau(t)} \left(X_{i,s} - \overline{X}_{i,\tau(t),R} \right) \left(r_{i,s} - \bar{r}_{i,\tau(t),R} \right) \right), \end{aligned}$$

where $X_{i,s} = (r_{M,s}, W_{i,s} r_{M,s})^\top$, and $\overline{X}_{i,\tau(t),R}$ and $\bar{r}_{i,\tau(t),R}$ are averages over the rolling window. We can define the composite estimator:

$$\widehat{\beta}_{i,\tau(t),R}(W_{i,t}) = \widehat{\beta}_{0,i,\tau(t),R} + \widehat{\beta}_{1,i,\tau(t),R}^\top W_{i,t}, \quad (3.8)$$

and its unobserved counterpart:

$$\beta_{i,\tau(t)}(W_{i,t}) = \beta_{0,i,\tau(t)} + \beta_{1,i,\tau(t)}^\top W_{i,t}, \quad (3.9)$$

so that:

$$\beta_{i,\tau(t)}(W_{i,t}) - \widehat{\beta}_{i,\tau(t),R}(W_{i,t}) = \left(\beta_{0,i,\tau(t)} - \widehat{\beta}_{0,i,\tau(t),R} \right) + \left(\beta_{1,i,\tau(t)} - \widehat{\beta}_{1,i,\tau(t),R} \right)^\top W_{i,t}.$$

Clearly, when the conditioning set $W_{i,t}$ contains common (e.g. macroeconomic) factors, then $\beta_{i,\tau(t)}(W_{i,t})$ as defined in (3.9) exhibits strong cross-sectional correlation as in Assumption 3.3. For simplicity, suppose $W_{i,t}$ is a scalar, and $W_{i,t} = W_t$, i.e. the conditioning variable is a common factor. This case implies strong cross-sectional correlation. Let $\widehat{\sigma}_{\beta_{0,i,\tau(t),R}}^2$ and $\widehat{\sigma}_{\beta_{1,i,\tau(t),R}}^2$ be the (possibly HAC) estimators of $\text{avar}\left(\sqrt{R}\left(\widehat{\beta}_{0,i,\tau(t),R} - \beta_{0,i,\tau(t)}\right)\right)$ and of $\text{avar}\left(\sqrt{R}\left(\widehat{\beta}_{1,i,\tau(t),R} - \beta_{1,i,\tau(t)}\right)\right)$ respectively. Next, we define:

$$\begin{aligned} \widehat{c}_{LIL,i,\tau(t),R}^L(W_t) &\equiv -\widehat{c}_{LIL,i,\tau(t),R}^L(W_t) = \widehat{c}_{LIL,i,\tau(t),R}^U(W_t) \\ &= \widehat{\sigma}_{\beta_{0,i,\tau(t),R}} \sqrt{\frac{2 \ln \ln R}{R}} + \widehat{\sigma}_{\beta_{1,i,\tau(t),R}} \sqrt{\frac{2 \ln \ln R}{R}} |W_t|. \end{aligned}$$

Finally, let $I_{i,\tau(t),R}^j(W_t) = \mathbb{1} \left\{ \widehat{\beta}_{\tau(t),R}^{p(j-1)}(W_t) < \widehat{\beta}_{i,\tau(t),R}(W_t) \leq \widehat{\beta}_{\tau(t),R}^{p(j)}(W_t) \right\}$, and define for $j = 2, \dots, \mathcal{N} - 1$,

$$\widehat{\beta}_{i,\tau(t),R}^{j,U}(W_t) = \widehat{\beta}_{i,\tau(t),R}(W_t) + I_{i,\tau(t),R}^j(W_t) \sup_{i \text{ s.t. } I_{i,\tau(t),R}^j(W_t)=1} \widehat{c}_{LIL,i,\tau(t),R}(W_t),$$

$$\widehat{\beta}_{i,\tau(t),R}^{j,L}(W_t) = \widehat{\beta}_{i,\tau(t),R}(W_t) - I_{i,\tau(t),R}^j(W_t) \sup_{i \text{ s.t. } I_{i,\tau(t),R}^j(W_t)=1} \widehat{c}_{LIL,i,\tau(t),R}(W_t).$$

Similarly, let $I_{i,\tau(t),R}^1(W_t)$ and $I_{i,\tau(t),R}^{\mathcal{N}}(W_t)$ be defined according to the same logic, and define for $j = 1, \mathcal{N}$,

$$\widehat{\beta}_{i,\tau(t),R}^{1,L}(W_t) = \widehat{\beta}_{i,\tau(t),R}(W_t) - I_{i,\tau(t),R}^1(W_t) \sup_{i \text{ s.t. } I_{i,\tau(t),R}^1(W_t)=1} \widehat{c}_{LIL,i,\tau(t),R}(W_t),$$

$$\widehat{\beta}_{i,\tau(t),R}^{\mathcal{N},U}(W_t) = \widehat{\beta}_{i,\tau(t),R}(W_t) + I_{i,\tau(t),R}^{\mathcal{N}}(W_t) \sup_{i \text{ s.t. } I_{i,\tau(t),R}^j(W_t)=1} \widehat{c}_{LIL,i,\tau(t),R}(W_t),$$

and let $\widehat{\beta}_{\tau(t),R}^{p(j),U}(W_t)$ and $\widehat{\beta}_{\tau(t),R}^{p(j),L}(W_t)$ be the $jN_{\tau(t)}/\mathcal{N}$ -th order statistics of $\widehat{\beta}_{i,\tau(t),R}^{j,U}(W_t)$ and $\widehat{\beta}_{i,\tau(t),R}^{j,L}(W_t)$ respectively. Then, the trimmed statistic for the conditional betas reads as:

$$\begin{aligned} \widehat{Z}\widehat{\beta}_{W,N,T,R} = \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ \widehat{\beta}_{i,\tau(t),R}(W_t) \leq \widehat{\beta}_{\tau(t),R}^{p(1),L}(W_t) \right\} \right. \\ \left. - \mathbb{1} \left\{ \widehat{\beta}_{\tau(t),R}^{p(\mathcal{N}-1),U}(W_t) \leq \widehat{\beta}_{i,\tau(t),R}(W_t) \right\} \right). \end{aligned} \quad (3.10)$$

We need to strengthen Assumptions B1 and B2, as follows.

Assumption B1' For $i = 1, \dots, N_{\tau(t)}$, and $X_{i,s} = (r_{M,s}, W_s r_{M,s})$, all elements of $X_{i,s}$ have $2(4 + \delta)$ bounded moments, for $\delta > 0$.

Assumption B2' For $i = 1, \dots, N_{\tau(t)}$, and $X_{i,s} = (r_{M,s}, W_s r_{M,s})$, $r_{i,t}$, $W_{i,s} r_{M,s}$ and $r_{M,t}$ are strictly stationary and strong mixing with size $-4(4 + \delta)/\delta$, for the same δ as in Assumption B1.

Proposition B' If Assumptions B1' and B2' hold, then conditionally on W_t :

(i)

$$\begin{aligned} \Pr \left(\liminf_{R \rightarrow \infty} \left(\widehat{\beta}_{0,i,\tau(t),R} + \widehat{\beta}_{1,i,\tau(t),R} W_t + \widehat{c}_{LIL,\beta_{0,i,\tau(t),R}}^L + \widehat{c}_{LIL,\beta_{1,i,\tau(t),R}}^L |W_t| \right) \leq \right. \\ \left. \beta_{i,\tau(t)}(W_t) \leq \limsup_{R \rightarrow \infty} \left(\widehat{\beta}_{0,i,\tau(t),R} + \widehat{\beta}_{1,i,\tau(t),R} W_t + \widehat{c}_{LIL,\beta_{0,i,\tau(t),R}}^U + \widehat{c}_{LIL,\beta_{1,i,\tau(t),R}}^U |W_t| \right) \right) = 1 \end{aligned}$$

(ii)

$$\begin{aligned} \frac{1}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ \widehat{\beta}_{i,\tau(t),R}(W_t) \leq \widehat{\beta}_{\tau(t),R}^{p(1),L}(W_t) \right\} \mathbb{1} \left\{ \widehat{\beta}_{\tau(t),R}^{p(1)}(W_t) < \beta_{i,\tau(t)}(W_t) \right\} \right. \\ \left. - \mathbb{1} \left\{ \widehat{\beta}_{\tau(t),R}^{p(\mathcal{N}-1),U}(W_t) < \widehat{\beta}_{i,\tau(t),R}(W_t) \right\} \mathbb{1} \left\{ \beta_{i,\tau(t)}(W_t) \leq \widehat{\beta}_{\tau(t),R}^{p(\mathcal{N}-1)}(W_t) \right\} \right) \\ = o_{a.s.}(1) \end{aligned}$$

Note that the bounds become less and less sharp as the dimensionality of the conditioning set increases.

3.6.3 Test Statistic and Critical Values

We now establish that the statements in Corollary 3.1 apply when sorting over unconditional and conditional betas. This requires an additional assumption:

Assumption 3.8 For $\epsilon \geq 0$ and all $i, i' \in N_{\tau(t)}$, let $\beta_{\tau(t)}^j$ and $\beta_{\tau(t)}^{j'}(W_t)$ denote the j -th cross-sectional order statistic of $\beta_{i,\tau(t)}$, and $\beta_{i,\tau(t)}(W_t)$ respectively. We require that almost surely there is a constant C such that:

$$\inf_{t \leq T} \inf_{i, i'} \left| \beta_{\tau(t-1)}^j - \beta_{\tau(t-1)}^{j'} \right| \geq C \frac{|i - i'|}{N^{1+\epsilon}},$$

and a constant $C(W_t)$ such that, conditionally on W_t :

$$\inf_{t \leq T} \inf_{i, i'} \left| \beta_{\tau(t-1)}^j(W_t) - \beta_{\tau(t-1)}^{j'}(W_t) \right| \geq C(W_t) \frac{|i - i'|}{N^{1+\epsilon}}.$$

Theorem B1 Let Assumptions B1 (B1'), B2 (B2'), 3.2 and 3.8 hold.

(i) Under H_0 , as $N, T, R \rightarrow \infty$ and $(T - R)/R \rightarrow \pi > 0$,

$$(a) \quad \widehat{Z}\beta_{N,T,R} \xrightarrow{d} N(0, \Omega_\beta),$$

where:

$$\Omega_\beta = V_\beta + V_{\beta^{q(1)}} + V_{\beta^{q(N)}} + covs,$$

$V_\beta, V_{\beta^{q(1)}}, V_{\beta^{q(N)}}$ and $covs$ are defined as in Corollary 3.1, with $\widehat{S}_{i,\tau(t),R}, \widehat{S}_{\tau(t),R}^{q(j)}, S^{q(j)}$ replaced by $\widehat{\beta}_{i,\tau(t),R}, \widehat{\beta}_{\tau(t),R}^{q(j)}, \beta^{q(j)}$

$$(b) \quad \widehat{Z}\beta_{W,N,T,R} \xrightarrow{d} N(0, \Omega_{\beta_w}),$$

where:

$$\Omega_{\beta_w} = V_{\beta_w} + V_{\beta_w^{q(1)}} + V_{\beta_w^{q(N)}} + covs,$$

where $V_{\beta_w}, V_{\beta_w^{q(1)}}, V_{\beta_w^{q(N)}}$ and $covs$ are as in part (i), with $\widehat{\beta}_{i,\tau(t),R}, \widehat{\beta}_{\tau(t),R}^{q(j)}, \beta^{q(j)}$ replaced by $\widehat{\beta}_{i,\tau(t),R}(W_t), \widehat{\beta}_{\tau(t),R}^{q(j)}(W_t), \beta^{q(j)}(W_t)$.

(ii) Under H_A there exists $\delta > 0$ such that:

$$(a) \quad \lim_{T,N,R \rightarrow \infty} \Pr \left(\frac{1}{\sqrt{T-R}} \left| \widehat{Z}\beta_{N,T,R} \right| > \delta \right) = 1.$$

$$(b) \quad \lim_{T,N,R \rightarrow \infty} \Pr \left(\frac{1}{\sqrt{T-R}} \left| \widehat{Z\beta}_{W,N,T,R} \right| > \delta \right) = 1.$$

Finally, the construction of the subsampled critical values is based on:

$$\widehat{Z\beta}_{N,T_\delta,R_\delta}^{(k)} = \frac{N}{\sqrt{T_\delta - R_\delta}} \sum_{t=1+R_\delta+k}^{1+R_\delta+k+T_\delta} \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ \widehat{\beta}_{i,\tau(t),R_\delta} \leq \widehat{\beta}_{\tau(t),R_\delta}^{p(1),L_\delta} \right\} - \mathbb{1} \left\{ \widehat{\beta}_{\tau(t),R_\delta}^{p(N-1),U_\delta} < \widehat{\beta}_{i,\tau(t),R_\delta} \right\} \right).$$

where $\widehat{\beta}_{i,\tau(t),R_\delta}$ is constructed using rolling windows of $R_\delta = \lfloor RT_\delta/T \rfloor$ observations. Then, provided that $\widehat{\beta}_{\tau(t),R_\delta}^{q(j)}$ satisfies Assumption 3.8, the statement in Theorem 3.4 applies. The subsample conditional beta follows by replacing $\widehat{\beta}_{i,\tau(t),R_\delta}, \widehat{\beta}_{\tau(t),R_\delta}^{p(1),L_\delta}, \widehat{\beta}_{\tau(t),R_\delta}^{p(N-1),U_\delta}$ with $\widehat{\beta}_{i,\tau(t),R_\delta}(W_t), \widehat{\beta}_{\tau(t),R_\delta}^{p(1),L_\delta}(W_t), \widehat{\beta}_{\tau(t),R_\delta}^{p(N-1),U_\delta}(W_t)$.

3.7 Simulations

In this section, we continue the motivating example of the previous section, and use simulations to show the validity of the theoretical results. In particular, in the first part of this section, we simulate stock returns to quantify the contamination error and the information loss error, both in the presence and in the absence of trimming. In the second part of this section, we simulate stock returns to show that trimming, coupled with the subsampling procedure, allows to construct a test statistic that has the right asymptotic size and unit power.

3.7.1 Data

We use data from the CRSP U.S. Stock database, which contains end-of-day prices on equity securities listed on U.S. stock exchanges. From the full universe, we retain only ordinary common shares (share codes 10 or 11) traded on the New York Stock Exchange, the American Stock Exchange, and the Nasdaq Stock Market (exchange codes 1, 2, and 3). We also collect the daily stock market factor, together with the risk-free rate, from Kenneth French's data library. We consider the period from January 1963 to December 2024⁷. Over this sample, the average number of trading days in a month is 21, while the number of listed stocks ranges from a minimum of 1956 to a maximum of 7465. Figure 3.3 reports the number of common stocks in the CRSP U.S. stock database over time. The resulting universe contains 15 606 trading days and 26 220 stocks.

⁷Data availability: for Nyse from December 31, 1925; for Nyse Market from July 2, 1962; for Nasdaq from December 14, 1972; for Arca from March 8, 2006; for the Fama-French factors from July 1, 1926.

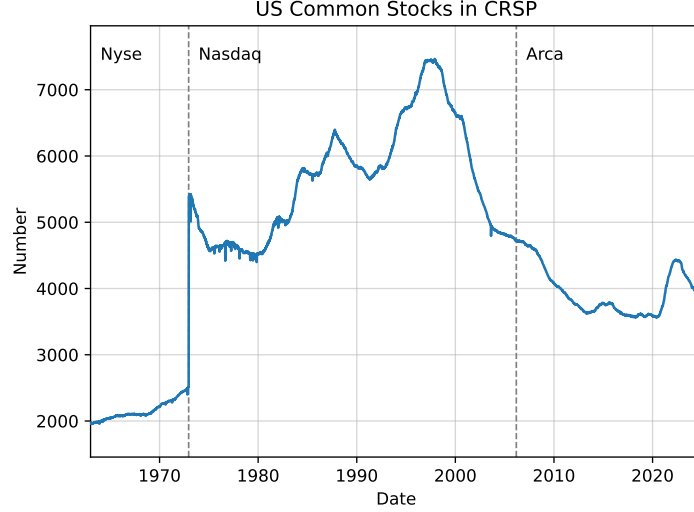


Figure 3.3: Number of common stocks in the CRSP U.S. stock database.

3.7.2 Simulations for Classification Errors

Since the true beta is not observable, classification errors cannot be measured directly from the data. Our simulation exercise provides a benchmark by treating the beta estimated from the historical sample at each rebalancing date as the latent “true” sorting variable, and then generating artificial return histories that are consistent with those estimated parameters. Re-estimating betas on the simulated samples yields alternative, noisy sorting variables that differ from the benchmark only because of estimation error. Comparing the portfolio assignments based on these simulated estimates with those based on the benchmark betas therefore gives a direct measure of contamination and information-loss errors. This approach is useful because it isolates the effect of beta estimation error while keeping the simulation close to the empirical environment observed in the data.

True (unobservable) sorting variable At each rebalancing date $\tau(t) \geq R$ (i.e. every 21 days), we regress the excess return of each stock over the market factor using data from the past R days, ignoring stocks with more than 5% of data points missing. We set R to be equal to a multiple E of the average number of trading days in a month, $R = 21E$ for $E \in \{12, 24, 60\}$, corresponding to an estimation period of one year, two years, and five years. Thus, we estimate the following regression for each stock and each rebalancing date:

$$r_{i,t}^e = \alpha_{i,\tau(t)} + \beta_{i,\tau(t)} r_{M,t} + \epsilon_{i,t}, \quad (3.11)$$

for $i = 1, \dots, N_{\tau(t)}$ and $t \in (\tau(t) - R, \tau(t)]$, where $r_{i,t}^e = r_{i,t} - r_{f,t}$. In the remainder of this section, we set $S_{i,\tau(t)} \equiv \widehat{\beta}_{i,\tau(t),R}$, i.e. $\widehat{\beta}_{i,\tau(t),R}$ is taken to be the true unobservable sorting variable for stock i at rebalancing date $\tau(t)$. This yields an unbalanced panel of market betas, where the time dimension is the rebalancing dates and the cross-section is the stock universe. From the same regression, we also compute the idiosyncratic volatility of each stock $\widehat{\sigma}_{i,\tau(t),R}$ as the sample standard deviation of the fitted residuals $\widehat{\epsilon}_{i,t,R}$. These values will be used to simulate stock returns in the next step.

Table C.1 shows some summary statistics for the market beta, estimated with $E = 12$, when stocks are sorted into $N = 10$ portfolios. The distribution is clearly asymmetric in the tails: P1 is strongly left-skewed, with mean below the median and a much longer lower tail, while P10 is strongly right-skewed, with mean above the median and a much longer upper tail. At the same time, dispersion is much larger in the two extreme portfolios: the standard deviation is 0.27 in P1 and 0.35 in P10, compared with values between 0.04 and 0.09 in the intermediate portfolios, and the range is also substantially wider in the tails. This highlights the importance of a trimming variable that accounts for the standard error of the sorting variable. Appendix C.2.1 reports the same table for other values of E and N .

Table 3.1: Summary statistics of the estimated betas across portfolios ($E = 12$, $N = 10$).

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Std. Dev.	0.27	0.06	0.05	0.04	0.04	0.04	0.05	0.06	0.09	0.35
Min	-2.32	0.13	0.34	0.50	0.64	0.78	0.92	1.09	1.30	1.61
Mean	-0.09	0.24	0.42	0.57	0.71	0.85	1.00	1.19	1.44	1.99
Median	-0.02	0.24	0.42	0.57	0.71	0.85	1.00	1.18	1.43	1.89
Max	0.13	0.34	0.50	0.64	0.78	0.92	1.09	1.30	1.61	4.12

Estimated (observable) sorting variable For each simulation $s = 1, \dots, 500$, at each rebalancing date $\tau(t) \geq R$, we simulate the excess returns of each stock over the past R days, using the historical values of the market factor, the estimated coefficients from regression (3.11), and drawing the innovations from a normal distribution with mean zero and standard deviation equal to the one computed from equation (3.11). Thus, we simulate excess returns using the following model for each stock and each rebalancing date⁸:

$$\widetilde{r}_{i,t,s}^e = \widehat{\alpha}_{i,\tau(t),R} + \widehat{\beta}_{i,\tau(t),R} r_{M,t} + \widetilde{\epsilon}_{i,t,s},$$

⁸We denote simulated values by a tilde and the simulation number is indexed by an s , e.g. $\widetilde{x}_s$.

for $i = 1, \dots, N_{\tau(t)}$ and $t \in (\tau(t) - R, \tau(t)]$, where $\tilde{\epsilon}_{i,t,s} \sim N(0, \hat{\sigma}_{i,\tau(t),R}^2)$. Then, for each stock and each rebalancing date we re-estimate regression (3.11), using the simulated excess returns $\tilde{r}_{i,t,s}^e$ in place of the historical ones $r_{i,t}^e$. In the remainder of this section, we set $\hat{S}_{i,\tau(t),R,s} \equiv \hat{\beta}_{i,\tau(t),R,s}$, i.e. $\hat{\beta}_{i,\tau(t),R,s}$ is taken to be the sorting variable estimated with measurement error for stock i at rebalancing date $\tau(t)$. This yields an unbalanced panel of market betas; a different one for each simulation.

Sorting stocks into untrimmed portfolios For each simulation s , at each rebalancing date $\tau(t) \geq R$, we sort the $N_{\tau(t)}$ stocks according to the values of $S_{i,\tau(t)}$ and $\hat{S}_{i,\tau(t),R,s}$ and we allocate them to $\mathcal{N} \in \{5, 10\}$ groups to obtain the untrimmed portfolios defined by the true unobservable sorting variable and the sorting variable estimated with measurement error, respectively. Thus, for the latter, we allocate the stocks according to the following conditions:

$$i \in \begin{cases} \text{Portfolio } j = 1: & \text{if } 1 = \mathbb{1} \left\{ \hat{S}_{i,\tau(t),R,s} \leq \hat{S}_{\tau(t),R,s}^{p(1)} \right\} \\ \text{Portfolio } j = 2, \dots, \mathcal{N} - 1: & \text{if } 1 = \mathbb{1} \left\{ \hat{S}_{\tau(t),R,s}^{p(j-1)} < \hat{S}_{i,\tau(t),R,s} \leq \hat{S}_{\tau(t),R,s}^{p(j)} \right\} \\ \text{Portfolio } j = \mathcal{N}: & \text{if } 1 = \mathbb{1} \left\{ \hat{S}_{\tau(t),R,s}^{p(\mathcal{N}-1)} < \hat{S}_{i,\tau(t),R,s} \right\} \end{cases}$$

and similarly for the former, where $S_{i,\tau(t)} \equiv \hat{\beta}_{i,\tau(t),R}$ and $S_{i,\tau(t)}^{p(j)} \equiv \hat{\beta}_{\tau(t),R}^{p(j)}$ are substituted for $\hat{S}_{i,\tau(t),R,s} \equiv \hat{\beta}_{i,\tau(t),R,s}$ and $\hat{S}_{\tau(t),R,s}^{p(j)} \equiv \hat{\beta}_{\tau(t),R,s}^{p(j)}$.

Sorting stocks into trimmed portfolios In order to define the trimmed portfolios, at each rebalancing date $\tau(t) \geq R$, we first compute an almost sure bound for the beta estimation error of each stock using the law of iterated logarithm:

$$\hat{c}_{i,\tau(t),R,s} = \sqrt{\frac{2 \ln(\ln R)}{R}} \text{Std}(\hat{S}_{i,\tau(t),R,s}),$$

for $i = 1, \dots, N_{\tau(t)}$. In particular, let $\hat{\epsilon}_{i,t,s}$ denote the fitted residuals from regression (3.11) estimated on the simulated excess returns $\tilde{r}_{i,t,s}^e$ over the window $u \in (\tau(t) - R, \tau(t)]$. Then we set:

$$\text{Std}(\hat{S}_{i,\tau(t),R,s}) = \frac{\hat{\sigma}_{\epsilon,i,\tau(t),R,s}}{\hat{\sigma}_{M,\tau(t),R}},$$

where:

$$\begin{aligned}\widehat{\sigma}_{\epsilon, i, \tau(t), R, s} &= \sqrt{\frac{1}{R-2} \sum_{u \in (\tau(t)-R, \tau(t)]} \widehat{\epsilon}_{i, u, s}^2}, \\ \widehat{\sigma}_{M, \tau(t), R} &= \sqrt{\frac{1}{R-1} \sum_{u \in (\tau(t)-R, \tau(t)]} (r_{M, u} - \bar{r}_{M, \tau(t), R})^2}, \\ \bar{r}_{M, \tau(t), R} &= \frac{1}{R} \sum_{u \in (\tau(t)-R, \tau(t)]} r_{M, u}.\end{aligned}$$

We proceed to compute the bounds of the trimmed portfolios. Recall that, for the 1-st (resp. $\mathcal{N}$ -th) portfolio, we only need to "adjust" the sorting variables by subtracting the respective almost-sure bounds from (resp. adding them to) the estimated sorting variables. Indeed, the lower bound (resp. upper bound) of the portfolio is unchanged. Instead, for portfolios $j = 2, \dots, \mathcal{N} - 1$, we need to "adjust" once adding and once subtracting the almost-sure bounds for the error. Thus, first we compute the adjusted sorting variables:

$$\text{Portfolio } j = 1: \quad \widehat{S}_{i, \tau(t), R, s}^{1, L} = \widehat{S}_{i, \tau(t), R, s} - \mathbb{1} \left\{ \widehat{S}_{i, \tau(t), R, s} \leq \widehat{S}_{\tau(t), R, s}^{p(1)} \right\} \widehat{c}_{i, \tau(t), R, s} \quad (3.12)$$

$$\text{Portfolio } j = 2, \dots, \mathcal{N} - 1: \quad \widehat{S}_{i, \tau(t), R, s}^{j, U} = \widehat{S}_{i, \tau(t), R, s} + \mathbb{1} \left\{ \widehat{S}_{\tau(t), R, s}^{p(j-1)} < \widehat{S}_{i, \tau(t), R, s} \leq \widehat{S}_{\tau(t), R, s}^{p(j)} \right\} \widehat{c}_{i, \tau(t), R, s}$$

$$\widehat{S}_{i, \tau(t), R, s}^{j, L} = \widehat{S}_{i, \tau(t), R, s} - \mathbb{1} \left\{ \widehat{S}_{\tau(t), R, s}^{p(j-1)} < \widehat{S}_{i, \tau(t), R, s} \leq \widehat{S}_{\tau(t), R, s}^{p(j)} \right\} \widehat{c}_{i, \tau(t), R, s}$$

$$\text{Portfolio } j = \mathcal{N}: \quad \widehat{S}_{i, \tau(t), R, s}^{\mathcal{N}, U} = \widehat{S}_{i, \tau(t), R, s} + \mathbb{1} \left\{ \widehat{S}_{\tau(t), R, s}^{p(\mathcal{N}-1)} < \widehat{S}_{i, \tau(t), R, s} \right\} \widehat{c}_{i, \tau(t), R, s} \quad (3.13)$$

Then, we compute the $jN_{\tau(t)}/\mathcal{N}$ -th order statistics $\widehat{S}_{\tau(t), R, s}^{p(j), U}$ and $\widehat{S}_{\tau(t), R, s}^{p(j), L}$ of $\widehat{S}_{i, \tau(t), R, s}^{j, U}$ and $\widehat{S}_{i, \tau(t), R, s}^{j, L}$ respectively. Finally, we allocate the stocks according to the following conditions:

$$i \in \begin{cases} \text{Portfolio } j = 1: & \text{if } 1 = \mathbb{1} \left\{ \widehat{S}_{i, \tau(t), R, s} \leq \widehat{S}_{\tau(t), R, s}^{p(1), L} \right\} \\ \text{Portfolio } j = 2, \dots, \mathcal{N} - 1: & \text{if } 1 = \mathbb{1} \left\{ \widehat{S}_{\tau(t), R, s}^{p(j-1), U} < \widehat{S}_{i, \tau(t), R, s} \leq \widehat{S}_{\tau(t), R, s}^{p(j), L} \right\} \\ \text{Portfolio } j = \mathcal{N}: & \text{if } 1 = \mathbb{1} \left\{ \widehat{S}_{\tau(t), R, s}^{p(\mathcal{N}-1), U} < \widehat{S}_{i, \tau(t), R, s} \right\}. \end{cases}$$

Figure 3.4 shows the average size of the trimmed and untrimmed portfolios across simulations and the number of stocks trimmed at each rebalancing date in the two extreme portfolios.

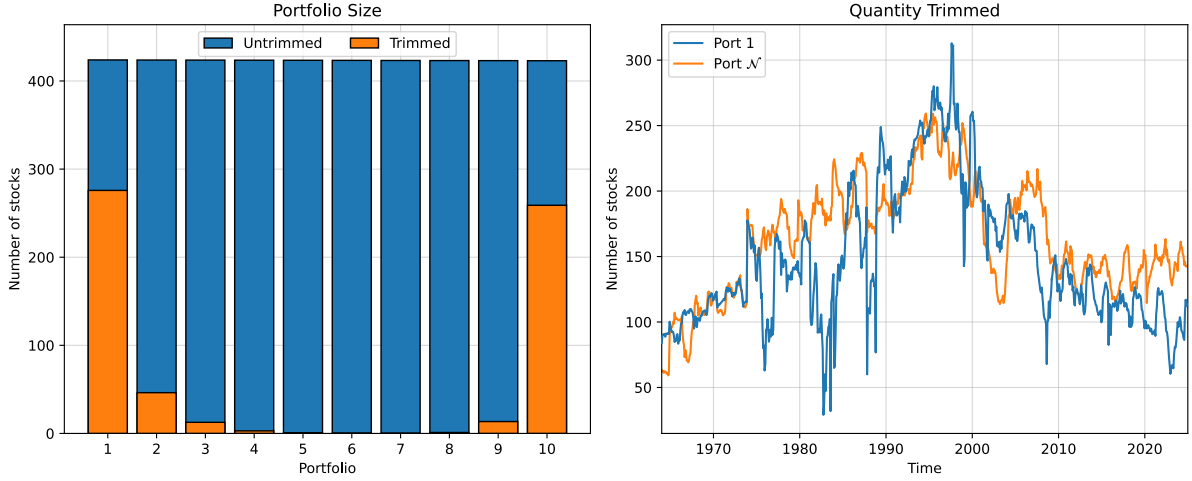


Figure 3.4: Left plot: average size of portfolios across simulations and time; Right plot: average quantity trimmed across simulations ($N = 10, E = 12$).

Classification errors At each rebalancing date $\tau(t) \geq R$, we measure the contamination error for the 1-st (resp. N -th) untrimmed portfolio determined by $\widehat{S}_{i,\tau(t),R,s}$ as the number of stocks in this portfolio that are included in each of the 2-nd to N -th (resp. 1-st to $(N-1)$ -th) untrimmed portfolios determined by $S_{i,\tau(t)}$, taken as a percentage of the size of the former. Thus, the contamination error in the 1-st (resp. N -th) portfolio due to portfolio $j = 2, \dots, N$ (resp. $j = 1, \dots, N-1$) is given by:

$$\begin{aligned} \text{Cont. in } 1 \text{ due to } j &= \frac{\sum_{i=1}^{N_{\tau(t)}} \mathbb{1} \left\{ S_{\tau(t)}^{p(j-1)} < S_{i,\tau(t)} \leq S_{\tau(t)}^{p(j)} \right\} \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R,s} \leq \widehat{S}_{\tau(t),R,s}^{p(1)} \right\}}{\sum_{i=1}^{N_{\tau(t)}} \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R,s} \leq \widehat{S}_{\tau(t),R,s}^{p(1)} \right\}} \\ \text{Cont. in } N \text{ due to } j &= \frac{\sum_{i=1}^{N_{\tau(t)}} \mathbb{1} \left\{ S_{\tau(t)}^{p(j-1)} < S_{i,\tau(t)} \leq S_{\tau(t)}^{p(j)} \right\} \mathbb{1} \left\{ \widehat{S}_{\tau(t),R,s}^{p(N-1)} < \widehat{S}_{i,\tau(t),R,s} \right\}}{\sum_{i=1}^{N_{\tau(t)}} \mathbb{1} \left\{ \widehat{S}_{\tau(t),R,s}^{p(N-1)} < \widehat{S}_{i,\tau(t),R,s} \right\}} \end{aligned}$$

Similarly, we measure the information loss error for the 1-st (resp. N -th) untrimmed portfolio determined by $S_{i,\tau(t)}$ as the number of stocks in this portfolio that are included in each of the 2-nd to N -th (resp. 1-st to $(N-1)$ -th) untrimmed portfolios determined by $\widehat{S}_{i,\tau(t),R,s}$, taken as a percentage of the size of the former. Thus, the information loss error in the 1-st (resp. N -th) portfolio due to portfolio $j = 2, \dots, N$ (resp. $j = 1, \dots, N-1$) is computed as:

$$\text{Info. loss in } 1 \text{ due to } j = \frac{\sum_{i=1}^{N_{\tau(t)}} \mathbb{1} \left\{ \widehat{S}_{\tau(t),R,s}^{p(j-1)} < \widehat{S}_{i,\tau(t),R,s} \leq \widehat{S}_{\tau(t),R,s}^{p(j)} \right\} \mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(1)} \right\}}{\sum_{i=1}^{N_{\tau(t)}} \mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(1)} \right\}}$$

$$\text{Inf. loss in } \mathcal{N} \text{ due to } j = \frac{\sum_{i=1}^{N_{\tau(t)}} \mathbb{1} \left\{ \widehat{S}_{\tau(t),R,s}^{p(j-1)} < \widehat{S}_{i,\tau(t),R,s} \leq \widehat{S}_{\tau(t),R,s}^{p(j)} \right\} \mathbb{1} \left\{ S_{\tau(t)}^{p(N-1)} < S_{i,\tau(t)} \right\}}{\sum_{i=1}^{N_{\tau(t)}} \mathbb{1} \left\{ S_{\tau(t)}^{p(N-1)} < S_{i,\tau(t)} \right\}}$$

Finally, we take the average of these values across simulations and time. A similar reasoning applies to the trimmed portfolios. Tables 3.2 and 3.3 show the contamination error and the information loss error without and with trimming, respectively. For example, consider the row "Contamination P1": columns P2-P10 show the percentage of stocks allocated to portfolio 1 that would actually belong to portfolios 2, ..., 10; column P1 is the total classification error, i.e. the sum of columns P2-P10. Two observations are in order. First, as expected, the two types of classification errors decrease as we move away from the extreme portfolios. Second, we can see how the trimming procedure correctly reduces the classification errors. Appendix C.2.2 reports the same table for other values of E and $\mathcal{N}$.

Table 3.2: Contamination and information loss errors without trimming (in %, $E = 12$, $\mathcal{N} = 10$).

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Contamination P1	32.92	19.19	6.59	3.09	1.69	0.99	0.61	0.37	0.23	0.16
Information loss P1	32.92	22.46	5.35	2.12	1.12	0.68	0.44	0.31	0.23	0.22
Contamination P10	0.22	0.16	0.21	0.30	0.45	0.74	1.46	3.79	17.51	24.83
Information loss P10	0.16	0.10	0.12	0.16	0.25	0.44	1.00	3.29	19.32	24.83
Portfolio size	424	424	424	424	423	423	423	423	423	423

Table 3.3: Contamination and information loss errors with trimming (in %, $E = 12$, $\mathcal{N} = 10$).

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Contamination P1	25.83	13.96	5.37	2.68	1.52	0.92	0.58	0.37	0.24	0.19
Information loss P1	25.83	2.93	0.17	0.00	0.00	0.00	0.00	0.00	0.01	0.23
Contamination P10	0.23	0.16	0.19	0.25	0.36	0.56	1.00	2.23	8.86	13.84
Information loss P10	0.20	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.52	13.84
Portfolio size	276	46	13	3	0	0	0	1	13	259
Trimmed quantity	148	378	411	421	423	423	423	422	410	164
Trimmed percentage	34.93	89.08	97.02	99.31	99.88	99.96	99.99	99.75	96.81	38.78

Image 3.5 shows the contamination error for different numbers of portfolios and estimation periods. The dashed line shows how the error is reduced when trimming is applied. Notice how a longer estimation period reduces the classification error across the board.

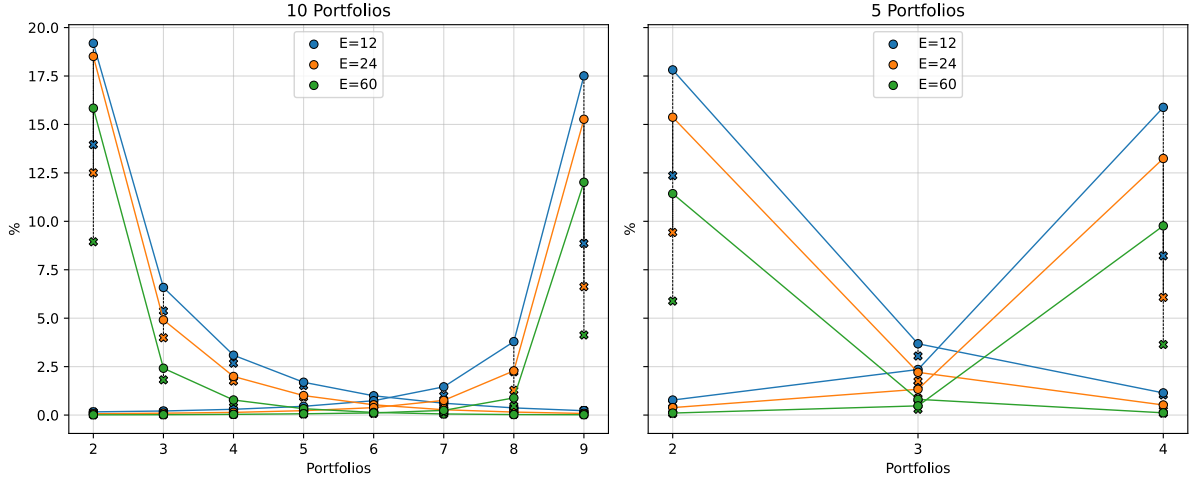


Figure 3.5: Contamination error with and without trimming for different numbers of portfolios and estimation periods.

3.7.3 Simulations for Hypothesis Testing

The purpose of this section is to study the finite-sample behavior of the hypothesis-testing procedure in the beta-sorting application. In the data, the true size and power of the test are not observable, because the data generating process and the true portfolio assignments are unknown. Our simulation exercise provides a benchmark by generating artificial return panels calibrated to the empirical distribution of estimated betas and then imposing either the null or the alternative hypothesis, under both weak and strong cross-sectional correlation of the sorting variable. We then apply the same sorting, trimming, and subsampling steps used in the empirical procedure to the simulated samples, and compare the resulting rejection frequencies across the untrimmed and trimmed statistics. This approach is useful because it isolates the effect of beta estimation error and the resulting portfolio contamination on the testing procedure, while keeping the simulation close to the empirical environment studied in the paper.

Simulated returns First, we calibrate the simulation using the historical sample. Throughout this section, we consider only the last 60 years of data (15 120 trading days). For each stock with at least 252 nonmissing excess-return observations in this period, we estimate a constant-coefficient market model using all available observations for that stock. Thus, we estimate:

$$r_{i,t}^e = \alpha_i + \beta_i r_{M,t} + \epsilon_{i,t}, \quad (3.14)$$

for $i = 1, \dots, N$ and all observed $t \in [\underline{T}_i, \bar{T}_i]$. In contrast with regression (3.11), regression (3.14) is estimated once over the full available history of each stock in the retained sample and therefore has time-invariant coefficients. We denote the resulting OLS estimates by $\hat{\alpha}_i$ and $\hat{\beta}_i$. From the same regression, we also compute the idiosyncratic volatility of each stock $\hat{\epsilon}_{i,t}$ as the sample standard deviation of the fitted residuals $\hat{\epsilon}_{i,t}$. This yields a cross section of historical estimates $\{(\hat{\alpha}_i, \hat{\beta}_i, \hat{\sigma}_i)\}_{i=1}^N$. Finally, to reduce the influence of extreme observations, we retain only the stocks whose $\hat{\alpha}_i$, $\hat{\beta}_i$, and $\hat{\sigma}_i$ all lie between their respective 1st and 99th empirical percentiles.

Figure 3.6 displays the empirical distribution of the full-sample beta estimates. The distribution spans a fairly wide range, from about -0.24 to 2.13 , and is mildly right-skewed, with a longer upper tail than lower tail. Most estimated betas are positive, although a small fraction are negative, indicating that the bulk of stocks comove positively with the market while a limited number exhibit negative market exposure.

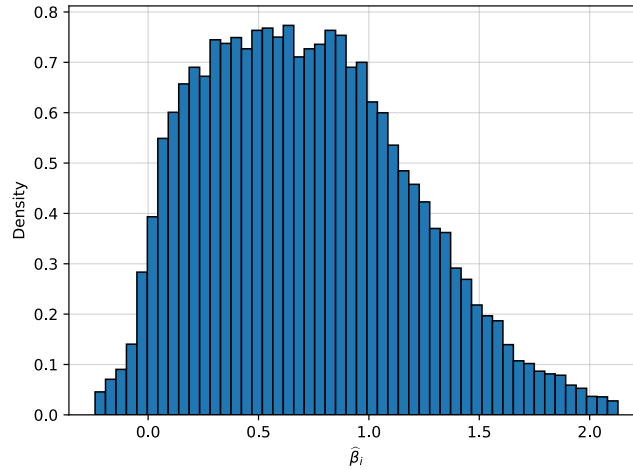


Figure 3.6: Empirical distribution of the sorting variable estimated on the whole sample.

Next, for each simulation $s = 1, \dots, 500$, we generate a balanced panel of daily excess returns for N stocks over T days according to:

$$r_{i,t,s}^e = \alpha_{i,t,s} + \beta_{i,t,s} r_{M,t,s} + \epsilon_{i,t,s},$$

for $i = 1, \dots, N$ and $t = 1, \dots, T$, where the simulated market factor is given by:

$$r_{M,t,s} = \mu_M + \sigma_M u_{t,s}, \quad u_{t,s} \stackrel{i.i.d.}{\sim} N(0, 1),$$

and μ_M and σ_M denote the empirical mean and standard deviation of the historical market factor. We take the baseline cross section of market betas to be evenly spaced over a compact interval:

$$\beta_{0,i} = \underline{\beta} + \frac{i-1}{N-1} (\bar{\beta} - \underline{\beta}), \quad i = 1, \dots, N.$$

We then partition the ordered grid $\{\beta_{0,i}\}_{i=1}^N$ into N equal-sized portfolios, denote the two extreme portfolios by $\mathcal{P}_1$ and $\mathcal{P}_N$, and define narrow boundary bands $\mathcal{B}_2 \subset \mathcal{P}_2$ and $\mathcal{B}_{N-1} \subset \mathcal{P}_{N-1}$ adjacent to the cutoffs between portfolios 1 and 2 and between portfolios $N-1$ and N , respectively. In the setting with weak cross-sectional correlation of the sorting variable, the true sorting variable is time invariant and we simply set:

$$\beta_{i,t,s} = \beta_{0,i}.$$

Instead, in the setting with strong cross-sectional correlation of the sorting variable, the true beta depends on a common state variable $W_{t,s}$ constructed from the simulated market factor as the standardized and clipped logarithm of the square root of the rolling $\mathcal{T}$ -day average of past squared market returns. Specifically, we set:

$$\beta_{i,t,s} = \beta_{0,i} + \beta_W W_{t,s}.$$

Because the same common shift $\beta_W W_{t,s}$ is added to all stocks, the cross-sectional ordering induced by $\beta_{0,i}$ is preserved over time. Moreover, the residual component is Gaussian and independent across i and t :

$$\varepsilon_{i,t,s} \sim N\left(0, \sigma_{\varepsilon,i}^2\right),$$

where $\sigma_{\varepsilon,i}$ equals a baseline scale outside the boundary bands and an inflated multiple of that scale inside $\mathcal{B}_2 \cup \mathcal{B}_{N-1}$. Finally, under H_0 we set:

$$\alpha_{i,t,s} = -\mu_M \beta_{i,t,s} + \begin{cases} -a_0 & \text{if } i \in \mathcal{B}_2, \\ +a_0 & \text{if } i \in \mathcal{B}_{N-1}, \\ 0 & \text{otherwise.} \end{cases}$$

Instead, under H_A we set:

$$\alpha_{i,t,s} = -\mu_M \beta_{i,t,s} + \begin{cases} -a_1 & \text{if } i \in \mathcal{P}_1, \\ +m_H a_1 & \text{if } i \in \mathcal{B}_2, \\ -m_H a_1 & \text{if } i \in \mathcal{B}_{N-1}, \\ +a_1 & \text{if } i \in \mathcal{P}_N, \\ 0 & \text{otherwise.} \end{cases}$$

Note that in the H_0 design, the additional alpha shifts are confined to narrow boundary bands in portfolios 2 and $N - 1$, so that the bottom and top portfolios continue to have the same expected return in the absence of sorting errors. In the H_A design, instead, the alpha shifts are negative in portfolio 1 and positive in portfolio N , with larger opposite-signed adjustments in the adjacent boundary bands, so that expected returns increase with beta while contamination from the neighboring portfolios attenuates the bottom-minus-top spread. The weak-correlation design keeps the true beta fixed over time, whereas the strong-correlation design makes it depend on the common state variable $W_{t,s}$, thereby inducing strong cross-sectional dependence in the sorting variable.

Subsampling scheme We set $R = 252$ days (12 months), which implies a ratio $\frac{T-R}{R} = 59$. To maintain the same speed of growth of $T - R$ relative to R in the subsampling scheme, we set $T_\delta = 5\,040$ days (20 years) and $R_\delta = 84$ days (7 months). We thus obtain 481 overlapping windows in the subsampling scheme. We now outline the steps to compute the test statistic $\widehat{Z}_{N,T,R,s}$, the subsampled counterpart $\widehat{Z}_{N,T_\delta,R_\delta,s}^{(k)}$, and recover the size and power of the test statistic.

First, we recompute the market beta of each stock at each rebalancing date using data for the past R days. In the remainder of this section, $\widehat{S}_{i,\tau(t),R,s} \equiv \widehat{\beta}_{i,\tau(t),s}$, i.e. $\widehat{\beta}_{i,\tau(t),s}$ is taken to be the sorting variable estimated with measurement error.

Next, we sort the stocks into trimmed portfolios at each rebalancing date, following the same steps outlined in the previous section in equations (3.12) - (3.13). Namely, at the last business day of each month, we sort the N stocks according to the values of $\widehat{S}_{i,\tau(t),R,s}$, we allocate them into N portfolios, and we compute the confidence interval around the true unobservable ordered statistics to trim the observations.

Finally, we compute the test statistic. More precisely, given the composition of the 1-st and the N -th trimmed portfolios at each rebalancing date, we compute the equally weighted average return of the stocks in these portfolios for each day until the next rebalancing date (i.e. we compute the average daily return of the portfolios over the month which starts the day after the rebalancing). Then we take the difference, thus obtaining the time series for the return of the 1-st trimmed portfolio minus the return of the N -th trimmed portfolio, where the two are dynamically updated. Lastly, we compute the test statistic as the average of this difference multiplied by the square root of the number of observations.

$$\widehat{Z}_{T,N,R,s} = \frac{N}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t,s} \left(\mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R,s} \leq \widehat{S}_{\tau(t),R,s}^{L,p(1)} \right\} - \mathbb{1} \left\{ \widehat{S}_{\tau(t),R,s}^{U,p(N-1)} < \widehat{S}_{i,\tau(t),R,s} \right\} \right),$$

where R in the square root and in the starting value of the summation should more accurately be the number of days until the first rebalancing date.

Next, we implement the subsampling procedure. We proceed as outlined above to recompute the market beta of the stocks, sort them into trimmed portfolios at each rebalancing date, and obtain a time series for the return of the 1-st trimmed portfolio minus the return of the N -th trimmed portfolio from which we compute the statistic. The only difference is that instead of considering the whole sample, we now consider overlapping windows of T_δ observations (shifting them by one month at the time), and that we now use an estimation period of R_δ days as opposed to R . For each subsample, we compute the test statistic $\widehat{Z}_{N,T_\delta,R_\delta}^{(k)}$ as the average of the return difference multiplied by the square root of the number of observations.

$$\widehat{Z}_{T_\delta,N,R_\delta,s}^{(k)} = \frac{N}{\sqrt{T_\delta}} \sum_{t=R_\delta+1}^{T_\delta} \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t,s} \left(\mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R_\delta,s} \leq \widehat{S}_{\tau(t),R_\delta,s}^{L,p(1)} \right\} - \mathbb{1} \left\{ \widehat{S}_{\tau(t),R_\delta,s}^{U,p(N-1)} < \widehat{S}_{i,\tau(t),R_\delta,s} \right\} \right),$$

Thus, we obtain an empirical distribution of K test statistics, of which we compute the 5% and 95% critical values, $c_{B,K,\alpha/2,s}$ and $c_{B,K,1-\alpha/2,s}$.

As a final step, we compute the percentage of times across simulations that $\widehat{Z}_{T,N,R}$ falls within and outside the critical values:

$$\frac{\sum_{s=1}^{500} \mathbb{1} \left\{ \widehat{c}_{B,K,\alpha/2,s} \leq \widehat{Z}_{T,N,R,s} \leq \widehat{c}_{B,K,1-\alpha/2,s} \right\}}{500}$$

$$\frac{\sum_{s=1}^{500} \mathbb{1} \left\{ \widehat{Z}_{T,N,R,s} \leq \widehat{c}_{B,K,\alpha/2,s} \right\} + \mathbb{1} \left\{ \widehat{c}_{B,K,1-\alpha/2,s} \leq \widehat{Z}_{T,N,R,s} \right\}}{500}$$

Table 3.4 shows the size and power of the test for different numbers of portfolios and under both the assumption of weak and strong cross-sectional correlation of the sorting variable. Under H_0 , contamination from portfolio 2 (resp. $N - 1$) lowers the return of portfolio 1 (resp. raises the return of portfolio N), thereby increasing the spread between the top and bottom portfolios relative to the uncontaminated case. As a result, the untrimmed statistic is too large in absolute value and falls outside the empirical confidence interval more often than the nominal 5% level, which explains its over-sizing. By contrast, trimming removes the stocks that are most likely to contaminate the extreme portfolios and brings the rejection frequency much closer to the nominal size. Under H_A , the same contamination works in the opposite direction for power: stocks from portfolio 2 that enter portfolio 1 raise the return of the bottom portfolio, while stocks from portfolio $N - 1$ that enter portfolio N lower the return of the top portfolio, so that the return spread is attenuated. Consequently, the untrimmed statistic is smaller than the trimmed one and rejects less often under the alternative, i.e. it is under-powered relative to the trimmed statistic. Even though the finite-sample results are not yet perfect, the table clearly shows that trimming, coupled with the subsampling procedure, moves the test in the right direction by correcting the size distortion under H_0 and improving power under H_A , in line with the theoretical prediction that the resulting procedure should have correct asymptotic size and unit power.

Table 3.4: Size and power of the untrimmed and trimmed test statistics (in %, $R=126$, $T_\delta=5040$, $R_\delta=42$).

	H_0		H_A	
	Untrimmed	Trimmed	Untrimmed	Trimmed
$N=10$, corr=weak	27.6	3.4	92.6	95.4
$N=10$, corr=strong	21.4	5.6	94.0	97.4
$N=5$, corr=weak	20.2	3.6	96.2	98.0
$N=5$, corr=strong	20.2	3.8	96.0	97.6

3.8 Conclusion

This paper studies how sorting errors affect tests for equality of expected returns across portfolios. When the sorting characteristic is observable, the relevant error comes from using

order statistics to approximate the unobservable population quantiles. We show that this error is asymptotically negligible only when the sorting variable is weakly cross-sectionally correlated; otherwise, it contributes to the asymptotic variance and has to be accounted for in inference. When the sorting characteristic is unobservable and must be estimated, measurement error creates contamination and information-loss errors in the resulting portfolios. The former occurs when a portfolio includes stocks that should belong to another portfolio, while the latter occurs when stocks that should belong to a given portfolio are excluded from it. These classification errors can distort the bottom-minus-top return spread and lead to incorrect inference.

We propose an adaptive trimming procedure that discards stocks close to the portfolio boundaries, where classification errors are most likely to occur. The trimming rule is designed to control contamination while requiring conditions under which the information loss induced by trimming is asymptotically negligible. Because the additional variance terms generated by estimated sorting variables are not generally available in closed form, we combine trimming with a subsampling procedure that preserves the relative growth rate of the testing and estimation windows. The theoretical results show that the resulting tests have correct asymptotic size and unit power. In simulations calibrated to beta-sorted portfolios, trimming substantially reduces contamination, and trimming coupled with subsampling moves rejection frequencies toward the nominal size under the null and toward unit power under the alternative.

Bibliography

Chapter 1: Disagreement in the Term Structure of Interest Rates

- P. Balduzzi, S. R. Das, and S. Foresi. The central tendency: A second factor in bond yields. *Review of Economics and Statistics*, 80(1):62–72, 1998.
- C. Banerghansa and M. W. McCracken. Forecast disagreement among fomc members. *Federal Reserve Bank of St. Louis Working Paper No*, 2009.
- M. Bongard, G. Galati, R. Moessner, and W. Nelson. Connecting the dots: market reactions to forecasts of policy rates and forward guidance provided by the fed. *International Journal of Finance & Economics*, 26(1):684–706, 2021.
- J. Y. Campbell. A defense of traditional hypotheses about the term structure of interest rates. *The Journal of Finance*, 41(1):183–193, 1986.
- J. Y. Campbell and R. J. Shiller. Yield spreads and interest rate movements: A bird’s eye view. *The Review of Economic Studies*, 58(3):495–514, 1991.
- J. Y. Campbell, R. J. Shiller, and L. M. Viceira. Understanding inflation-indexed bond markets. Technical report, National Bureau of Economic Research, 2009.
- L. Chen. *Stochastic mean and stochastic volatility: a three-factor model of the term structure of interest rates and its applications in derivatives pricing and risk management*. Blackwell publishers, 1996.
- J. C. Cox, J. E. Ingersoll Jr, and S. A. Ross. A re-examination of traditional hypotheses about the term structure of interest rates. *The Journal of Finance*, 36(4):769–799, 1981.
- J. M. Culbertson. The term structure of interest rates. *The quarterly journal of economics*, 71(4): 485–517, 1957.

- G.-A. Detmers. Forward guidance under disagreement: Evidence from the fed’s dot projections. Technical report, SFB 649 Discussion Paper, 2016.
- M. Ehrmann and M. Fratzscher. Dispersed communication by central bank committees and the predictability of monetary policy decisions. *Public Choice*, 157(1):223–244, 2013.
- E. F. Fama and R. R. Bliss. The information in long-maturity forward rates. *The American Economic Review*, pages 680–692, 1987.
- R. Greenwood, S. Hanson, and D. Vayanos. Forward guidance in the yield curve: Short rates versus bond supply. Technical report, National Bureau of Economic Research, 2015.
- R. Greenwood, S. G. Hanson, and G. Y. Liao. Asset price dynamics in partially segmented markets. *The Review of Financial Studies*, 31(9):3307–3343, 2018.
- R. Greenwood, S. Hanson, J. C. Stein, and A. Sunderam. A quantity-driven theory of term premia and exchange rates. *The Quarterly Journal of Economics*, 138(4):2327–2389, 2023a.
- R. Greenwood, S. Hanson, and D. Vayanos. Supply and demand and the term structure of interest rates. *Annual Review of Financial Economics*, 16, 2023b.
- R. S. Gürkaynak, B. Sack, and J. H. Wright. The us treasury yield curve: 1961 to the present. *Journal of monetary Economics*, 54(8):2291–2304, 2007.
- S. G. Hanson, D. O. Lucca, and J. H. Wright. Rate-amplifying demand and the excess sensitivity of long-term rates. *The Quarterly Journal of Economics*, 136(3):1719–1781, 2021.
- S. G. Hanson, A. Malkhozov, and G. Venter. Demand-and-supply imbalance risk and long-term swap spreads. *Journal of Financial Economics*, 154:103814, 2024.
- F. Hayashi. Computing equilibrium bond prices in the vayanos-vila model. *Research in Economics*, 72(2):181–195, 2018.
- Z. He, S. Nagel, and Z. Song. Treasury inconvenience yields during the covid-19 crisis. *Journal of Financial Economics*, 143(1):57–79, 2022.
- R. Kekre, M. Lenel, and F. Mainardi. Monetary policy, segmentation, and the term structure. Technical report, National Bureau of Economic Research, 2024.

- T. B. King. Expectation and duration at the effective lower bound. *Journal of Financial Economics*, 134(3):736–760, 2019.
- A. Krishnamurthy and A. Vissing-Jorgensen. The aggregate demand for treasury debt. *Journal of Political Economy*, 120(2):233–267, 2012.
- O. P. Machado, C. Carvalho, and T. Berriel. Lift-off uncertainty: What can we infer from the fomc’s summary of economic projections? In *2015 Meeting Papers*, number 903. Society for Economic Dynamics, 2015.
- F. Modigliani and R. Sutch. Innovations in interest rate policy. *The American economic review*, 56(1/2):178–197, 1966.
- W. Ray, M. Droste, and Y. Gorodnichenko. Unbundling quantitative easing: Taking a cue from treasury auctions. *Journal of Political Economy*, 132(9):3115–3172, 2024.
- A. Riboni and F. Ruge-Murcia. Dissent in monetary policy decisions. *Journal of Monetary Economics*, 66:137–154, 2014.
- A. Sangvinatsos. *Expectations Hypothesis*. Wiley Online Library, 2010. doi: <https://doi.org/10.1002/9780470061602.eqf03005>.
- D. Vayanos and J.-L. Vila. A preferred-habitat model of the term structure of interest rates. *Econometrica*, 89(1):77–112, 2021a.
- D. Vayanos and J.-L. Vila. Supplement to ‘a preferred-habitat model of the term structure of interest rates’. *Econometrica Supplemental Material*, 89(1), 2021b.

Chapter 2: Market Failure, the QE Trap and Fiscal Escape

- M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. U.S. Government Printing Office, Washington, D.C., tenth printing edition, 1972.
- T. Adrian and H. S. Shin. Procyclical leverage and value-at-risk. *Review of Financial Studies*, 27(2):373–403, 2014.

- J. Bai and P. Perron. Estimating and testing linear models with multiple structural changes. *Econometrica*, 66(1):47–78, 1998.
- R. Bansal and A. Yaron. Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance*, 59(4):1481–1509, 2004.
- S. Basak and D. Cuoco. An equilibrium model with restricted stock market participation. *Review of Financial Studies*, 11(2):309–341, 1998.
- A. N. Berger and R. A. Roman. Did saving wall street really save main street? the real effects of TARP on local economic conditions. *Journal of Financial and Quantitative Analysis*, 52(5): 1827–1867, 2017.
- B. S. Bernanke. Nonmonetary effects of the financial crisis in the propagation of the great depression. *American Economic Review*, 73(3):257–276, 1983.
- B. S. Bernanke. Central banking after the great recession: Lessons learned and challenges ahead. Discussion with Liaquat Ahamed, The Brookings Institution, Washington, D.C., 2014. January 16.
- S. Bigio. Endogenous liquidity and the business cycle. *American Economic Review*, 105(6): 1883–1927, 2015.
- A. N. Borodin and P. Salminen. *Handbook of Brownian Motion: Facts and Formulae*. Birkhäuser, Basel, second edition, 2002.
- M. K. Brunnermeier and L. H. Pedersen. Market liquidity and funding liquidity. *Review of Financial Studies*, 22(6):2201–2238, 2009.
- M. K. Brunnermeier and Y. Sannikov. A macroeconomic model with a financial sector. *American Economic Review*, 104(2):379–421, 2014.
- M. K. Brunnermeier, H. James, and J.-P. Landau. *The Euro and the Battle of Ideas*. Princeton University Press, Princeton, 2016.
- R. J. Caballero and E. Farhi. The safety trap. *Review of Economic Studies*, 85(1):223–274, 2018.
- J. Y. Campbell and J. H. Cochrane. By force of habit: A consumption-based explanation of aggregate stock market behavior. *Journal of Political Economy*, 107(2):205–251, 1999.

- Z. Chen and L. Epstein. Ambiguity, risk, and asset returns in continuous time. *Econometrica*, 70(4):1403–1443, 2002.
- R. Chetty, A. Guren, D. Manoli, and A. Weber. Are micro and macro labor supply elasticities consistent? a review of evidence on the intensive and extensive margins. *American Economic Review*, 101(3):471–475, 2011.
- P. Cochran, L. Petrasek, Z. Saravay, M. Tian, and E. Wu. Assessment of dealer capacity to intermediate in treasury and agency MBS markets. FEDS Notes, Board of Governors of the Federal Reserve System, 2024.
- W. Cui and V. Sterk. Quantitative easing with heterogeneous agents. *Journal of Monetary Economics*, 123:68–90, 2021.
- V. Cúrdia and M. Woodford. The central-bank balance sheet as an instrument of monetary policy. *Journal of Monetary Economics*, 58(1):54–79, 2011.
- M. Del Negro, G. Eggertsson, A. Ferrero, and N. Kiyotaki. The great escape? a quantitative evaluation of the Fed’s liquidity facilities. *American Economic Review*, 107(3):824–857, 2017.
- F. Delbaen and W. Schachermayer. Arbitrage possibilities in Bessel processes and their relations to local martingales. *Probability Theory and Related Fields*, 102(3):357–366, 1995.
- A. Dembo and O. Zeitouni. *Large Deviations Techniques and Applications*. Springer, Berlin, second edition, 2010.
- M. Demuth and J. A. van Casteren. *Stochastic Spectral Theory for Selfadjoint Feller Operators: A Functional Integration Approach*. Birkhäuser, Basel, 2000.
- S. Di Tella. Uncertainty shocks and balance sheet recessions. *Journal of Political Economy*, 125(6):2038–2081, 2017.
- D. Duffie and L. G. Epstein. Stochastic differential utility. *Econometrica*, 60(2):353–394, 1992.
- G. B. Eggertsson and P. Krugman. Debt, deleveraging, and the liquidity trap: A fisher-minsky-koo approach. *Quarterly Journal of Economics*, 127(3):1469–1513, 2012.
- H.-J. Engelbert and W. Schmidt. On one-dimensional stochastic differential equations with generalized drift. In *Stochastic Differential Systems*, volume 69 of *Lecture Notes in Control and Information Sciences*, pages 143–155. Springer, 1985.

- Federal Deposit Insurance Corporation. Federal deposit insurance corporation improvement act of 1991, 1991. Public Law 102–242.
- M. I. Freidlin and A. D. Wentzell. *Random Perturbations of Dynamical Systems*. Springer, Berlin, third edition, 2012.
- J. Gagnon, M. Raskin, J. Remache, and B. Sack. The financial market effects of the federal reserve’s large-scale asset purchases. *International Journal of Central Banking*, 7(1):3–43, 2011.
- M. Gertler and P. Karadi. A model of unconventional monetary policy. *Journal of Monetary Economics*, 58(1):17–34, 2011.
- M. Gertler and N. Kiyotaki. Financial intermediation and credit policy in business cycle analysis. In *Handbook of Monetary Economics*, volume 3A, pages 547–599. Elsevier, 2010.
- R. Greenwood, S. G. Hanson, and J. C. Stein. A comparative-advantage approach to government debt maturity. *Journal of Finance*, 70(4):1683–1722, 2015.
- V. Guerrieri and G. Lorenzoni. Credit crises, precautionary savings, and the liquidity trap. *Quarterly Journal of Economics*, 132(3):1427–1467, 2017.
- C. Gulati and A. L. Smith. The evolving role of the fed’s balance sheet: Effects and challenges. *Economic Review*, 107(4):59–84, 2022.
- Z. He and A. Krishnamurthy. A model of capital and crises. *Review of Economic Studies*, 79(2): 735–777, 2012.
- Z. He and A. Krishnamurthy. Intermediary asset pricing. *American Economic Review*, 103(2): 732–770, 2013.
- Z. He and A. Krishnamurthy. A macroeconomic framework for quantifying systemic risk. *American Economic Journal: Macroeconomics*, 11(4):1–37, 2019.
- Z. He, B. Kelly, and A. Manela. Intermediary asset pricing: New evidence from many asset classes. *Journal of Financial Economics*, 126(1):1–35, 2017.
- T. Homar and S. J. G. van Wijnbergen. Bank recapitalization and economic recovery after financial crises. *Journal of Financial Intermediation*, 32:16–28, 2017.

- J. Hugonnier. Rational asset pricing bubbles and portfolio constraints. *Journal of Economic Theory*, 147(6):2260–2302, 2012.
- I. Karatzas and S. E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer-Verlag, New York, second edition, 1991.
- S. Karlin and H. M. Taylor. *A Second Course in Stochastic Processes*. Academic Press, New York, 1981.
- R. C. Koo. *The Escape from Balance Sheet Recession and the QE Trap: A Hazardous Road for the World Economy*. John Wiley & Sons, Hoboken, NJ, 2014.
- A. Krishnamurthy and A. Vissing-Jørgensen. The effects of quantitative easing on interest rates: Channels and implications for policy. *Brookings Papers on Economic Activity*, (Fall):215–287, 2011.
- Y. Liu and J. C. Wu. Reconstructing the yield curve. *Journal of Financial Economics*, 142(3):1395–1425, 2021.
- A. Mijatović and M. Urusov. On the martingale property of certain local martingales. *Probability Theory and Related Fields*, 152(1):1–30, 2012.
- Organisation for Economic Co-operation and Development. *OECD Economic Outlook, Volume 2025 Issue 1: Tackling Uncertainty, Reviving Growth*. OECD Publishing, Paris, 2025.
- S. Pal and P. Protter. Analysis of continuous strict local martingales via h -transforms. *Stochastic Processes and their Applications*, 120(8):1424–1443, 2010.
- T. Philippon and P. Schnabl. Efficient recapitalization. *Journal of Finance*, 68(1):1–42, 2013.
- M. H. Protter and H. F. Weinberger. *Maximum Principles in Differential Equations*. Springer-Verlag, New York, 1984.
- M. Reed and B. Simon. *Methods of Modern Mathematical Physics IV: Analysis of Operators*. Academic Press, New York, 1978.
- R. Reis. Different types of central bank insolvency and the central role of seignorage. NBER Working Paper 21226, National Bureau of Economic Research, 2015.

- R. Reis. Funding quantitative easing to target inflation. In *Designing Resilient Monetary Policy Frameworks for the Future*. Federal Reserve Bank of Kansas City, Kansas City, 2016.
- R. Reis. Can the central bank alleviate fiscal burdens? In D. G. Mayes, P. L. Siklos, and J.-E. Sturm, editors, *The Oxford Handbook of the Economics of Central Banking*, pages 131–170. Oxford University Press, Oxford, 2019.
- D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer-Verlag, Berlin, third edition, 1999.
- M. Rognlie, A. Shleifer, and A. Simsek. Investment hangover and the great recession. *American Economic Journal: Macroeconomics*, 10(2):113–153, 2018.
- D. H. Silva. The risk channel of unconventional monetary policy. Working paper, University of Illinois at Urbana-Champaign, 2020.
- H. Stehfest. Numerical inversion of laplace transforms. *Communications of the ACM*, 13(1):47–49, 1970.
- D. Vayanos and J.-L. Vila. A preferred-habitat model of the term structure of interest rates. *Econometrica*, 89(1):77–112, 2021.
- P. Veronesi and L. Zingales. Paulson’s gift. *Journal of Financial Economics*, 97(3):339–368, 2010.
- A. Zettl. *Sturm–Liouville Theory*, volume 121 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2005.

Chapter 3: The Econometrics of Portfolio Sorts

- D. Amaya, P. Christoffersen, K. Jacobs, and A. Vasquez. Does realized skewness predict the cross-section of equity returns? *Journal of Financial Economics*, 118(1):135–167, 2015.
- A. Ang, R. J. Hodrick, Y. Xing, and X. Zhang. The cross-section of volatility and expected returns. *The Journal of Finance*, 61(1):259–299, 2006.
- F. Black, M. C. Jensen, and M. Scholes. The capital asset pricing model: Some empirical tests. In M. C. Jensen, editor, *Studies in the Theory of Capital Markets*, pages 79–121. Praeger, New York, 1972.

- M. D. Cattaneo, R. K. Crump, M. H. Farrell, and E. Schaumburg. Characteristic-sorted portfolios: Estimation and inference. *Review of Economics and Statistics*, 102(3):531–551, 2020.
- M. D. Cattaneo, R. K. Crump, and W. Wang. Beta-sorted portfolios. *arXiv preprint arXiv:2208.10974*, 2022.
- S. Cederburg and M. S. O’Doherty. Does it pay to bet against beta? on the conditional performance of the beta anomaly. *The Journal of finance*, 71(2):737–774, 2016.
- V. Corradi. Deciding between $i(0)$ and $i(1)$ via flil-based bounds. *Econometric Theory*, 15(5): 643–663, 1999.
- E. Eberlein. On strong invariance principles under dependence assumptions. *The Annals of Probability*, pages 260–270, 1986.
- E. F. Fama and K. R. French. The cross-section of expected stock returns. *the Journal of Finance*, 47(2):427–465, 1992.
- E. F. Fama and K. R. French. Common risk factors in the returns on stocks and bonds. *Journal of financial economics*, 33(1):3–56, 1993.
- E. F. Fama and K. R. French. The value premium and the capm. *The Journal of Finance*, 61(5): 2163–2185, 2006.
- E. F. Fama and J. D. MacBeth. Risk, return, and equilibrium: Empirical tests. *Journal of political economy*, 81(3):607–636, 1973.
- A. Frazzini and L. H. Pedersen. Betting against beta. *Journal of financial economics*, 111(1):1–25, 2014.
- S. Gonçalves. The moving blocks bootstrap for panel linear regression models with individual fixed effects. *Econometric Theory*, 27(5):1048–1082, 2011.
- C. R. Harvey and A. Siddique. Conditional skewness in asset pricing tests. *The Journal of finance*, 55(3):1263–1295, 2000.
- N. Jegadeesh and S. Titman. Returns to buying winners and selling losers: Implications for stock market efficiency. *The Journal of finance*, 48(1):65–91, 1993.

- B. Liang. Portfolio formation, measurement errors, and beta shifts: A random sampling approach. *Journal of Financial Research*, 23(3):261–284, 2000.
- J. Lintner. The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics*, 47(1):13–37, 1965.
- A. W. Lo and A. C. MacKinlay. Stock market prices do not follow random walks: Evidence from a simple specification test. *The review of financial studies*, 1(1):41–66, 1988.
- M. Lopez de Prado and V. Zoonekynd. Why has factor investing failed?: The role of specification errors. *The Role of Specification Errors (January 17, 2024)*, 2024.
- D. N. Politis, J. P. Romano, and M. Wolf. Subsampling for heteroskedastic time series. *Journal of Econometrics*, 81(2):281–317, 1997.
- R.-D. Reiss. *Approximate distributions of order statistics: with applications to nonparametric statistics*. Springer science & business media, 2012.
- E. Rio et al. *Asymptotic theory of weakly dependent random processes*, volume 80. Springer, 2017.
- J. Shanken. Intertemporal asset pricing: An empirical investigation. *Journal of Econometrics*, 45(1-2):99–120, 1990.
- W. F. Sharpe. Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19(3):425–442, 1964.

Appendix A

Disagreement in the Term Structure of Interest Rates

Nicolo Ceneda

A.1 Proofs

A.1.1 Proposition 1.1

A complication when solving the system is that the coefficients of $A_r(\tau)$ in each equation depend on integrals involving the functions $(A_r(\tau), A_\eta(\tau), C(\omega, \tau))$ themselves. To solve the system, I follow a three-step procedure.

(Step 1) I take the integrals as given, and solve the first two equations as a system of linear ODEs with constant coefficients. Thus, I take $\kappa_{rr}^* \equiv \kappa_r - \lambda_{rr}$ as given, and solve the first differential equation. The solution is:

$$A_r(\tau) = \frac{1 - e^{-\tau \kappa_{rr}^*}}{\kappa_{rr}^*}.$$

Let $g_1(\tau, x) \equiv \frac{1 - e^{-\tau x}}{x}$. It is straightforward to verify that the function $g_1(\tau, x) \geq 0$, and hence $A_r(\tau) \geq 0$. Similarly, I take $\kappa_{r\eta}^* \equiv \kappa_r + \lambda_{r\eta}$ as given, and solve the second differential equation. The solution is:

$$A_\eta(\tau) = \frac{(1 - e^{-\tau \kappa_\eta}) \kappa_{rr}^* \kappa_{r\eta}^* - (1 - e^{-\tau \kappa_{rr}^*}) \kappa_\eta \kappa_{r\eta}^*}{(\kappa_{rr}^* - \kappa_\eta) \kappa_{rr}^* \kappa_\eta}.$$

Note that $A_\eta(\tau)$ can be written as $\frac{g_1(\tau, \kappa_\eta) - g_1(\tau, \kappa_{rr}^*)}{\kappa_{rr}^* - \kappa_\eta} \kappa_{r\eta}^*$. Given that $\partial_x g_1(\tau, x) < 0$, the first term is non-negative and, if $\kappa_{r\eta}^* > 0$ (which I will prove below), then also $A_\eta(\tau) \geq 0$.

(Step 2) I require that $(\kappa_{rr}^*, \kappa_{r\eta}^*)$ solve $\kappa_{rr}^* = \kappa_r - \lambda_{rr}$ and $\kappa_{r\eta}^* = \kappa_r + \lambda_{r\eta}$ given the solutions $(A_r(\tau), A_\eta(\tau))$. Substituting $A_r(\tau)$ into $\kappa_{rr}^* \equiv \kappa_r - \lambda_{rr}$, yields the fixed-point condition in κ_{rr}^* :

$$\underbrace{\kappa_{rr}^*}_{LHS} = \underbrace{\kappa_r + a\sigma_r^2 \int_0^\infty \beta(\tau) \left(\frac{1 - e^{-\tau\kappa_{rr}^*}}{\kappa_{rr}^*} \right)^2 d\tau}_{RHS}.$$

Note that $LHS|_{\kappa_{rr}^*=0} = 0$, $\partial_{\kappa_{rr}^*} LHS > 0$ and $\lim_{\kappa_{rr}^* \rightarrow \infty} LHS = \infty$; $RHS|_{\kappa_{rr}^*=0} > \kappa_r$, $\partial_{\kappa_{rr}^*} RHS < 0$ and $\lim_{\kappa_{rr}^* \rightarrow \infty} RHS = \kappa_r$. Therefore, the equation has a unique solution $\kappa_{rr}^* > \kappa_r$. Similarly, substituting $A_r(\tau)$ and $A_\eta(\tau)$ into $\kappa_{r\eta}^* = \kappa_r + \lambda_{r\eta}$:

$$\kappa_{r\eta}^* = \frac{\kappa_r}{1 + a\sigma_r^2 \int_0^\infty \beta(\tau) \left(\frac{(1 - e^{-\tau\kappa_\eta})\kappa_{rr}^* - (1 - e^{-\tau\kappa_{rr}^*})\kappa_\eta}{(\kappa_{rr}^* - \kappa_\eta)\kappa_{rr}^*\kappa_\eta} \right) \left(\frac{1 - e^{-\tau\kappa_{rr}^*}}{\kappa_{rr}^*} \right) d\tau}.$$

Note that the second term in the integral can be written as $\frac{g_1(\tau, \kappa_\eta) - g_1(\tau, \kappa_{rr}^*)}{\kappa_{rr}^* - \kappa_\eta}$. Therefore, the denominator is positive and greater than one, and hence $\kappa_{r\eta}^* < \kappa_r$.

(Step 3) Given $A_r(\tau)$ and $A_\eta(\tau)$, I solve the system of ODEs for $C(\omega, \tau)$ numerically.

A.1.2 Proposition 1.2

Multiplying both sides of the target short-rate dynamics by the integrating factor $e^{t\kappa_\eta}$ and integrating from t to $t + \tau$:

$$\begin{aligned} e^{t\kappa_\eta} d\eta_t &= e^{t\kappa_\eta} \kappa_\eta (r_t^*(\omega) - \eta_t) dt + e^{t\kappa_\eta} \sigma_\eta dz_{\eta,t} \\ \int_t^{t+\tau} d(e^{u\kappa_\eta} \eta_u) &= r_t^*(\omega) \int_t^{t+\tau} e^{u\kappa_\eta} \kappa_\eta du + \sigma_\eta \int_t^{t+\tau} e^{u\kappa_\eta} dz_{\eta,u} \\ \eta_{t+\tau} &= e^{-\tau\kappa_\eta} \eta_t + (1 - e^{-\tau\kappa_\eta}) r_t^*(\omega) + \sigma_\eta \int_t^{t+\tau} e^{-(t+\tau-u)\kappa_\eta} dz_{\eta,u} \end{aligned}$$

Multiplying both sides of the short-rate dynamics by the integrating factor $e^{t\kappa_r}$, integrating from t to $t + \tau$, substituting in η_u and taking the expectation:

$$\begin{aligned} e^{t\kappa_r} dr_t &= e^{t\kappa_r} \kappa_r (\eta_t - r_t) dt + e^{t\kappa_r} \sigma_r dz_{r,t} \\ \int_t^{t+\tau} d(e^{u\kappa_r} r_u) &= \kappa_r \int_t^{t+\tau} e^{u\kappa_r} \eta_u du + \sigma_r \int_t^{t+\tau} e^{u\kappa_r} dz_{r,u} \\ \mathbb{E}_t(r_{t+\tau}) &= e^{-\tau\kappa_r} r_t + \frac{(e^{-\tau\kappa_\eta} - e^{-\tau\kappa_r}) \kappa_r}{\kappa_r - \kappa_\eta} \eta_t + \frac{(1 - e^{-\tau\kappa_\eta}) \kappa_r - (1 - e^{-\tau\kappa_r}) \kappa_\eta}{\kappa_r - \kappa_\eta} r_t^*(\omega). \end{aligned}$$

It follows that the effects of shocks to the current and target short-rate on the expected future short-rate are, respectively:

$$\partial_{r_t} \mathbb{E}_t(r_{t+\tau}) = e^{-\tau \kappa_r},$$

$$\partial_{\eta_t} \mathbb{E}_t(r_{t+\tau}) = \frac{(e^{-\tau \kappa_\eta} - e^{-\tau \kappa_r}) \kappa_r}{\kappa_r - \kappa_\eta}.$$

Moreover, from the definition of instantaneous forward rate for maturity τ :

$$f_t(\tau) = \partial_\tau A_r(\tau) r_t + \partial_\tau A_\eta(\tau) \eta_t + \partial_\tau C(\omega, \tau).$$

It follows that the effects of shocks to the current and target short-rate on the instantaneous forward rate are, respectively:

$$\begin{aligned} \partial_{r_t} f_t(\tau) &= \partial_\tau A_r(\tau) \\ &= e^{-\tau \kappa_{rr}^*}, \end{aligned}$$

$$\begin{aligned} \partial_{\eta_t} f_t(\tau) &= \partial_\tau A_\eta(\tau) \\ &= \frac{(e^{-\tau \kappa_\eta} - e^{-\tau \kappa_{rr}^*}) \kappa_{r\eta}^*}{\kappa_{rr}^* - \kappa_\eta}. \end{aligned}$$

First, note that $\partial_{r_t} \mathbb{E}_t(r_{t+\tau}) > \partial_{r_t} f_t(\tau)$ because $\kappa_r < \kappa_{rr}^*$. Second, let $g_1(\kappa_1, \kappa_2) = \frac{(e^{-\tau \kappa_2} - e^{-\tau \kappa_1})}{\kappa_1 - \kappa_2}$ and note that $\partial_{\eta_t} \mathbb{E}_t(r_{t+\tau}) = g_1(\kappa_r, \kappa_\eta) \kappa_r$ and $\partial_{\eta_t} f_t(\tau) = g_1(\kappa_{rr}^*, \kappa_\eta) \kappa_{r\eta}^*$. Moreover, note that $\partial_{\kappa_1} g_1(\kappa_1, \kappa_2) = \frac{e^{-\tau \kappa_1} (1 + \tau(\kappa_1 - \kappa_2)) - e^{-\tau \kappa_2}}{(\kappa_1 - \kappa_2)^2} < 0$ because $e^{-\tau \kappa_1} (1 + \tau(\kappa_1 - \kappa_2)) - e^{-\tau \kappa_2} < 0$ (one can rewrite the inequality as $1 + \tau(\kappa_1 - \kappa_2) < e^{\tau(\kappa_1 - \kappa_2)}$, which always holds). Therefore, $g_1(\kappa_r, \kappa_\eta) > g_1(\kappa_{rr}^*, \kappa_\eta)$ because $\kappa_r < \kappa_{rr}^*$ and hence $\partial_{\eta_t} \mathbb{E}_t(r_{t+\tau}) > \partial_{\eta_t} f_t(\tau)$ because $\kappa_r > \kappa_{r\eta}^*$.

A.1.3 Proposition 1.3

From the definition of continuously compounded yield to maturity of the τ -year bond:

$$y_t(\tau) = \frac{A_r(\tau)}{\tau} r_t + \frac{A_\eta(\tau)}{\tau} \eta_t + \frac{C(\omega, \tau)}{\tau}.$$

First, the effect of a shock to the current short-rate on the bond yield is:

$$\partial_{r_t} y_t(\tau) = \frac{A_r(\tau)}{\tau} \geq 0.$$

where $\lim_{\tau \rightarrow 0} \partial_{r_t} y_t(\tau) = 1$ and $\lim_{\tau \rightarrow \infty} \partial_{r_t} y_t(\tau) = 0$. Moreover, differentiating:

$$\partial_{\tau} \partial_{r_t} y_t(\tau) = \frac{e^{-\tau \kappa_{rr}^*} \tau \kappa_{rr}^* - (1 - e^{-\tau \kappa_{rr}^*})}{\tau^2 \kappa_{rr}^*},$$

Let $g_1(\tau, \kappa) \equiv e^{-\tau \kappa} \tau \kappa - (1 - e^{-\tau \kappa})$ and note that $\partial_{\tau} \partial_{r_t} y_t(\tau) = \frac{g_1(\tau, \kappa_{rr}^*)}{\tau^2 \kappa_{rr}^*}$. Moreover, note that $g_1(0, \kappa) = 0$ and $\partial_{\tau} g_1(\tau, \kappa) = -e^{-\tau \kappa} \tau \kappa^2 < 0$. Therefore, $g_1(\tau, \kappa) < 0$ and hence $\partial_{\tau} \partial_{r_t} y_t(\tau) < 0$.

Second, the effect of a shock to the target short-rate on the bond yield is:

$$\partial_{\eta_t} y_t(\tau) = \frac{A_{\eta}(\tau)}{\tau} \geq 0.$$

where $\lim_{\tau \rightarrow 0} \partial_{\eta_t} y_t(\tau) = 0$ and $\lim_{\tau \rightarrow \infty} \partial_{\eta_t} y_t(\tau) = 0$. Moreover, differentiating:

$$\partial_{\tau} \partial_{\eta_t} y_t(\tau) = \frac{\kappa_{r\eta}^*}{(\kappa_{rr}^* - \kappa_{\eta}) \tau^2} \left(\frac{e^{-\tau \kappa_{\eta}} \tau \kappa_{\eta} - (1 - e^{-\tau \kappa_{\eta}})}{\kappa_{\eta}} - \frac{e^{-\tau \kappa_{rr}^*} \tau \kappa_{rr}^* - (1 - e^{-\tau \kappa_{rr}^*})}{\kappa_{rr}^*} \right).$$

Let $g_2(\tau, \kappa) \equiv \frac{e^{-\tau \kappa} \tau \kappa - (1 - e^{-\tau \kappa})}{\kappa}$ and $g_3(\tau) = \frac{g_2(\tau, \kappa_{\eta}) - g_2(\tau, \kappa_{rr}^*)}{\kappa_{rr}^* - \kappa_{\eta}}$ and note that $\partial_{\tau} \partial_{\eta_t} y_t(\tau) = \frac{\kappa_{r\eta}^*}{\tau^2} g_3(\tau)$. Moreover, note that $g_3(0) = 0$, $\lim_{\tau \rightarrow \infty} g_3(\tau) = -\frac{1}{\kappa_{\eta} \kappa_{rr}^*} < 0$, and $\partial_{\tau} g_3(\tau) = \frac{e^{-\tau \kappa_{rr}^*} \tau \kappa_{rr}^* - e^{-\tau \kappa_{\eta}} \tau \kappa_{\eta}}{\kappa_{rr}^* - \kappa_{\eta}} > 0$ for $\tau \in \left(0, \frac{\ln(\kappa_{\eta}/\kappa_{rr}^*)}{\kappa_{\eta} - \kappa_{rr}^*}\right)$. Therefore, $g_3(\tau)$ and hence $\partial_{\tau} \partial_{\eta_t} y_t(\tau)$ start positive and eventually turn negative.

A.1.4 Proposition 1.4

Setting $\sigma_{\eta} = 0$ in the proof of Proposition 1.7, the b_{FB} coefficient reduces to:

$$b_{FB} = \frac{\partial_{\tau} A_r(\tau) + A_r(\tau) \kappa_r - 1}{\partial_{\tau} A_r(\tau) - 1}.$$

From the ODE for $A_r(\tau)$ and the solution for κ_{rr}^* :

$$\partial_{\tau} A_r(\tau) + A_r(\tau) \kappa_r - 1 = A_r(\tau) \lambda_{rr}$$

$$\partial_{\tau} A_r(\tau) + A_r(\tau) \kappa_r - 1 = A_r(\tau) (\kappa_r - \kappa_{rr}^*),$$

$$\partial_\tau A_r(\tau) - 1 = A_r(\tau) (\lambda_{rr} - \kappa_r)$$

$$\partial_\tau A_r(\tau) - 1 = -A_r(\tau) \kappa_{rr}^*,$$

Therefore:

$$\begin{aligned} b_{FB} &= \frac{A_r(\tau) (\kappa_{rr}^* - \kappa_r)}{A_r(\tau) \kappa_{rr}^*} \\ &= \frac{\kappa_{rr}^* - \kappa_r}{\kappa_{rr}^*}. \end{aligned}$$

Since $\kappa_{rr}^* > \kappa_r$ if $a > 0$ and $\beta(\tau) > 0$ in a positive-measure subset of $(0, \infty)$, then $b_{FB} > 0$. Moreover, since κ_{rr}^* is increasing in a , σ_r^2 , and $\beta(\tau)$, it must be that b_{FB} increases in a , σ_r^2 , and $\beta(\tau)$.

Similarly, the b_{CS} coefficient reduces to:

$$b_{CS} = 1 - \frac{\partial_\tau A_r(\tau) + A_r(\tau) \kappa_r - 1}{\frac{A_r(\tau)}{\tau} - 1}.$$

From the ODE for $A_r(\tau)$ and the solution for κ_{rr}^* :

$$\begin{aligned} b_{CS} &= 1 - \frac{A_r(\tau) (\kappa_{rr}^* - \kappa_r)}{1 - \frac{A_r(\tau)}{\tau}} \\ &= 1 - \frac{A_r(\tau) \tau (\kappa_{rr}^* - \kappa_r)}{\tau - A_r(\tau)}. \end{aligned}$$

Since $\kappa_{rr}^* > \kappa_r$ if $a > 0$ and $\beta(\tau) > 0$ in a positive-measure subset of $(0, \infty)$, then $b_{CS} < 1$. Moreover, let $g_1(\tau) = \frac{A_r(\tau) \tau}{\tau - A_r(\tau)}$ and note that $\partial_\tau g_1(\tau) = \frac{1}{(\tau - A_r(\tau))^2} \left(e^{-\tau \kappa_{rr}^*} - \left(\frac{1 - e^{-\tau \kappa_{rr}^*}}{\tau \kappa_{rr}^*} \right)^2 \right) < 0$ because $e^{-x} < \left(\frac{1 - e^{-x}}{x} \right)^2$. Given that $\kappa_{rr}^* - \kappa_r > 0$ under these conditions, b_{CS} increases in τ .

A.1.5 Proposition 1.5

A complication when solving the system is that the coefficients of $(A_r(\tau), A_\eta(\tau))$ in each equation depend on integrals involving the functions $(A_r(\tau), A_\eta(\tau), C(\omega, \tau))$ themselves. To solve the system, I follow a two-step procedure.

(Step 1) I take the integrals as given, and solve the system as a system of linear ODEs with constant coefficients. Thus, I consider the first differential equation, and I take the integrals in Ψ as given. I diagonalize the matrix Ψ to transform the coupled system of linear ODEs into a system of two decoupled linear ODEs, making them easier to solve. Let v_j for $j = 1, 2$ denote

an eigenvalue of Ψ . The elements of the associated eigenvector are given by the left-eigenvalue formula:

$$\begin{bmatrix} \lambda_{j,1} & \lambda_{j,2} \end{bmatrix} \begin{bmatrix} \Psi_{11} & \Psi_{12} \\ \Psi_{21} & \Psi_{22} \end{bmatrix} = v_j \begin{bmatrix} \lambda_{j,1} & \lambda_{j,2} \end{bmatrix}$$

$$\begin{cases} \lambda_{j,1}\Psi_{11} + \lambda_{j,2}\Psi_{21} = v_j\lambda_{j,1} \\ \lambda_{j,1}\Psi_{12} + \lambda_{j,2}\Psi_{22} = v_j\lambda_{j,2}. \end{cases}$$

The ratio of the elements of the eigenvector is:

$$\lambda_j = \frac{\lambda_{j,2}}{\lambda_{j,1}} = \frac{v_j - \Psi_{11}}{\Psi_{21}}.$$

It follows that Ψ can be diagonalised as follows:

$$\Psi = P^{-1}NP,$$

where:

$$N = \begin{bmatrix} v_1 & 0 \\ 0 & v_2 \end{bmatrix}, \quad P = \begin{bmatrix} 1 & \lambda_1 \\ 1 & \lambda_2 \end{bmatrix}, \quad P^{-1} = \begin{bmatrix} -\frac{\lambda_2}{\lambda_1 - \lambda_2} & \frac{\lambda_1}{\lambda_1 - \lambda_2} \\ \frac{1}{\lambda_1 - \lambda_2} & -\frac{1}{\lambda_1 - \lambda_2} \end{bmatrix}.$$

Substituting into the ODE:

$$\partial_\tau A(\tau) + P^{-1}NPA(\tau) = \iota$$

Multiplying the ODE from the left by P and defining $\tilde{A}(\tau) \equiv PA(\tau)$, I solve the ODE by means of an integrating factor:

$$\begin{aligned} \partial_\tau \tilde{A}(\tau) + N\tilde{A}(\tau) &= P\iota \\ e^{\tau N} \partial_\tau \tilde{A}(\tau) + e^{\tau N} N\tilde{A}(\tau) &= e^{\tau N} P\iota \\ \partial_\tau (e^{\tau N} \tilde{A}(\tau)) &= e^{\tau N} P\iota \\ \int_0^\tau d(e^{uN} \tilde{A}(u)) &= \left(\int_0^\tau e^{uN} du \right) P\iota \\ A(\tau) &= P^{-1} \text{Diag} \left[\frac{1 - e^{-\tau v_1}}{v_1}, \frac{1 - e^{-\tau v_2}}{v_2} \right] P\iota. \end{aligned}$$

Note that:

$$\text{Diag} \left[\frac{1 - e^{-\tau\nu_1}}{\nu_1}, \frac{1 - e^{-\tau\nu_2}}{\nu_2} \right] = \text{Diag} \left[0, \frac{1 - e^{-\tau\nu_2}}{\nu_2} - \frac{1 - e^{-\tau\nu_1}}{\nu_1} \right] + \frac{1 - e^{-\tau\nu_1}}{\nu_1} \mathbf{I}_2,$$

where $\mathbf{I}_2$ denotes the 2×2 identity matrix. It follows that:

$$\begin{bmatrix} A_r(\tau) \\ A_\eta(\tau) \end{bmatrix} = \begin{bmatrix} \phi_r \left(\frac{1 - e^{-\tau\nu_2}}{\nu_2} - \frac{1 - e^{-\tau\nu_1}}{\nu_1} \right) + \frac{1 - e^{-\tau\nu_1}}{\nu_1} \\ \phi_\eta \left(\frac{1 - e^{-\tau\nu_2}}{\nu_2} - \frac{1 - e^{-\tau\nu_1}}{\nu_1} \right) \end{bmatrix},$$

where:

$$\phi_r = \frac{\lambda_1}{\lambda_1 - \lambda_2},$$

$$\phi_\eta = -\frac{1}{\lambda_1 - \lambda_2}.$$

Similarly, I take the integral ψ as given and solve the second differential equation by simple integration:

$$C(\tau) = \left(\int_0^\tau A(u)^\top du \right) \psi - \frac{1}{2} \int_0^\tau A(u)^\top \Sigma \Sigma^\top A(u) du$$

(Step 2) I compute the integrals in Ψ given the solutions $(A_r(\tau), A_\eta(\tau))$ and I require that they are consistent with the value of Ψ used to compute $(A_r(\tau), A_\eta(\tau))$. There are 4 integrals in the first differential equation. These integrals involve the 2 functions $(A_r(\tau), A_\eta(\tau))$, and determine the elements of the 2×2 matrix Ψ . In turn, the eigenvalues and eigenvectors of Ψ determine the solutions for $(A_r(\tau), A_\eta(\tau))$, and those solutions determine the values of the integrals. This amounts to solving a non-linear system of 4 equations in the 4 integrals. Given the solutions $A_r(\tau)$ and $A_\eta(\tau)$, the elements of ψ can be derived from a linear system of 2 equations to find the solution $C(\tau)$.

A.1.6 Lemma 1.1

First, multiplying the ODE system by $\beta(\tau)$ and integrating:

$$\int_0^\infty \beta(\tau) \partial_\tau A(\tau) d\tau + \Psi \int_0^\infty \beta(\tau) A(\tau) d\tau - \int_0^\infty \beta(\tau) d\tau = 0$$

$$\int_0^\infty \beta(\tau) \partial_\tau \mathbf{A}(\tau) d\tau + \mathbf{\Psi} \boldsymbol{\ell} - \boldsymbol{\iota} \int_0^\infty \beta(\tau) d\tau = 0.$$

Integrating by parts, and assuming $\lim_{\tau \rightarrow \infty} \beta(\tau) \mathbf{A}(\tau) = 0$, a property which is required for the matrix $\mathbf{\Psi}$ to be finite (and that holds for the solution provided in Theorem 1.1, as shown in its proof):

$$\begin{aligned} \int_0^\infty \beta(\tau) \partial_\tau \mathbf{A}(\tau) d\tau &= [\beta(\tau) \mathbf{A}(\tau)]_0^\infty - \int_0^\infty \partial_\tau \beta(\tau) \mathbf{A}(\tau) d\tau \\ &= \delta_\beta \boldsymbol{\ell}. \end{aligned}$$

It follows that:

$$\begin{aligned} \delta_\beta \boldsymbol{\ell} + \mathbf{\Psi} \boldsymbol{\ell} - \boldsymbol{\iota} \int_0^\infty \beta(\tau) d\tau &= 0 \\ (\delta_\beta \mathbf{I}_2 + \mathbf{\Psi}) \boldsymbol{\ell} &= \frac{\beta}{\delta_\beta} \boldsymbol{\iota}. \end{aligned}$$

This vector equation yields the two scalar equations:

$$\left(\delta_\beta + \kappa_r + a \sigma_r^2 \ell_{r,r} \right) \ell_r + a \sigma_\eta^2 \ell_{r,\eta} \ell_\eta = \frac{\beta}{\delta_\beta},$$

$$\left(\delta_\beta + \kappa_\eta + a \sigma_\eta^2 \ell_{\eta,\eta} \right) \ell_\eta - \left(\kappa_r - a \sigma_r^2 \ell_{r,\eta} \right) \ell_r = 0.$$

Second, multiplying the ODE system from the right by $\beta(\tau) \mathbf{A}(\tau)^\top$, adding to the resulting 2×2 matrix its transpose, and integrating:

$$\begin{aligned} &\int_0^\infty \beta(\tau) \left(\partial_\tau \mathbf{A}(\tau) \mathbf{A}(\tau)^\top + \mathbf{A}(\tau) \partial_\tau \mathbf{A}(\tau)^\top \right) d\tau + \\ &+ \int_0^\infty \beta(\tau) \left(\mathbf{\Psi} \mathbf{A}(\tau) \mathbf{A}(\tau)^\top + \mathbf{A}(\tau) \mathbf{A}(\tau)^\top \mathbf{\Psi}^\top \right) d\tau - \int_0^\infty \beta(\tau) \left(\boldsymbol{\iota} \mathbf{A}(\tau)^\top + \mathbf{A}(\tau) \boldsymbol{\iota}^\top \right) d\tau = 0 \\ &\int_0^\infty \beta(\tau) \left(\partial_\tau \mathbf{A}(\tau) \mathbf{A}(\tau)^\top + \mathbf{A}(\tau) \partial_\tau \mathbf{A}(\tau)^\top \right) d\tau + \mathbf{\Psi} \mathbf{L} + \mathbf{L} \mathbf{\Psi}^\top - \boldsymbol{\iota} \boldsymbol{\ell}^\top - \boldsymbol{\ell} \boldsymbol{\iota}^\top = 0. \end{aligned}$$

Integrating by parts, and assuming $\lim_{\tau \rightarrow \infty} \beta(\tau) \mathbf{A}(\tau) = 0$

$$\begin{aligned} \int_0^\infty \beta(\tau) \left(\partial_\tau \mathbf{A}(\tau) \mathbf{A}(\tau)^\top + \mathbf{A}(\tau) \partial_\tau \mathbf{A}(\tau)^\top \right) d\tau &= [\beta(\tau) \mathbf{A}(\tau) \mathbf{A}(\tau)^\top]_0^\infty - \int_0^\infty \partial_\tau \beta(\tau) \mathbf{A}(\tau) \mathbf{A}(\tau)^\top d\tau \\ &= \delta_\beta \mathbf{L}. \end{aligned}$$

It follows that:

$$\begin{aligned}\delta_\beta \mathbf{L} + \mathbf{\Psi} \mathbf{L} + \mathbf{L} \mathbf{\Psi}^\top - \mathbf{l} \mathbf{l}^\top - \mathbf{\ell} \mathbf{\ell}^\top &= 0 \\ (\delta_\beta \mathbf{I}_2 + \mathbf{\Psi}) \mathbf{L} + \mathbf{L} \mathbf{\Psi}^\top &= \mathbf{l} \mathbf{l}^\top + \mathbf{\ell} \mathbf{\ell}^\top.\end{aligned}$$

This matrix equation yields the three scalar equations:

$$\begin{aligned}\left(\frac{\delta_\beta}{2} + \kappa_r + a\sigma_r^2 \ell_{r,r}\right) \ell_{r,r} + a\sigma_\eta^2 \ell_{r,\eta}^2 &= \ell_r, \\ \left(\delta_\beta + \kappa_r + \kappa_\eta + a\sigma_r^2 \ell_{r,r} + a\sigma_\eta^2 \ell_{\eta,\eta}\right) \ell_{r,\eta} - \left(\kappa_r - a\sigma_r^2 \ell_{r,\eta}\right) \ell_{r,r} + a\sigma_\eta^2 \ell_{r,\eta} \ell_{\eta,\eta} &= \ell_\eta, \\ \left(\frac{\delta_\beta}{2} + \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta}\right) \ell_{\eta,\eta} - \left(\kappa_r - a\sigma_r^2 \ell_{r,\eta}\right) \ell_{r,\eta} &= 0.\end{aligned}$$

The total number of equations is 5, the same as the number of unknown Laplace transforms: the first (vector) equation yields 2 scalar equations, and the third (matrix) equation yields 3 scalar equations because the matrices are symmetric.

A.1.7 Lemma 1.2

I define:

$$n(\tau) \equiv \int_0^\tau \mathbf{A}(u)^\top \mathbf{\Sigma} \mathbf{\Sigma}^\top \mathbf{A}(u) du, \quad \mathbf{n}_0(\tau) \equiv \begin{bmatrix} n_{0,r}(\tau) & n_{0,\eta}(\tau) \end{bmatrix} \equiv \int_0^\tau \mathbf{A}(u)^\top du,$$

$$\mathbf{n}_1 \equiv \begin{bmatrix} n_{1,r} \\ n_{1,\eta} \end{bmatrix} \equiv \int_0^\infty \alpha_0(\tau) \mathbf{A}(\tau) d\tau, \quad \mathbf{n}_2 \equiv \begin{bmatrix} n_{2,r} \\ n_{2,\eta} \end{bmatrix} \equiv \frac{1}{2} \int_0^\infty \beta(\tau) n(\tau) \mathbf{A}(\tau) d\tau,$$

$$\mathbf{N} \equiv \begin{bmatrix} n_{r,r} & n_{r,\eta} \\ n_{\eta,r} & n_{\eta,\eta} \end{bmatrix} \equiv \int_0^\infty \beta(\tau) \mathbf{A}(\tau) \mathbf{n}_0(\tau) d\tau,$$

where $n(\tau)$ is a scalar, the vector $\mathbf{n}_0(\tau)$ has dimension 1×2 , the vector $\mathbf{n}_1$ has dimension 2×1 , the vector $\mathbf{n}_2$ has dimension 2×1 , and the matrix $\mathbf{N}$ has dimension 2×2 . Then the function

$C(\tau)$ can be written as:

$$C(\tau) = n_0(\tau)\psi - \frac{1}{2}n(\tau).$$

Substituting $C(\tau)$ into ψ :

$$\psi = r^*K\mathbf{1} + a\Sigma\Sigma^\top \int_0^\infty \left(\alpha_0(\tau) - \beta(\tau) \left(n_0(\tau)\psi - \frac{1}{2}n(\tau) \right) \right) A(\tau) d\tau$$

$$\psi = r^*K\mathbf{1} + a\Sigma\Sigma^\top (n_1 - N\psi + n_2)$$

$$\psi = (\mathcal{I}_2 + a\Sigma\Sigma^\top N)^{-1} (r^*K\mathbf{1} + a\Sigma\Sigma^\top (n_1 + n_2)) ,$$

where:

$$(\mathcal{I}_2 + a\Sigma\Sigma^\top N)^{-1} = \frac{1}{d} \begin{bmatrix} 1 + a\sigma_\eta^2 n_{\eta,\eta} & -a\sigma_r^2 n_{r,\eta} \\ -a\sigma_\eta^2 n_{\eta,r} & 1 + a\sigma_r^2 n_{r,r} \end{bmatrix} ,$$

$$d \equiv (1 + a\sigma_r^2 n_{r,r}) (1 + a\sigma_\eta^2 n_{\eta,\eta}) - a^2 \sigma_r^2 \sigma_\eta^2 n_{r,\eta} n_{\eta,r} .$$

It follows that:

$$\psi \equiv \begin{bmatrix} \psi_r \\ \psi_\eta \end{bmatrix} = \frac{1}{d} \begin{bmatrix} a\sigma_r^2 (1 + a\sigma_\eta^2 n_{\eta,\eta}) (n_{1,r} + n_{2,r}) - a\sigma_r^2 n_{r,\eta} (r^* \kappa_\eta + a\sigma_\eta^2 (n_{1,\eta} + n_{2,\eta})) \\ -a^2 \sigma_r^2 \sigma_\eta^2 n_{\eta,r} (n_{1,r} + n_{2,r}) + (1 + a\sigma_r^2 n_{r,r}) (r^* \kappa_\eta + a\sigma_\eta^2 (n_{1,\eta} + n_{2,\eta})) \end{bmatrix} .$$

Substituting ψ back into $C(\tau)$:

$$C(\tau) = n_{0,r}(\tau)\psi_r + n_{0,\eta}(\tau)\psi_\eta - \frac{1}{2}n(\tau)$$

A.1.8 Theorem 1.1

Substituting ℓ_r and ℓ_η into the system, I obtain a system of 3 equations in the 3 unknowns $(\ell_{r,r}, \ell_{\eta,\eta}, \ell_{r,\eta})$:

$$0 = R_{r,r}(\ell_{r,r}, \ell_{\eta,\eta}, \ell_{r,\eta}) \equiv \left(\frac{\delta_\beta}{2} + \kappa_r + a\sigma_r^2 \ell_{r,r} \right) \ell_{r,r} + a\sigma_\eta^2 \ell_{r,\eta}^2 - \frac{N_r}{D} ,$$

$$0 = R_{r,\eta}(\ell_{r,r}, \ell_{\eta,\eta}, \ell_{r,\eta}) \equiv \left(\delta_\beta + \kappa_r + \kappa_\eta + a\sigma_r^2 \ell_{r,r} + a\sigma_\eta^2 \ell_{\eta,\eta} \right) \ell_{r,\eta} - \left(\kappa_r - a\sigma_r^2 \ell_{r,\eta} \right) \ell_{r,r} \\ + a\sigma_\eta^2 \ell_{r,\eta} \ell_{\eta,\eta} - \frac{N_\eta}{D},$$

$$0 = R_{\eta,\eta}(\ell_{\eta,\eta}, \ell_{r,\eta}) \equiv \left(\frac{\delta_\beta}{2} + \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta} \right) \ell_{\eta,\eta} - \left(\kappa_r - a\sigma_r^2 \ell_{r,\eta} \right) \ell_{r,\eta},$$

where:

$$N_r \equiv \frac{\beta}{\delta_\beta} \left(\delta_\beta + \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta} \right),$$

$$N_\eta \equiv \frac{\beta}{\delta_\beta} \left(\kappa_r - a\sigma_r^2 \ell_{r,\eta} \right),$$

$$D \equiv \left(\delta_\beta + \kappa_r + a\sigma_r^2 \ell_{r,r} \right) \left(\delta_\beta + \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta} \right) + \left(\kappa_r - a\sigma_r^2 \ell_{r,\eta} \right) a\sigma_\eta^2 \ell_{r,\eta}.$$

To show that the system has a solution, I proceed in two steps. First, I take $0 < \ell_{r,\eta} < \frac{\kappa_r}{a\sigma_r^2}$ as given, and construct $\ell_{r,r} > 0$ and $\ell_{\eta,\eta} > 0$ uniquely from the first and third equations of the system. Second, I treat $(\ell_{r,r}, \ell_{\eta,\eta})$ as implicit functions of $\ell_{r,\eta}$ and show that the second equation of the system has a solution $0 < \ell_{r,\eta} < \frac{\kappa_r}{a\sigma_r^2}$.

(Step 1) First, taking $0 < \ell_{r,\eta} < \frac{\kappa_r}{a\sigma_r^2}$ as given, I construct $\ell_{\eta,\eta} > 0$ from the third equation:

$$R_{\eta,\eta}(\ell_{\eta,\eta}; \ell_{r,\eta}) = a\sigma_\eta^2 \ell_{\eta,\eta}^2 + \left(\frac{\delta_\beta}{2} + \kappa_\eta \right) \ell_{\eta,\eta} - \left(\kappa_r - a\sigma_r^2 \ell_{r,\eta} \right) \ell_{r,\eta} = 0.$$

Since $\ell_{\eta,\eta} = \int_0^\infty \beta(\tau) A_\eta^2(\tau) d\tau$, we require $\ell_{\eta,\eta} \geq 0$. We need to show that this equation has a unique positive solution for $\ell_{\eta,\eta}$. To check for uniqueness, I examine the derivative of $R_{\eta,\eta}$ with respect to $\ell_{\eta,\eta}$:

$$\partial_{\ell_{\eta,\eta}} R_{\eta,\eta} = \frac{\delta_\beta}{2} + \kappa_\eta + 2a\sigma_\eta^2 \ell_{\eta,\eta}.$$

Note that $\partial_{\ell_{\eta,\eta}} R_{\eta,\eta} > 0$ and $R_{\eta,\eta}$ is strictly increasing in $\ell_{\eta,\eta}$ for $\ell_{\eta,\eta} \geq 0$. Therefore, if a solution

$\ell_{\eta,\eta}$ exists, it is unique. To check for existence, I examine the limits:

$$R_{\eta,\eta}(0; \ell_{r,\eta}) = - \left(\kappa_r - a\sigma_r^2 \ell_{r,\eta} \right) \ell_{r,\eta},$$

$$\lim_{\ell_{\eta,\eta} \rightarrow \infty} R_{\eta,\eta}(\ell_{\eta,\eta}; \ell_{r,\eta}) = \infty.$$

Hence, a unique positive solution $\ell_{\eta,\eta}$ exists if and only if $R_{\eta,\eta}(0; \ell_{r,\eta}) < 0$, which always holds for $0 < \ell_{r,\eta} < \frac{\kappa_r}{a\sigma_r^2}$. Therefore, given $0 < \ell_{r,\eta} < \frac{\kappa_r}{a\sigma_r^2}$, the third equation yields a unique solution $\ell_{\eta,\eta}(\ell_{r,\eta}) > 0$.

Second, taking $0 < \ell_{r,\eta} < \frac{\kappa_r}{a\sigma_r^2}$ as given, and treating $\ell_{\eta,\eta}(\ell_{r,\eta}) > 0$ as an implicit function of $\ell_{r,\eta}$, I construct $\ell_{r,r} > 0$ from the first equation:

$$R_{r,r}(\ell_{r,r}; \ell_{\eta,\eta}(\ell_{r,\eta}), \ell_{r,\eta}) = \left(\frac{\delta_\beta}{2} + \kappa_r + a\sigma_r^2 \ell_{r,r} \right) \ell_{r,r} + a\sigma_\eta^2 \ell_{r,\eta}^2 - \frac{N_r(\ell_{\eta,\eta}(\ell_{r,\eta}))}{D(\ell_{r,r}, \ell_{\eta,\eta}(\ell_{r,\eta}), \ell_{r,\eta})} = 0.$$

Since $\ell_{r,r} = \int_0^\infty \beta(\tau) A_r^2(\tau) d\tau$, we require $\ell_{r,r} \geq 0$. We need to show that this equation has a unique positive solution for $\ell_{r,r}$. To check for uniqueness, I examine the derivative of $R_{r,r}$ with respect to $\ell_{r,r}$:

$$\partial_{\ell_{r,r}} R_{r,r} = \frac{\delta_\beta}{2} + \kappa_r + 2a\sigma_r^2 \ell_{r,r} + \frac{N_r(\ell_{\eta,\eta}(\ell_{r,\eta}))}{D^2(\ell_{r,r}, \ell_{\eta,\eta}(\ell_{r,\eta}), \ell_{r,\eta})} \partial_{\ell_{r,r}} D.$$

Note that $\frac{\delta_\beta}{2} + \kappa_r + 2a\sigma_r^2 \ell_{r,r} > 0$ for $\ell_{r,r} \geq 0$. Moreover, $N_r = \frac{\beta}{\delta_\beta} (\delta_\beta + \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta}) > 0$ and $\partial_{\ell_{r,r}} D = a\sigma_r^2 (\delta_\beta + \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta}) > 0$ for $\ell_{\eta,\eta} > 0$, so $\frac{N_r}{D^2} \partial_{\ell_{r,r}} D > 0$. Finally, for $\ell_{r,r} \geq 0$, given that $\ell_{\eta,\eta} > 0$ as established, and given that $0 < \ell_{r,\eta} < \frac{\kappa_r}{a\sigma_r^2}$ holds by assumption, it holds that $D > 0$, so $\frac{N_r}{D^2} \partial_{\ell_{r,r}} D$ is well defined. Thus, $\partial_{\ell_{r,r}} R_{r,r} > 0$ and $R_{r,r}$ is strictly increasing in $\ell_{r,r}$ for $\ell_{r,r} \geq 0$. Therefore, if a solution $\ell_{r,r}$ exists, it is unique. To check for existence, I examine the limits:

$$R_{r,r}(0; \ell_{\eta,\eta}(\ell_{r,\eta}), \ell_{r,\eta}) = a\sigma_\eta^2 \ell_{r,\eta}^2 - \frac{N_r(\ell_{\eta,\eta}(\ell_{r,\eta}))}{D(0, \ell_{\eta,\eta}(\ell_{r,\eta}), \ell_{r,\eta})},$$

$$\lim_{\ell_{r,r} \rightarrow \infty} R_{r,r}(\ell_{r,r}; \ell_{\eta,\eta}(\ell_{r,\eta}), \ell_{r,\eta}) = \infty.$$

Hence, a unique positive solution $\ell_{r,r}$ exists if and only if $R_{r,r}(0; \ell_{\eta,\eta}(\ell_{r,\eta}), \ell_{r,\eta}) < 0$. I assume that $(\ell_{\eta,\eta}, \ell_{r,\eta})$ are such that the condition $a\sigma_\eta^2 \ell_{r,\eta}^2 < \frac{N_r(\ell_{\eta,\eta}(\ell_{r,\eta}))}{D(0, \ell_{\eta,\eta}(\ell_{r,\eta}), \ell_{r,\eta})}$ holds, and return to this issue in

Step 2. Therefore, given $0 < \ell_{r,\eta} < \frac{\kappa_r}{a\sigma_r^2}$, the first equation yields a unique solution $\ell_{r,r}(\ell_{r,\eta}) > 0$.

(Step 2) The constraint required in Step 1 for the existence of $\ell_{r,r}(\ell_{r,\eta}) > 0$ is:

$$\begin{aligned} R_{r,r}(0; \ell_{\eta,\eta}(\ell_{r,\eta}), \ell_{r,\eta}) &= a\sigma_\eta^2 \ell_{r,\eta}^2 - \frac{N_r(\ell_{\eta,\eta}(\ell_{r,\eta}))}{D(0, \ell_{\eta,\eta}(\ell_{r,\eta}), \ell_{r,\eta})} \\ &= a\sigma_\eta^2 \ell_{r,\eta}^2 - \frac{\frac{\beta}{\delta_\beta} \left(\delta_\beta + \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta}(\ell_{r,\eta}) \right)}{(\delta_\beta + \kappa_r) \left(\delta_\beta + \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta}(\ell_{r,\eta}) \right) + (\kappa_r - a\sigma_r^2 \ell_{r,\eta}) a\sigma_\eta^2 \ell_{r,\eta}} < 0 \end{aligned}$$

This positivity constraint must hold, at least for $\ell_{r,\eta}$ close to zero, which is the lower bound of the domain $0 < \ell_{r,\eta} < \frac{\kappa_r}{a\sigma_r^2}$. To verify this, I compute $\lim_{\ell_{r,\eta} \rightarrow 0^+} R_{r,r}(0; \ell_{\eta,\eta}(\ell_{r,\eta}), \ell_{r,\eta})$. First, taking the limit of both sides of the third equation as $\ell_{r,\eta} \rightarrow 0^+$:

$$0 = \left(\frac{\delta_\beta}{2} + \kappa_\eta + a\sigma_\eta^2 \underline{\ell}_{\eta,\eta} \right) \underline{\ell}_{\eta,\eta},$$

where $\underline{\ell}_{\eta,\eta} \equiv \lim_{\ell_{r,\eta} \rightarrow 0^+} \ell_{\eta,\eta}$. The only non-negative solution is $\underline{\ell}_{\eta,\eta} = 0$. Given that from Step 1 $\ell_{\eta,\eta} > 0$ for $0 < \ell_{r,\eta} < \frac{\kappa_r}{a\sigma_r^2}$, it must be that $\lim_{\ell_{r,\eta} \rightarrow 0^+} \ell_{\eta,\eta} = 0^+$. Second, taking the limit of the first equation as $\ell_{r,\eta} \rightarrow 0^+$:

$$0 = \left(\frac{\delta_\beta}{2} + \kappa_r + a\sigma_r^2 \underline{\ell}_{r,r} \right) \underline{\ell}_{r,r} - \frac{\frac{\beta}{\delta_\beta}}{\delta_\beta + \kappa_r + a\sigma_r^2 \underline{\ell}_{r,r}},$$

where $\underline{\ell}_{r,r} \equiv \lim_{\ell_{r,\eta} \rightarrow 0^+} \ell_{r,r}$. This equation has a unique positive solution $\underline{\ell}_{r,r}$ because the function $g(x) = \left(\frac{\delta_\beta}{2} + \kappa_r + a\sigma_r^2 x \right) x - \frac{\beta/\delta_\beta}{\delta_\beta + \kappa_r + a\sigma_r^2 x}$ is strictly increasing in $x > 0$, $g(0) < 0$ and $\lim_{x \rightarrow \infty} g(x) = \infty$. It follows that:

$$\lim_{\ell_{r,\eta} \rightarrow 0^+} R_{r,r}(0; \ell_{\eta,\eta}(\ell_{r,\eta}), \ell_{r,\eta}) = -\frac{\frac{\beta}{\delta_\beta}}{\delta_\beta + \kappa_r} < 0.$$

Thus, the condition required in Step 1 for the existence of a unique $\ell_{r,r}(\ell_{r,\eta}) > 0$ holds for $\ell_{r,\eta}$ sufficiently close to zero. Finally, I analyse the sign of $\mathcal{R}(\ell_{r,\eta}) \equiv R_{r,\eta}(\ell_{r,r}(\ell_{r,\eta}), \ell_{\eta,\eta}(\ell_{r,\eta}), \ell_{r,\eta})$ as $\ell_{r,\eta} \rightarrow 0^+$:

$$\lim_{\ell_{r,\eta} \rightarrow 0^+} \mathcal{R}(\ell_{r,\eta}) = -\kappa_r \underline{\ell}_{r,r} - \frac{\frac{\beta}{\delta_\beta} \kappa_r}{\left(\delta_\beta + \kappa_r + a\sigma_r^2 \underline{\ell}_{r,r} \right) (\delta_\beta + \kappa_\eta)} < 0.$$

Next, I consider the upper bound of the domain of $\ell_{r,\eta}$. Taking the limit of both sides of the

third equation as $\ell_{r,\eta} \rightarrow \kappa_r/(a\sigma_r^2)^-$:

$$0 = \left(\frac{\delta_\beta}{2} + \kappa_\eta + a\sigma_\eta^2 \bar{\ell}_{\eta,\eta} \right) \bar{\ell}_{\eta,\eta},$$

where $\bar{\ell}_{\eta,\eta} \equiv \lim_{\ell_{r,\eta} \rightarrow \kappa_r/(a\sigma_r^2)^-} \ell_{\eta,\eta}$. The only non-negative solution is $\bar{\ell}_{\eta,\eta} = 0$. Given that from Step 1 $\ell_{\eta,\eta} > 0$ for $0 < \ell_{r,\eta} < \frac{\kappa_r}{a\sigma_r^2}$, it must be that $\lim_{\ell_{r,\eta} \rightarrow \kappa_r/(a\sigma_r^2)^-} \ell_{\eta,\eta} = 0^+$. Second, taking the limit of both sides of the first equation as $\ell_{r,\eta} \rightarrow \kappa_r/(a\sigma_r^2)^-$:

$$0 = \left(\frac{\delta_\beta}{2} + \kappa_r + a\sigma_r^2 \bar{\ell}_{r,r} \right) \bar{\ell}_{r,r} + \frac{\kappa_r^2 \sigma_\eta^2}{a\sigma_r^4} - \frac{\frac{\beta}{\delta_\beta}}{\delta_\beta + \kappa_r + a\sigma_r^2 \bar{\ell}_{r,r}},$$

where $\bar{\ell}_{r,r} \equiv \lim_{\ell_{r,\eta} \rightarrow \kappa_r/(a\sigma_r^2)^-} \ell_{r,r}$. The function $g(x) = \left(\frac{\delta_\beta}{2} + \kappa_r + a\sigma_r^2 x \right) x + \frac{\kappa_r^2 \sigma_\eta^2}{a\sigma_r^4} - \frac{\beta/\delta_\beta}{\delta_\beta + \kappa_r + a\sigma_r^2 x}$ is strictly increasing in $x > 0$ and $\lim_{x \rightarrow \infty} g(x) = \infty$. Therefore the equation has a unique positive solution $\bar{\ell}_{r,r}$ if $g(0) < 0$, which is the case if $\frac{\kappa_r^2 \sigma_\eta^2}{a\sigma_r^4} < \frac{\beta/\delta_\beta}{\delta_\beta + \kappa_r}$, or equivalently $\kappa_r^2 \sigma_\eta^2 \delta_\beta (\delta_\beta + \kappa_r) < a\sigma_r^4 \beta$. It follows that:

$$\lim_{\ell_{r,\eta} \rightarrow \kappa_r/(a\sigma_r^2)^-} R_{r,r}(0; \ell_{\eta,\eta}(\ell_{r,\eta}), \ell_{r,\eta}) = \frac{\kappa_r^2 \sigma_\eta^2}{a\sigma_r^4} - \frac{\frac{\beta}{\delta_\beta}}{\delta_\beta + \kappa_r} < 0,$$

provided that $\kappa_r^2 \sigma_\eta^2 \delta_\beta (\delta_\beta + \kappa_r) < a\sigma_r^4 \beta$. Finally, I analyse the sign of $\mathcal{R}(\ell_{r,\eta}) \equiv R_{r,\eta}(\ell_{r,r}(\ell_{r,\eta}), \ell_{\eta,\eta}(\ell_{r,\eta}), \ell_{r,\eta})$ as $\ell_{r,\eta} \rightarrow \kappa_r/(a\sigma_r^2)^-$:

$$\lim_{\ell_{r,\eta} \rightarrow \kappa_r/(a\sigma_r^2)^-} \mathcal{R}(\ell_{r,\eta}) = \left(\delta_\beta + \kappa_r + \kappa_\eta + a\sigma_r^2 \bar{\ell}_{r,r} \right) \frac{\kappa_r}{a\sigma_r^2} > 0.$$

I have shown that $\mathcal{R}(\ell_{r,\eta})$ is continuous in the domain. The continuity follows from the fact that $\ell_{r,r}(\ell_{r,\eta})$ and $\ell_{\eta,\eta}(\ell_{r,\eta})$ are continuous functions of $\ell_{r,\eta}$ by the implicit function theorem, since the partial derivatives $\partial_{\ell_{r,r}} R_{r,r}$ and $\partial_{\ell_{\eta,\eta}} R_{\eta,\eta}$ are non-zero. Additionally, I have shown that $\lim_{\ell_{r,\eta} \rightarrow 0^+} \mathcal{R}(\ell_{r,\eta}) < 0$ and $\lim_{\ell_{r,\eta} \rightarrow \kappa_r/(a\sigma_r^2)^-} \mathcal{R}(\ell_{r,\eta}) > 0$. By the intermediate value theorem there exists at least one $0 < \ell_{r,\eta} < \frac{\kappa_r}{a\sigma_r^2}$ such that $R(\ell_{r,\eta}) = 0$. For this $\ell_{r,\eta}$, the corresponding $\ell_{r,r}(\ell_{r,\eta})$ and $\ell_{\eta,\eta}(\ell_{r,\eta})$ are positive and satisfy the first and third equations of the system. Thus the system has a solution $(\ell_{r,r}, \ell_{\eta,\eta}, \ell_{r,\eta})$ with all components positive.

The solution satisfies $\ell_{r,r} > 0$, $\ell_{\eta,\eta} > 0$ and $0 < \ell_{r,\eta} < \frac{\kappa_r}{a\sigma_r^2}$. Moreover, recall $\Psi_{r,r} = \kappa_r + a\sigma_r^2 \ell_{r,r}$, $\Psi_{r,\eta} = a\sigma_\eta^2 \ell_{r,\eta}$, $\Psi_{\eta,r} = -\kappa_r + a\sigma_r^2 \ell_{r,\eta}$ and $\Psi_{\eta,\eta} = \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta}$. It follows that $\Psi_{r,r} > \kappa_r$, $0 < \Psi_{r,\eta} < \frac{\kappa_r \sigma_\eta^2}{\sigma_r^2}$, $-\kappa_r < \Psi_{\eta,r} < 0$, $\Psi_{\eta,\eta} > \kappa_\eta$.

To complete the proof of existence of a solution, I show that the integrals in the Laplace

transforms converge. This property is required when performing the integration by parts in Proof of Proposition 1.1. In particular, the vector ℓ and the matrix L involve the integral terms $\beta e^{-\tau\delta_\beta} A_i(\tau)$ and $\beta e^{-\tau\delta_\beta} A_i(\tau) A_j(\tau)$ for $i, j \in \{r, \eta\}$. Given the solutions $A_r(\tau)$ and $A_\eta(\tau)$, these integral terms are just linear combinations of terms like $\beta e^{-\tau\delta_\beta} \frac{1-e^{-\tau v_i}}{v_i}$ and $\beta e^{-\tau\delta_\beta} \frac{1-e^{-\tau v_i}}{v_i} \frac{1-e^{-\tau v_j}}{v_j}$ or, focusing on the exponential components, $e^{-\tau\delta_\beta} e^{-\tau v_i}$ and $e^{-\tau\delta_\beta} e^{-\tau(v_i+v_j)}$, for $i, j \in \{r, \eta\}$. Convergence occurs if $\delta_\beta + \text{Re}(v_i) > 0$ and $\delta_\beta + \text{Re}(v_i + v_j) > 0$ or, denoting by $v_{\min}$ the eigenvalue with the smallest real part, if $\text{Re}(v_{\min}) > -\delta_\beta$ and $\text{Re}(v_{\min}) > -\frac{\delta_\beta}{2}$. Given that $-\frac{\delta_\beta}{2} > -\delta_\beta$, I conclude that the Laplace-transform integrals converge if the real parts of the eigenvalues of Ψ exceed $-\frac{\delta_\beta}{2}$. To check this condition, I examine the characteristic polynomial of Ψ :

$$\begin{aligned} P(v) &\equiv \det(\Psi - v\mathcal{I}_2) \\ &= (\Psi_{r,r} - v)(\Psi_{\eta,\eta} - v) - \Psi_{r,\eta}\Psi_{\eta,r} \end{aligned}$$

Evaluating $P(v)$ at $v = -\frac{\delta_\beta}{2}$:

$$\begin{aligned} P\left(-\frac{\delta_\beta}{2}\right) &= \left(-\frac{\delta_\beta}{2}\right)^2 - (\Psi_{r,r} + \Psi_{\eta,\eta})\left(-\frac{\delta_\beta}{2}\right) + \Psi_{r,r}\Psi_{\eta,\eta} - \Psi_{r,\eta}\Psi_{\eta,r} \\ &= \frac{\delta_\beta^2}{4} + (\Psi_{r,r} + \Psi_{\eta,\eta})\frac{\delta_\beta}{2} + \Psi_{r,r}\Psi_{\eta,\eta} - \Psi_{r,\eta}\Psi_{\eta,r} > 0 \end{aligned}$$

Therefore, if the eigenvalues (i.e. the roots of the characteristic equation $P(v) = 0$) are real, they must exceed $-\frac{\delta_\beta}{2}$. If the eigenvalues are complex, their real part is:

$$\text{Re}(v) = \frac{\Psi_{r,r} + \Psi_{\eta,\eta}}{2} > 0$$

and exceeds $-\frac{\delta_\beta}{2}$.

A.1.9 Corollary 1.1

Substituting Ψ into the ODE system, I get a system of two first-order ODEs:

$$\partial_\tau A_r(\tau) + A_r(\tau) \left(\kappa_r + a\sigma_r^2 \ell_{r,r} \right) + A_\eta(\tau) a\sigma_\eta^2 \ell_{r,\eta} - 1 = 0,$$

$$\partial_\tau A_\eta(\tau) - A_r(\tau) \left(\kappa_r - a\sigma_r^2 \ell_{r,\eta} \right) + A_\eta(\tau) \left(\kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta} \right) = 0.$$

Substituting in the solutions for $A_r(\tau)$ and $A_\eta(\tau)$:

$$\begin{aligned} & \left(\phi_r (v_1 - v_2) + \kappa_r + a\sigma_r^2 \ell_{r,r} - v_1 \right) \frac{1 - e^{-\tau v_1}}{v_1} \\ & + \left(\phi_r \left(\kappa_r + a\sigma_r^2 \ell_{r,r} - v_2 \right) + \phi_\eta a\sigma_\eta^2 \ell_{r,\eta} \right) \left(\frac{1 - e^{-\tau v_2}}{v_2} - \frac{1 - e^{-\tau v_1}}{v_1} \right) = 0, \end{aligned}$$

$$\begin{aligned} & \left(\phi_\eta (v_1 - v_2) - \kappa_r + a\sigma_r^2 \ell_{r,\eta} \right) \frac{1 - e^{-\tau v_1}}{v_1} \\ & \left(-\phi_r \left(\kappa_r - a\sigma_r^2 \ell_{r,\eta} \right) + \phi_\eta \left(\kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta} - v_2 \right) \right) \left(\frac{1 - e^{-\tau v_2}}{v_2} - \frac{1 - e^{-\tau v_1}}{v_1} \right) = 0. \end{aligned}$$

Identifying the terms linear in $\frac{1-e^{-\tau v_1}}{v_1}$ and $\frac{1-e^{-\tau v_2}}{v_2} - \frac{1-e^{-\tau v_1}}{v_1}$:

$$\phi_r (v_1 - v_2) + \kappa_r + a\sigma_r^2 \ell_{r,r} - v_1 = 0,$$

$$\phi_r \left(\kappa_r + a\sigma_r^2 \ell_{r,r} - v_2 \right) + \phi_\eta a\sigma_\eta^2 \ell_{r,\eta} = 0,$$

$$\phi_\eta (v_1 - v_2) - \kappa_r + a\sigma_r^2 \ell_{r,\eta} = 0,$$

$$-\phi_r \left(\kappa_r - a\sigma_r^2 \ell_{r,\eta} \right) + \phi_\eta \left(\kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta} - v_2 \right) = 0.$$

Substituting the first expression into the second and the third expression into the fourth:

$$\phi_r (1 - \phi_r) (v_1 - v_2) + \phi_\eta a\sigma_\eta^2 \ell_{r,\eta} = 0,$$

$$-\phi_r (v_1 - v_2) + \kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta} - v_2 = 0.$$

The first and third equations, and the two equations above constitute a system of four equations in the four unknowns $(v_1, v_2, \phi_r, \phi_\eta)$. Substituting the solutions for $A_r(\tau)$ and $A_\eta(\tau)$ into the Laplace transforms:

$$\ell_{r,r} = F(v_1, v_1) + 2\phi_r \hat{F}(v_1, v_1) + \phi_r^2 \hat{\hat{F}}(v_1, v_2),$$

$$\ell_{\eta,\eta} = \phi_{\eta}^2 \hat{\hat{F}}(v_1, v_2),$$

$$\ell_{r,\eta} = \phi_{\eta} \left(\hat{F}(v_1, v_2) + \phi_r \hat{\hat{F}}(v_1, v_2) \right)$$

where:

$$F(v, v') \equiv \int_0^{\infty} \beta(\tau) \frac{1 - e^{-\tau v}}{v} \frac{1 - e^{-\tau v'}}{v'} d\tau,$$

$$\hat{F}(v, v') \equiv F(v, v') - F(v, v),$$

$$\hat{\hat{F}}(v, v') \equiv F(v, v) + F(v', v') - 2F(v, v').$$

Substituting the Laplace transforms into the system:

$$\phi_r (v_1 - v_2) + \kappa_r + a\sigma_r^2 \left(F(v_1, v_1) + 2\phi_r \hat{F}(v_1, v_2) + \phi_r^2 \hat{\hat{F}}(v_1, v_2) \right) - v_1 = 0,$$

$$\phi_{\eta} (v_1 - v_2) - \kappa_r + a\sigma_r^2 \phi_{\eta} \left(\hat{F}(v_1, v_2) + \phi_r \hat{\hat{F}}(v_1, v_2) \right) = 0,$$

$$\phi_{\eta} (1 - \phi_r) (v_1 - v_2) + a\sigma_{\eta}^2 \phi_{\eta}^2 \left(\hat{F}(v_1, v_2) + \phi_r \hat{\hat{F}}(v_1, v_2) \right) = 0,$$

$$-\phi_r (v_1 - v_2) + \kappa_{\eta} + a\sigma_{\eta}^2 \phi_{\eta}^2 \hat{\hat{F}}(v_1, v_2) - v_2 = 0.$$

Let $a \approx 0$. Setting $(\phi_r, \phi_{\eta}) = (ac_r, c_{\eta})$, the system can be re-written as:

$$ac_r (v_1 - v_2) + \kappa_r + a\sigma_r^2 \left(F(v_1, v_1) + 2ac_r \hat{F}(v_1, v_2) + a^2 c_r^2 \hat{\hat{F}}(v_1, v_2) \right) - v_1 = 0,$$

$$c_{\eta} (v_1 - v_2) - \kappa_r + a\sigma_r^2 c_{\eta} \left(\hat{F}(v_1, v_2) + ac_r \hat{\hat{F}}(v_1, v_2) \right) = 0,$$

$$c_r (1 - ac_r) (v_1 - v_2) + \sigma_\eta^2 c_\eta^2 \left(\hat{F}(v_1, v_2) + ac_r \hat{\hat{F}}(v_1, v_2) \right) = 0,$$

$$-ac_r (v_1 - v_2) + \kappa_\eta + a\sigma_\eta^2 c_\eta^2 \hat{\hat{F}}(v_1, v_2) - v_2 = 0.$$

Taking the limit as $a \rightarrow 0$:

$$v_1 = \kappa_r,$$

$$c_\eta = \frac{\kappa_r}{\kappa_r - \kappa_\eta} \equiv \underline{c}_\eta,$$

$$c_r = -\frac{\kappa_r^2 \sigma_\eta^2}{(\kappa_r - \kappa_\eta)^3} \left(\int_0^\infty \beta(\tau) \frac{1 - e^{-\tau \kappa_r}}{\kappa_r} \frac{1 - e^{-\tau \kappa_\eta}}{\kappa_\eta} d\tau - \int_0^\infty \beta(\tau) \left(\frac{1 - e^{-\tau \kappa_r}}{\kappa_r} \right)^2 d\tau \right) \equiv \underline{c}_r,$$

$$v_2 = \kappa_\eta.$$

A.1.10 Proposition 1.6

Recall the system of ODEs:

$$\partial_\tau A_r(\tau) + A_r(\tau) \left(\kappa_r + a\sigma_r^2 \ell_{r,r} \right) + A_\eta(\tau) a\sigma_\eta^2 \ell_{r,\eta} - 1 = 0,$$

$$\partial_\tau A_\eta(\tau) - A_r(\tau) \left(\kappa_r - a\sigma_r^2 \ell_{r,\eta} \right) + A_\eta(\tau) \left(\kappa_\eta + a\sigma_\eta^2 \ell_{\eta,\eta} \right) = 0.$$

First, using the initial conditions $A_r(0) = 0$ and $A_\eta(0) = 0$, the first ODE implies $\partial_\tau A_r(0) = 1$. Therefore, $A_r(\tau) > 0$ for small τ . Second, using the initial conditions $A_r(0) = 0$ and $A_\eta(0) = 0$, the second ODE implies $\partial_\tau A_\eta(0) = 0$. Differentiating the second ODE at zero, and using $-\ell_{r,\eta} > 0$, which follows from $\Psi_{\eta,r} < 0$ and Ψ , I find $\partial_\tau^2 A_\eta(0) > 0$. Therefore, $\partial_\tau A_\eta(\tau) > 0$ and $A_\eta(\tau) > 0$ for small τ .

Recall the solutions $A_r(\tau) = \phi_r \left(\frac{1-e^{-\tau v_2}}{v_2} - \frac{1-e^{-\tau v_1}}{v_1} \right) + \frac{1-e^{-\tau v_1}}{v_1}$ and $A_\eta(\tau) = \phi_\eta \left(\frac{1-e^{-\tau v_2}}{v_2} - \frac{1-e^{-\tau v_1}}{v_1} \right)$.

Differentiating:

$$\partial_\tau A_r(\tau) = \phi_r (e^{-\tau v_2} - e^{-\tau v_1}) + e^{-\tau v_1},$$

$$\partial_\tau A_\eta(\tau) = \phi_\eta (e^{-\tau v_2} - e^{-\tau v_1}),$$

$$\partial_\tau \frac{A_r(\tau)}{A_\eta(\tau)} = \frac{e^{-\tau(v_1+v_2)} (v_1 (e^{\tau v_2} - 1) - v_2 (e^{\tau v_1} - 1))}{v_1 v_2 \phi_\eta \left(\frac{1-e^{-\tau v_2}}{v_2} - \frac{1-e^{-\tau v_1}}{v_1} \right)^2}.$$

Suppose that the two eigenvalues of Ψ are real, and without loss of generality set $v_1 > v_2$. First, let $g_1(v, \tau) = \frac{1-e^{-\tau v}}{v}$ and note that $\partial_v g_1(v, \tau) = \frac{e^{-\tau v}(1+\tau v)-1}{v^2} < 0$, because $(1+\tau v) < e^{\tau v}$. Therefore, the term in parentheses of $A_\eta(\tau)$ is positive. Given that $A_\eta(\tau) > 0$ for small τ , then $\phi_\eta > 0$, and hence $A_\eta(\tau) > 0$ for all τ . Second, let $g_2(v, \tau) = e^{-\tau v}$ and note that $\partial_v g_2(v, \tau) = -\tau e^{-\tau v} < 0$. Therefore, the term in parentheses of $\partial_\tau A_\eta(\tau)$ is positive. Given $\phi_\eta > 0$, then $\partial_\tau A_\eta(\tau) > 0$ for all τ . Third, let $g_3(v_1, v_2, \tau) = \frac{v_1(e^{\tau v_2}-1)-v_2(e^{\tau v_1}-1)}{v_1 v_2}$ and note that $g_3(v_1, v_2, 0) = 0$ and $\partial_\tau g_3(v_1, v_2, \tau) = e^{\tau v_2} - e^{\tau v_1} < 0$. Therefore, $\partial_\tau \frac{A_r(\tau)}{A_\eta(\tau)} < 0$. Fourth, note that $\partial_\tau A_r(\tau)$ has at most one root. Therefore, it can change sign at most once. Given $\partial_\tau A_r(\tau) > 0$ for small τ , then $\partial_\tau A_r(\tau) > 0$ for $\tau \in (0, \bar{\tau}')$ and $\partial_\tau A_r(\tau) < 0$ for $\tau \in (\bar{\tau}', \infty)$, where $\bar{\tau}'$ is a threshold in $(0, \infty)$. Given the behavior of $\partial_\tau A_r(\tau)$, and $A_r(\tau) > 0$ for small τ , then $A_r(\tau) > 0$ for $\tau \in (0, \bar{\tau})$ and $A_r(\tau) < 0$ for $\tau \in (\bar{\tau}, \infty)$, where $\bar{\tau}$ is a threshold in $(0, \infty)$. When $a \approx 0$, Corollary 1 implies $\phi_r \approx 0$, $v_1 \approx \kappa_r > 0$ and $v_2 \approx \kappa_s > 0$. From the solution for $A_r(\tau)$ it follows $A_r(\tau) > 0$. Moreover, when $\beta(\tau) = 0$, the definition of L implies $\ell_{r,r} = \ell_{r,s} = 0$. From the ODE of $A_r(\tau)$ it follows that $A_r(\tau) = \frac{1-e^{-\tau \kappa_r}}{\kappa_r} > 0$. In both cases, $\bar{\tau} = \infty$. When $a \approx \infty$, Corollary 1 implies that for τ bounded away from zero:

$$\begin{aligned} A_r(\tau) &\approx \frac{1}{a^{\frac{1}{3}}} \left(\frac{1}{\bar{n}_0} + \bar{c}_r \frac{1-e^{-\tau \bar{v}_1}}{\bar{v}_1} \right) \\ &= \frac{1}{a^{\frac{1}{3}} \bar{n}_0} \left(1 - \frac{\int_0^\infty \beta(u) \frac{1-e^{-u \bar{v}_1}}{\bar{v}_1} du}{\int_0^\infty \beta(u) \left(\frac{1-e^{-u \bar{v}_1}}{\bar{v}_1} \right)^2 du} \frac{1-e^{-\tau \bar{v}_1}}{\bar{v}_1} \right) \\ &= \frac{1}{a^{\frac{1}{3}} \bar{n}_0} \frac{\int_0^\infty \beta(u) (1-e^{-u \bar{v}_1}) (e^{-\tau \bar{v}_1} - e^{-u \bar{v}_1}) du}{\int_0^\infty \beta(u) (1-e^{-u \bar{v}_1})^2 du}. \end{aligned}$$

Since this is negative for τ close to ∞ , then $\bar{\tau} < \infty$.

Suppose the eigenvalues of Ψ are complex. Since they are conjugates, I set $v_1 = \mu + i\xi$ and $v_2 = \mu - i\xi$ for (μ, ξ) real. Substituting in the solutions for $A_r(\tau)$ and $A_\eta(\tau)$, and using Euler's equation $e^{\pm ix} = \cos(x) \pm i \sin(x)$:

$$\begin{aligned} A_r(\tau) &= \frac{(2\phi_r - 1)i\xi + \mu}{\mu^2 + \xi^2} - \frac{(2\phi_r - 1)i\xi + \mu}{\mu^2 + \xi^2} e^{-\tau\mu} \cos(\tau\xi) - \frac{(2\phi_r - 1)i\mu - \xi}{\mu^2 + \xi^2} e^{-\tau\mu} \sin(\tau\xi) \\ &= \phi_{r,0} + \phi_{r,1} e^{-\tau\mu} \cos(\tau\xi) + \phi_{r,2} e^{-\tau\mu} \sin(\tau\xi), \end{aligned}$$

$$\begin{aligned} A_\eta(\tau) &= \frac{2\phi_\eta i\xi}{\mu^2 + \xi^2} - \frac{2\phi_\eta i\xi}{\mu^2 + \xi^2} e^{-\tau\mu} \cos(\tau\xi) - \frac{2\phi_\eta i\mu}{\mu^2 + \xi^2} e^{-\tau\mu} \sin(\tau\xi) \\ &= \phi_{\eta,0} + \phi_{\eta,1} e^{-\tau\mu} \cos(\tau\xi) + \phi_{\eta,2} e^{-\tau\mu} \sin(\tau\xi), \end{aligned}$$

for real numbers $\{\phi_{j,n}\}_{j=r,\eta;n=0,1,2}$. Since the initial conditions $A_r(0) = 0$ and $A_\eta(0) = 0$ imply $\phi_{r,0} + \phi_{r,1} = 0$ and $\phi_{\eta,0} + \phi_{\eta,1} = 0$, the condition $\partial_\tau A_r(0) = 1$ implies $-\mu\phi_{r,1} + \xi\phi_{r,2} = 1$, and the condition $\partial_\tau A_\eta(0) = 0$ implies $-\mu\phi_{\eta,1} + \xi\phi_{\eta,2} = 0$:

$$A_r(\tau) = \phi_{r,0} \left(1 - e^{-\tau\mu} \cos(\tau\xi) - \frac{\mu}{\xi} e^{-\tau\mu} \sin(\tau\xi) \right) + \frac{1}{\xi} e^{-\tau\mu} \sin(\tau\xi),$$

$$A_\eta(\tau) = \phi_{\eta,0} \left(1 - e^{-\tau\mu} \cos(\tau\xi) - \frac{\mu}{\xi} e^{-\tau\mu} \sin(\tau\xi) \right).$$

Differentiating:

$$\partial_\tau A_r(\tau) = \phi_{r,0} \frac{\xi^2 + \mu^2}{\xi} e^{-\tau\mu} \sin(\tau\xi) + e^{-\tau\mu} \left(\cos(\tau\xi) - \frac{\mu}{\xi} \sin(\tau\xi) \right),$$

$$\partial_\tau A_\eta(\tau) = \phi_{\eta,0} \frac{\xi^2 + \mu^2}{\xi} e^{-\tau\mu} \sin(\tau\xi),$$

$$\partial_\tau \frac{A_r(\tau)}{A_\eta(\tau)} = \frac{\phi_{\eta,0} e^{-\tau\mu} \left(\cos(\tau\xi) - \frac{\mu}{\xi} \sin(\tau\xi) - e^{-\tau\mu} \right)}{\phi_{\eta,0}^2 \left(1 - e^{-\tau\mu} \cos(\tau\xi) - \frac{\mu}{\xi} e^{-\tau\mu} \sin(\tau\xi) \right)^2}.$$

First, note that if $\xi > 0$, then $\sin(\tau\xi) > 0$ for $0 < \tau\xi < \pi \rightarrow 0 < \tau < \frac{\pi}{\xi}$. Dividing by $\xi > 0$ yields a positive number. If $\xi < 0$, then $\sin(\tau\xi) < 0$ for $-\pi < \tau\xi < 0 \rightarrow 0 < \tau < \frac{\pi}{-\xi}$. Dividing by $\xi < 0$

yields a positive number. Given that $\partial_\tau A_\eta(\tau) > 0$ for small τ , then $\phi_{\eta,0} > 0$. Hence $\partial_\tau A_\eta(\tau) > 0$ for $\tau \in \left(0, \frac{\pi}{|\xi|}\right)$. Second, let $g_4(\tau) = \cos(\tau\xi) - \frac{\mu}{\xi} \sin(\tau\xi) - e^{-\tau\mu}$. Note that since $\phi_{\eta,0} > 0$, then $\partial_\tau \frac{A_r(\tau)}{A_\eta(\tau)} < 0$ when $g_4(\tau) < 0$. Moreover, since $\partial_\mu^2 g_4(\tau) = -\tau^2 e^{-\tau\mu}$, the function $g_4(\tau)$ is concave in μ and is maximised for $\mu = -\frac{1}{\tau} \ln \frac{\sin(\tau\xi)}{\tau\xi}$. The maximum is:

$$\frac{\sin(\tau\xi)}{\tau\xi} \left(\frac{\tau\xi \cos(\tau\xi)}{\sin(\tau\xi)} + \ln \frac{\sin(\tau\xi)}{\tau\xi} - 1 \right) = \frac{\sin(\tau\xi)}{\tau\xi} g_5(\tau\xi),$$

where:

$$g_5(x) \equiv \frac{x \cos(x)}{\sin(x)} + \ln \frac{\sin(x)}{x} - 1.$$

Differentiating:

$$\partial_x g_5(x) = -\frac{x^2 - 2x \cos(x) \sin(x) + \sin^2(x)}{x \sin^2(x)}.$$

Since:

$$x^2 - 2x \cos(x) \sin(x) + \sin^2(x) > x^2 - 2|x \sin(x)| + \sin^2(x) = (|x| - |\sin(x)|)^2 > 0$$

for $x \neq 0$, $\partial_x g_5(x) > 0$ for $x < 0$, and $\partial_x g_5(x) < 0$ for $x > 0$. Since, in addition, $\lim_{x \rightarrow 0} g_5(x) = 0$, then $g_5(x) < 0$. Hence, the maximum is negative for $\tau \in \left(0, \frac{\pi}{|\xi|}\right)$ and so is $\partial_\tau \frac{A_r(\tau)}{A_\eta(\tau)}$. In the proposition I define $\hat{\tau}_1 \equiv \frac{\pi}{|\xi|}$. Third, the result for $A_\eta(\tau)$ and the threshold $\hat{\tau}_2 > \hat{\tau}_1$ follows because $A_\eta(0) = 0$ and $\partial_\tau A_\eta(\tau) > 0$ for $\tau \in (0, \hat{\tau}_1)$ imply $A_\eta(\tau) > 0$ for $\tau \in (0, \hat{\tau}_1]$. Fourth, if $\hat{\tau}_2 < \infty$, then $A_\eta(\hat{\tau}_2) = 0$ and $\partial_\tau A_\eta(\hat{\tau}_2) \leq 0$. If $\partial_\tau A_\eta(\hat{\tau}_2) < 0$, then $-\ell_{r,\eta} > 0$ and the ODE of $A_\eta(\tau)$ imply $A_r(\hat{\tau}_2) < 0$. If $\partial_\tau A_\eta(\hat{\tau}_2) = 0$, then $-\ell_{r,\eta} > 0$ and the ODE of $A_\eta(\tau)$ imply $A_r(\hat{\tau}_2) = 0$, and the ODE of $A_r(\tau)$ implies $\partial_\tau A_r(\hat{\tau}_2) = 1$. Hence, in both cases, $A_r(\tau) < 0$ for τ smaller than and close to $\hat{\tau}_2$.

A.1.11 Proposition 1.7

Using the conjectured price of the bond, the dependent and independent variables in the Fama and Bliss (1987) regression are:

$$\frac{p_{t+\Delta\tau}(\tau - \Delta\tau) - p_t(\tau)}{\Delta\tau} - y_t(\Delta\tau) = \frac{A_r(\tau)r_t + A_\eta(\tau)\eta_t + C(\omega, \tau)}{\Delta\tau}$$

$$\begin{aligned}
& - \frac{A_r(\tau - \Delta\tau)r_{t+\Delta\tau} + A_\eta(\tau - \Delta\tau)\eta_{t+\Delta\tau} + C(\omega, \tau - \Delta\tau)}{\Delta\tau} \\
& - \frac{A_r(\Delta\tau)r_t + A_\eta(\Delta\tau)\eta_t + C(\omega, \Delta\tau)}{\Delta\tau}, \\
f_t(\tau - \Delta\tau, \tau) - y_t(\Delta\tau) &= \frac{A_r(\tau)r_t + A_\eta(\tau)\eta_t + C(\omega, \tau)}{\Delta\tau} \\
& - \frac{A_r(\tau - \Delta\tau)r_t + A_\eta(\tau - \Delta\tau)\eta_t + C(\omega, \tau - \Delta\tau)}{\Delta\tau} \\
& - \frac{A_r(\Delta\tau)r_t + A_\eta(\Delta\tau)\eta_t + C(\omega, \Delta\tau)}{\Delta\tau}.
\end{aligned}$$

The regression coefficient is:

$$b_{FB} = \frac{\mathbb{Cov}\left(f_t(\tau - \Delta\tau, \tau) - y_t(\Delta\tau), \frac{p_{t+\Delta\tau}(\tau - \Delta\tau) - p_t(\tau)}{\Delta\tau} - y_t(\Delta\tau)\right)}{\mathbb{Var}(f_t(\tau - \Delta\tau, \tau) - y_t(\Delta\tau))}.$$

Note that I can take out $\Delta\tau^{-2}$ from both the covariance in the numerator and the variance in the denominator. Therefore, these terms cancel out when computing b_{FB} , and I ignore them in the following calculations. Moreover, conditioning on the state ω , the terms that contain the function $C(\omega, \cdot)$ are constant, and therefore their covariances and variances with other terms are zero. Thus, the numerator of the b_{FB} coefficient can be written as:

$$\begin{aligned}
N_{FB} &= \mathbb{Cov}\left((A_r(\tau) - A_r(\tau - \Delta\tau) - A_r(\Delta\tau))r_t + (A_\eta(\tau) - A_\eta(\tau - \Delta\tau) - A_\eta(\Delta\tau))\eta_t, \right. \\
&\quad \left. (A_r(\tau) - A_r(\Delta\tau))r_t - A_r(\tau - \Delta\tau)r_{t+\Delta\tau} + (A_\eta(\tau) - A_\eta(\Delta\tau))\eta_t - A_\eta(\tau - \Delta\tau)\eta_{t+\Delta\tau}\right) \\
&= b_{FB,r,1}b_{FB,r,2}\mathbb{Var}(r_t) + b_{FB,\eta,1}b_{FB,\eta,2}\mathbb{Var}(\eta_t) \\
&\quad - b_{FB,r,1}b_{FB,r,3}\mathbb{Cov}(r_t, r_{t+\Delta\tau}) - b_{FB,\eta,1}b_{FB,\eta,3}\mathbb{Cov}(\eta_t, \eta_{t+\Delta\tau}) \\
&\quad - b_{FB,r,1}b_{FB,\eta,3}\mathbb{Cov}(r_t, \eta_{t+\Delta\tau}) - b_{FB,\eta,1}b_{FB,r,3}\mathbb{Cov}(\eta_t, r_{t+\Delta\tau}) \\
&\quad + (b_{FB,r,1}b_{FB,\eta,2} + b_{FB,\eta,1}b_{FB,r,2})\mathbb{Cov}(r_t, \eta_t),
\end{aligned}$$

and the denominator can be written as:

$$\begin{aligned}
D_{FB} &= \mathbb{Var}\left((A_r(\tau) - A_r(\tau - \Delta\tau) - A_r(\Delta\tau))r_t + (A_\eta(\tau) - A_\eta(\tau - \Delta\tau) - A_\eta(\Delta\tau))\eta_t\right) \\
&= b_{FB,r,1}^2\mathbb{Var}(r_t) + b_{FB,\eta,1}^2\mathbb{Var}(\eta_t) + 2b_{FB,r,1}b_{FB,\eta,1}\mathbb{Cov}(r_t, \eta_t),
\end{aligned}$$

where, for $j \in \{r, \eta\}$:

$$b_{FB,j,1} \equiv A_j(\tau) - A_j(\tau - \Delta\tau) - A_j(\Delta\tau),$$

$$b_{FB,j,2} \equiv A_j(\tau) - A_j(\Delta\tau),$$

$$b_{FB,j,3} \equiv A_j(\tau - \Delta\tau),$$

Note that r_t and η_t can be written as a multivariate Ornstein-Uhlenbeck process:

$$d\mathbf{x}_t = (\mathbf{R}\mathbf{x}_t + \mathbf{b}) dt + \mathbf{\Sigma} d\mathbf{z}_t,$$

where:

$$\mathbf{x}_t = \begin{bmatrix} r_t \\ \eta_t \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} -\kappa_r & \kappa_r \\ 0 & -\kappa_\eta \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} 0 \\ \kappa_\eta r_t^*(\omega) \end{bmatrix}, \quad d\mathbf{z}_t = \begin{bmatrix} dz_{r,t} \\ dz_{\eta,t} \end{bmatrix}.$$

Let the covariance matrix of the process $\mathbf{x}_t$ be:

$$\mathbf{V}_t \equiv \begin{bmatrix} \text{Var}(r_t) & \text{Cov}(r_t, \eta_t) \\ \text{Cov}(r_t, \eta_t) & \text{Var}(\eta_t) \end{bmatrix} \equiv \begin{bmatrix} V_{rr,t} & V_{r\eta,t} \\ V_{r\eta,t} & V_{\eta\eta,t} \end{bmatrix}.$$

To find variances and covariances of these processes I derive and solve the differential equations for these moments. The time evolution of the covariance matrix $\mathbf{V}_t$ for a multivariate Ornstein-Uhlenbeck process is given by the Lyapunov equation:

$$\begin{aligned} \frac{d\mathbf{V}_t}{dt} &= \mathbf{R}\mathbf{V}_t + \mathbf{V}_t\mathbf{R}^\top + \mathbf{\Sigma}\mathbf{\Sigma}^\top \\ &= \begin{bmatrix} -2\kappa_r V_{rr,t} + 2\kappa_r V_{r\eta,t} + \sigma_r^2 & -(\kappa_r + \kappa_\eta) V_{r\eta,t} + \kappa_r V_{\eta\eta,t} \\ -(\kappa_r + \kappa_\eta) V_{r\eta,t} + \kappa_r V_{\eta\eta,t} & -2\kappa_\eta V_{\eta\eta,t} + \sigma_\eta^2 \end{bmatrix}. \end{aligned}$$

This gives a system of ordinary differential equations for the elements of the covariance matrix:

$$\frac{dV_{rr,t}}{dt} = -2\kappa_r V_{rr,t} + 2\kappa_r V_{r\eta,t} + \sigma_r^2,$$

$$\frac{dV_{r\eta,t}}{dt} = -(\kappa_r + \kappa_\eta) V_{r\eta,t} + \kappa_r V_{\eta\eta,t},$$

$$\frac{dV_{\eta\eta,t}}{dt} = -2\kappa_\eta V_{\eta\eta,t} + \sigma_\eta^2.$$

Focusing on the stationary distribution (as $t \rightarrow \infty$), where the variances and covariances become constant, I can set the time derivatives to zero. Solving the system, and denoting by V_{rr} , $V_{r\eta}$ and $V_{\eta\eta}$ the stationary values yields:

$$V_{rr} = \frac{\sigma_r^2}{2\kappa_r} + \frac{\kappa_r \sigma_\eta^2}{2\kappa_\eta (\kappa_r + \kappa_\eta)},$$

$$V_{r\eta} = \frac{\kappa_r \sigma_\eta^2}{2\kappa_\eta (\kappa_r + \kappa_\eta)},$$

$$V_{\eta\eta} = \frac{\sigma_\eta^2}{2\kappa_\eta}.$$

Let the stationary cross-covariance at lag $\Delta\tau$ be:

$$\mathbf{C}(\Delta\tau) \equiv \begin{bmatrix} \text{Cov}(r_t, r_{t+\Delta\tau}) & \text{Cov}(r_t, \eta_{t+\Delta\tau}) \\ \text{Cov}(\eta_t, r_{t+\Delta\tau}) & \text{Cov}(\eta_t, \eta_{t+\Delta\tau}) \end{bmatrix} \equiv \begin{bmatrix} C_{rr,t} & C_{r\eta,t} \\ C_{\eta r,t} & C_{\eta\eta,t} \end{bmatrix}.$$

For the multivariate Ornstein-Uhlenbeck process $\mathbf{x}_t$ with drift matrix $\mathbf{R}$ and stationary covariance $\mathbf{V}$, this covariance is given by:

$$\mathbf{C}(\Delta\tau) = \mathbf{V} \left(e^{\Delta\tau \mathbf{R}} \right)^\top = \begin{bmatrix} V_{rr} e^{-\Delta\tau \kappa_r} + V_{r\eta} \frac{\kappa_r (e^{-\Delta\tau \kappa_\eta} - e^{-\Delta\tau \kappa_r})}{\kappa_r - \kappa_\eta} & V_{r\eta} e^{-\Delta\tau \kappa_\eta} \\ V_{r\eta} e^{-\Delta\tau \kappa_r} + V_{\eta\eta} \frac{\kappa_r (e^{-\Delta\tau \kappa_\eta} - e^{-\Delta\tau \kappa_r})}{\kappa_r - \kappa_\eta} & V_{\eta\eta} e^{-\Delta\tau \kappa_\eta} \end{bmatrix}.$$

This yields:

$$C_{rr} = e^{-\Delta\tau \kappa_r} V_{rr} + \frac{\kappa_r (e^{-\Delta\tau \kappa_\eta} - e^{-\Delta\tau \kappa_r})}{\kappa_r - \kappa_\eta} V_{r\eta},$$

$$C_{r\eta} = e^{-\Delta\tau \kappa_\eta} V_{r\eta},$$

$$C_{\eta r} = e^{-\Delta\tau\kappa_r} V_{r\eta} + \frac{\kappa_r (e^{-\Delta\tau\kappa_\eta} - e^{-\Delta\tau\kappa_r})}{\kappa_r - \kappa_\eta} V_{\eta\eta},$$

$$C_{\eta\eta} = e^{-\Delta\tau\kappa_\eta} V_{\eta\eta}.$$

Therefore, the numerator of the b_{FB} coefficient can be written as:

$$\begin{aligned} N_{FB} = & b_{FB,r,1} b_{FB,r,2} V_{rr} + b_{FB,\eta,1} b_{FB,\eta,2} V_{\eta\eta} - b_{FB,r,1} b_{FB,r,3} C_{rr} - b_{FB,\eta,1} b_{FB,\eta,3} C_{\eta\eta} \\ & + (b_{FB,r,1} b_{FB,\eta,2} + b_{FB,\eta,1} b_{FB,r,2}) V_{r\eta} - b_{FB,r,1} b_{FB,\eta,3} C_{r\eta} - b_{FB,\eta,1} b_{FB,r,3} C_{\eta r}, \end{aligned}$$

and the denominator can be written as:

$$D_{FB} = b_{FB,r,1}^2 V_{rr} + b_{FB,\eta,1}^2 V_{\eta\eta} + 2b_{FB,r,1} b_{FB,\eta,1} V_{r\eta}.$$

Taylor expanding $A_j(\tau - \Delta\tau)$ around τ and $A_j(\Delta\tau)$ around 0, so as to consider a small deviation $\Delta\tau$, and recalling that $\partial_\tau A_r(0) = 1$ and $\partial_\tau A_\eta(0) = 0$:

$$A_r(\tau - \Delta\tau) = A_r(\tau) - \partial_\tau A_r(\tau) \Delta\tau + \mathcal{O}(\Delta\tau^2),$$

$$A_r(\Delta\tau) = \Delta\tau + \mathcal{O}(\Delta\tau^2),$$

$$A_\eta(\tau - \Delta\tau) = A_\eta(\tau) - \partial_\tau A_\eta(\tau) \Delta\tau + \mathcal{O}(\Delta\tau^2),$$

$$A_\eta(\Delta\tau) = \mathcal{O}(\Delta\tau^2),$$

It follows that:

$$b_{FB,r,1} = \partial_\tau A_r(\tau) \Delta\tau - \Delta\tau + \mathcal{O}(\Delta\tau^2),$$

$$b_{FB,r,2} = A_r(\tau) - \Delta\tau + \mathcal{O}(\Delta\tau^2),$$

$$b_{FB,r,3} = A_r(\tau) - \partial_\tau A_r(\tau) \Delta\tau + \mathcal{O}(\Delta\tau^2),$$

$$b_{FB,\eta,1} = \partial_\tau A_\eta(\tau) \Delta\tau + \mathcal{O}(\Delta\tau^2),$$

$$b_{FB,\eta,2} = A_\eta(\tau) + \mathcal{O}(\Delta\tau^2),$$

$$b_{FB,\eta,3} = A_\eta(\tau) - \partial_\tau A_\eta(\tau) \Delta\tau + \mathcal{O}(\Delta\tau^2),$$

Similarly, Taylor expanding $e^{-\Delta\tau\kappa_j}$ around 0:

$$C_{rr} = (1 - \Delta\tau\kappa_r) V_{rr} + \Delta\tau\kappa_r V_{r\eta},$$

$$C_{r\eta} = (1 - \Delta\tau\kappa_\eta) V_{r\eta},$$

$$C_{\eta r} = (1 - \Delta\tau\kappa_r) V_{r\eta} + \Delta\tau\kappa_r V_{\eta\eta},$$

$$C_{\eta\eta} = (1 - \Delta\tau\kappa_\eta) V_{\eta\eta}.$$

Therefore, the numerator of the b_{FB} coefficient can be written as:

$$\begin{aligned} N_{FB} = & \Delta\tau^2 (\partial_\tau A_r(\tau) + A_r(\tau)\kappa_r - 1) ((\partial_\tau A_r(\tau) - 1) V_{rr} + \partial_\tau A_\eta(\tau) V_{r\eta}) \\ & + \Delta\tau^2 (\partial_\tau A_\eta(\tau) + A_\eta(\tau)\kappa_\eta - A_r(\tau)\kappa_r) ((\partial_\tau A_r(\tau) - 1) V_{r\eta} + \partial_\tau A_\eta(\tau) V_{\eta\eta}), \end{aligned}$$

and the denominator can be written as:

$$D_{FB} = \Delta\tau^2 (\partial_\tau A_r(\tau) - 1)^2 V_{rr} + \Delta\tau^2 (\partial_\tau A_\eta(\tau))^2 V_{\eta\eta} + 2\Delta\tau^2 (\partial_\tau A_r(\tau) - 1) \partial_\tau A_\eta(\tau) V_{r\eta}.$$

It follows that the b_{FB} coefficient can be written as:

$$b_{FB} = ((\partial_\tau A_r(\tau) + A_r(\tau)\kappa_r - 1) ((\partial_\tau A_r(\tau) - 1) V_{rr} + \partial_\tau A_\eta(\tau) V_{r\eta}) \\ + (\partial_\tau A_\eta(\tau) + A_\eta(\tau)\kappa_\eta - A_r(\tau)\kappa_r) ((\partial_\tau A_r(\tau) - 1) V_{r\eta} + \partial_\tau A_\eta(\tau) V_{\eta\eta})) \\ \left((\partial_\tau A_r(\tau) - 1)^2 V_{rr} + (\partial_\tau A_\eta(\tau))^2 V_{\eta\eta} + 2 (\partial_\tau A_r(\tau) - 1) \partial_\tau A_\eta(\tau) V_{r\eta} \right)^{-1}.$$

First, consider the denominator, and note that it can be written as:

$$\begin{bmatrix} \partial_\tau A_r(\tau) - 1 & \partial_\tau A_\eta(\tau) \end{bmatrix} \begin{bmatrix} V_{rr} & V_{r\eta} \\ V_{r\eta} & V_{\eta\eta} \end{bmatrix} \begin{bmatrix} \partial_\tau A_r(\tau) - 1 \\ \partial_\tau A_\eta(\tau) \end{bmatrix},$$

First, looking at the matrix V , one can see that the first leading principal minor V_{rr} is positive, and that its determinant $V_{rr}V_{\eta\eta} - V_{r\eta}^2 = \frac{\sigma_r^2\sigma_\eta^2}{4\kappa_r\kappa_\eta} + \frac{\kappa_r\sigma_\eta^4}{4\kappa_\eta(\kappa_r+\kappa_\eta)^2}$ is positive. Therefore, V is positive definite, and the denominator is positive. Second, consider the numerator. For $\tau < \min\{\bar{\tau}, \hat{\tau}_1\}$, $A_r(\tau) > 0$, $A_\eta(\tau) > 0$ and $\partial_\tau A_\eta(\tau) > 0$. Moreover, the system of ODEs implies:

$$\partial_\tau A_r(\tau) + A_r(\tau)\kappa_r - 1 = -A_r(\tau)a\sigma_r^2\ell_{r,r} - A_\eta(\tau)a\sigma_\eta^2\ell_{r,\eta} < 0,$$

$$\partial_\tau A_r(\tau) - 1 = -A_r(\tau) \left(\kappa_r + a\sigma_r^2\ell_{r,r} \right) - A_\eta(\tau)a\sigma_\eta^2\ell_{r,\eta} < 0,$$

$$\partial_\tau A_\eta(\tau) + A_\eta(\tau)\kappa_\eta - A_r(\tau)\kappa_r = -A_r(\tau)a\sigma_r^2\ell_{r,\eta} - A_\eta(\tau)a\sigma_\eta^2\ell_{\eta,\eta} < 0,$$

where the inequalities follow from the fact that $A_r(\tau) > 0$, $A_\eta(\tau) > 0$, $\ell_{r,r} \geq 0$, $\ell_{\eta,\eta} \geq 0$ and $\ell_{r,\eta} \geq 0$, which in turn follows from $\Psi_{r,\eta} > 0$. It follows that the leading coefficient of each of the two terms in the numerator is negative. Therefore, a sufficient condition for the numerator to be positive is that the second leading coefficient of each term is negative:

$$(\partial_\tau A_r(\tau) - 1) V_{rr} + \partial_\tau A_\eta(\tau) V_{r\eta} < 0 \\ -\frac{\partial_\tau A_\eta(\tau)}{\partial_\tau A_r(\tau) - 1} < \frac{V_{rr}}{V_{r\eta}},$$

$$(\partial_\tau A_r(\tau) - 1) V_{r\eta} + \partial_\tau A_\eta(\tau) V_{\eta\eta} < 0$$

$$-\frac{\partial_\tau A_\eta(\tau)}{\partial_\tau A_r(\tau) - 1} < \frac{V_{r\eta}}{V_{\eta\eta}}.$$

Both of these conditions hold for $\tau < \min\{\bar{\tau}, \hat{\tau}_1\}$ if:

$$-\frac{\partial_\tau A_\eta(\tau)}{\partial_\tau A_r(\tau) - 1} < \min\left\{\frac{V_{rr}}{V_{r\eta}}, \frac{V_{r\eta}}{V_{\eta\eta}}\right\} = \frac{\kappa_r}{\kappa_r + \kappa_\eta}.$$

When $a \approx 0$, the solutions for $A_r(\tau)$ and $A_\eta(\tau)$ and $(v_1, v_2, \phi_r, \phi_\eta) = (\kappa_r, \kappa_\eta, a\underline{c}_r, \underline{c}_\eta)$ imply:

$$b_{FB} = \frac{N_{FB,1} + N_{FB,2}}{D_{FB}} + \mathcal{O}(a^2),$$

where:

$$N_{FB,1} = \left(-a\sigma_r^2 \frac{1 - e^{-\tau\kappa_r}}{\kappa_r} F(\kappa_r, \kappa_r) - a\sigma_\eta^2 \underline{c}_\eta^2 \left(\frac{1 - e^{-\tau\kappa_\eta}}{\kappa_\eta} - \frac{1 - e^{-\tau\kappa_r}}{\kappa_r} \right) \hat{F}(\kappa_r, \kappa_\eta) \right) \\ \left((e^{-\tau\kappa_r} - 1) V_{rr} + \underline{c}_\eta (e^{-\tau\kappa_\eta} - e^{-\tau\kappa_r}) V_{r\eta} \right) + \mathcal{O}(a^2),$$

$$N_{FB,2} = \left(-a\sigma_r^2 \underline{c}_\eta \frac{1 - e^{-\tau\kappa_r}}{\kappa_r} \hat{F}(\kappa_r, \kappa_\eta) - a\sigma_\eta^2 \underline{c}_\eta^3 \left(\frac{1 - e^{-\tau\kappa_\eta}}{\kappa_\eta} - \frac{1 - e^{-\tau\kappa_r}}{\kappa_r} \right) \hat{F}(\kappa_r, \kappa_\eta) \right) \\ \left((e^{-\tau\kappa_r} - 1) V_{r\eta} + \underline{c}_\eta (e^{-\tau\kappa_\eta} - e^{-\tau\kappa_r}) V_{\eta\eta} \right) + \mathcal{O}(a^2),$$

$$D_{FB} = \left(2a\underline{c}_r (e^{-\tau\kappa_\eta} - e^{-\tau\kappa_r}) (e^{-\tau\kappa_r} - 1) + (e^{-\tau\kappa_r} - 1)^2 \right) V_{rr} \\ + \left(2a\underline{c}_r \underline{c}_\eta (e^{-\tau\kappa_\eta} - e^{-\tau\kappa_r})^2 + 2\underline{c}_\eta (e^{-\tau\kappa_r} - 1) (e^{-\tau\kappa_\eta} - e^{-\tau\kappa_r}) \right) V_{r\eta} \\ + \underline{c}_\eta^2 (e^{-\tau\kappa_\eta} - e^{-\tau\kappa_r})^2 V_{\eta\eta} + \mathcal{O}(a^2).$$

First, the denominator is positive because of an analogous reasoning to the one before. Second, consider the numerator. The leading coefficient of each of the two terms in the numerator is negative. Therefore, a sufficient condition for the numerator to be positive is that the second leading coefficient of each term is negative:

$$(e^{-\tau\kappa_r} - 1) V_{rr} + \underline{c}_\eta (e^{-\tau\kappa_\eta} - e^{-\tau\kappa_r}) V_{r\eta} < 0 \\ -\frac{\underline{c}_\eta (e^{-\tau\kappa_\eta} - e^{-\tau\kappa_r})}{e^{-\tau\kappa_r} - 1} < \frac{V_{rr}}{V_{r\eta}},$$

$$(e^{-\tau\kappa_r} - 1) V_{r\eta} + \underline{c}_\eta (e^{-\tau\kappa_\eta} - e^{-\tau\kappa_r}) V_{\eta\eta} < 0$$

$$-\frac{\underline{c}_\eta (e^{-\tau\kappa_\eta} - e^{-\tau\kappa_r})}{e^{-\tau\kappa_r} - 1} < \frac{V_{r\eta}}{V_{\eta\eta}}.$$

Both of these conditions hold if:

$$-\frac{\underline{c}_\eta (e^{-\tau\kappa_\eta} - e^{-\tau\kappa_r})}{e^{-\tau\kappa_r} - 1} < \min \left\{ \frac{V_{rr}}{V_{r\eta}}, \frac{V_{r\eta}}{V_{\eta\eta}} \right\} = \frac{\kappa_r}{\kappa_r + \kappa_\eta}.$$

Substituting the value of $\underline{c}_\eta$, the condition holds for τ such that:

$$\frac{e^{-\tau\kappa_\eta} - e^{-\tau\kappa_r}}{1 - e^{-\tau\kappa_r}} < \frac{\kappa_r - \kappa_\eta}{\kappa_r + \kappa_\eta}.$$

Using the conjectured price of the bond, the dependent and independent variables in the [Campbell and Shiller \(1991\)](#) regression are:

$$y_{t+\Delta\tau}(\tau - \Delta\tau) - y_t(\tau) = \frac{A_r(\tau - \Delta\tau)r_{t+\Delta\tau} + A_\eta(\tau - \Delta\tau)\eta_{t+\Delta\tau} + C(\omega, \tau - \Delta\tau)}{\tau - \Delta\tau}$$

$$- \frac{A_r(\tau)r_t + A_\eta(\tau)\eta_t + C(\omega, \tau)}{\tau},$$

$$\frac{\Delta\tau}{\tau - \Delta\tau} (y_t(\tau) - y_t(\Delta\tau)) = \frac{\Delta\tau}{\tau - \Delta\tau} \left(\frac{A_r(\tau)r_t + A_\eta(\tau)\eta_t + C(\omega, \tau)}{\tau} \right.$$

$$\left. - \frac{A_r(\Delta\tau)r_t + A_\eta(\Delta\tau)\eta_t + C(\omega, \Delta\tau)}{\Delta\tau} \right).$$

The regression coefficient is:

$$b_{CS} = \frac{\text{Cov} \left(\frac{\Delta\tau}{\tau - \Delta\tau} (y_t(\tau) - y_t(\Delta\tau)), y_{t+\Delta\tau}(\tau - \Delta\tau) - y_t(\tau) \right)}{\text{Var} \left(\frac{\Delta\tau}{\tau - \Delta\tau} (y_t(\tau) - y_t(\Delta\tau)) \right)}.$$

Note that I can take out $\frac{\Delta\tau}{\tau - \Delta\tau}$ from the covariance in the numerator and $\frac{\Delta\tau^2}{(\tau - \Delta\tau)^2}$ from the variance in the denominator. Therefore, these terms partially cancel out when computing b_{CS} , and only $\frac{\Delta\tau}{\tau - \Delta\tau}$ in the denominator survives in the following calculations. Moreover, conditioning on the state ω , the terms that contain the function $C(\omega, \cdot)$ are constant, and therefore their covariances and variances with other terms are zero. Thus, the numerator of the b_{CS} coefficient can be

written as:

$$\begin{aligned}
N_{CS} &= \mathbb{Cov} \left(\left(\frac{A_r(\tau)}{\tau} - \frac{A_r(\Delta\tau)}{\Delta\tau} \right) r_t + \left(\frac{A_\eta(\tau)}{\tau} - \frac{A_\eta(\Delta\tau)}{\Delta\tau} \right) \eta_t, \right. \\
&\quad \left. \frac{A_r(\tau - \Delta\tau)}{\tau - \Delta\tau} r_{t+\Delta\tau} + \frac{A_\eta(\tau - \Delta\tau)}{\tau - \Delta\tau} \eta_{t+\Delta\tau} - \frac{A_r(\tau)}{\tau} r_t - \frac{A_\eta(\tau)}{\tau} \eta_t \right) \\
&= -b_{CS,r,1} b_{CS,r,3} \mathbb{Var}(r_t) - b_{CS,\eta,1} b_{CS,\eta,3} \mathbb{Var}(\eta_t) + b_{CS,r,2} b_{CS,r,3} \mathbb{Cov}(r_t, r_{t+\Delta\tau}) \\
&\quad + b_{CS,\eta,2} b_{CS,\eta,3} \mathbb{Cov}(\eta_t, \eta_{t+\Delta\tau}) - (b_{CS,\eta,1} b_{CS,r,3} + b_{CS,r,1} b_{CS,\eta,3}) \mathbb{Cov}(r_t, \eta_t) \\
&\quad + b_{CS,\eta,2} b_{CS,r,3} \mathbb{Cov}(r_t, \eta_{t+\Delta\tau}) + b_{CS,r,2} b_{CS,\eta,3} \mathbb{Cov}(\eta_t, r_{t+\Delta\tau}),
\end{aligned}$$

and the denominator can be written as:

$$\begin{aligned}
D_{CS} &= \frac{\Delta\tau}{\tau - \Delta\tau} \mathbb{Var} \left(\left(\frac{A_r(\tau)}{\tau} - \frac{A_r(\Delta\tau)}{\Delta\tau} \right) r_t + \left(\frac{A_\eta(\tau)}{\tau} - \frac{A_\eta(\Delta\tau)}{\Delta\tau} \right) \eta_t \right) \\
&= \frac{\Delta\tau}{\tau - \Delta\tau} \left(b_{CS,r,3}^2 \mathbb{Var}(r_t) + b_{CS,\eta,3}^2 \mathbb{Var}(\eta_t) + 2b_{CS,r,3} b_{CS,\eta,3} \mathbb{Cov}(r_t, \eta_t) \right),
\end{aligned}$$

where, for $j \in \{r, \eta\}$:

$$b_{CS,j,1} = \frac{A_j(\tau)}{\tau},$$

$$b_{CS,j,2} = \frac{A_j(\tau - \Delta\tau)}{\tau - \Delta\tau},$$

$$b_{CS,j,3} = \frac{A_j(\tau)}{\tau} - \frac{A_j(\Delta\tau)}{\Delta\tau}.$$

Substituting the values of the variances and covariances as before, the numerator of the b_{CS} coefficient can be written as:

$$\begin{aligned}
N_{CS} &= -b_{CS,r,1} b_{CS,r,3} V_{rr} - b_{CS,\eta,1} b_{CS,\eta,3} V_{\eta\eta} + b_{CS,r,2} b_{CS,r,3} C_{rr} + b_{CS,\eta,2} b_{CS,\eta,3} C_{\eta\eta} \\
&\quad - (b_{CS,\eta,1} b_{CS,r,3} + b_{CS,r,1} b_{CS,\eta,3}) V_{r\eta} + b_{CS,\eta,2} b_{CS,r,3} C_{r\eta} + b_{CS,r,2} b_{CS,\eta,3} C_{\eta r},
\end{aligned}$$

and the denominator can be written as:

$$D_{CS} = \frac{\Delta\tau}{\tau - \Delta\tau} \left(b_{CS,r,3}^2 V_{rr} + b_{CS,\eta,3}^2 V_{\eta\eta} + 2b_{CS,r,3} b_{CS,\eta,3} V_{r\eta} \right).$$

Taylor expanding as before:

$$b_{CS,r,1} = \frac{A_r(\tau)}{\tau},$$

$$b_{CS,r,2} = \frac{A_r(\tau) - \partial_\tau A_r(\tau)\Delta\tau}{\tau - \Delta\tau},$$

$$b_{CS,r,3} = \frac{A_r(\tau)}{\tau} - 1,$$

$$b_{CS,\eta,1} = \frac{A_\eta(\tau)}{\tau},$$

$$b_{CS,\eta,2} = \frac{A_\eta(\tau) - \partial_\tau A_\eta(\tau)\Delta\tau}{\tau - \Delta\tau},$$

$$b_{CS,\eta,3} = \frac{A_\eta(\tau)}{\tau}.$$

Therefore, the numerator of the b_{CS} coefficient can be written as:

$$\begin{aligned} N_{CS} = & \frac{\Delta\tau}{\tau - \Delta\tau} \left(\frac{A_r(\tau)}{\tau} - \partial_\tau A_r(\tau) - A_r(\tau)\kappa_r \right) \left(\left(\frac{A_r(\tau)}{\tau} - 1 \right) V_{rr} + \frac{A_\eta(\tau)}{\tau} V_{r\eta} \right) \\ & + \frac{\Delta\tau}{\tau - \Delta\tau} \left(\frac{A_\eta(\tau)}{\tau} - \partial_\tau A_\eta(\tau) - A_\eta(\tau)\kappa_\eta + A_r(\tau)\kappa_r \right) \left(\frac{A_\eta(\tau)}{\tau} V_{\eta\eta} + \left(\frac{A_r(\tau)}{\tau} - 1 \right) V_{r\eta} \right), \end{aligned}$$

and the denominator can be written as:

$$D_{CS} = \frac{\Delta\tau}{\tau - \Delta\tau} \left(\left(\frac{A_r(\tau)}{\tau} - 1 \right)^2 V_{rr} + \left(\frac{A_\eta(\tau)}{\tau} \right)^2 V_{\eta\eta} + 2 \left(\frac{A_r(\tau)}{\tau} - 1 \right) \left(\frac{A_\eta(\tau)}{\tau} \right) V_{r\eta} \right).$$

It follows that the b_{CS} coefficient can be written as:

$$\begin{aligned} b_{CS} = & 1 - \left((\partial_\tau A_r(\tau) + A_r(\tau)\kappa_r - 1) \left(\left(\frac{A_r(\tau)}{\tau} - 1 \right) V_{rr} + \frac{A_\eta(\tau)}{\tau} V_{r\eta} \right) \right. \\ & \left. + (\partial_\tau A_\eta(\tau) + A_\eta(\tau)\kappa_\eta - A_r(\tau)\kappa_r) \left(\frac{A_\eta(\tau)}{\tau} V_{\eta\eta} + \left(\frac{A_r(\tau)}{\tau} - 1 \right) V_{r\eta} \right) \right) \end{aligned}$$

$$\left(\left(\frac{A_r(\tau)}{\tau} - 1 \right)^2 V_{rr} + \left(\frac{A_\eta(\tau)}{\tau} \right)^2 V_{\eta\eta} + 2 \left(\frac{A_r(\tau)}{\tau} - 1 \right) \left(\frac{A_\eta(\tau)}{\tau} \right) V_{r\eta} \right)^{-1}.$$

First, the denominator is positive because of an analogous reasoning to the one before. Second, consider the numerator. For $\tau < \min\{\bar{\tau}, \hat{\tau}_1\}$, $A_r(\tau) > 0$, $A_\eta(\tau) > 0$, $\partial_\tau A_\eta(\tau) > 0$, and $\partial_\tau A_r(\tau) + A_r(\tau)\kappa_r - 1 < 0$ and $\partial_\tau A_\eta(\tau) + A_\eta(\tau)\kappa_\eta - A_r(\tau)\kappa_r < 0$. Moreover, let $g(\tau) = A_r(\tau) - \tau$. The initial condition $A_r(0) = 0$ implies that $g(0) = 0$, and the condition $\partial_\tau A_r(\tau) - 1 < 0$ obtained above implies that $\partial_\tau g(\tau) = \partial_\tau A_r(\tau) - 1 < 0$. It follows that $A_r(\tau) - \tau < 0$. Therefore, a sufficient condition for the numerator to be positive is that the second leading coefficient of each term is negative:

$$\begin{aligned} \left(\frac{A_r(\tau)}{\tau} - 1 \right) V_{rr} + \frac{A_\eta(\tau)}{\tau} V_{r\eta} &< 0 \\ -\frac{A_\eta(\tau)}{A_r(\tau) - \tau} &< \frac{V_{rr}}{V_{r\eta}}, \end{aligned}$$

$$\begin{aligned} \frac{A_\eta(\tau)}{\tau} V_{\eta\eta} + \left(\frac{A_r(\tau)}{\tau} - 1 \right) V_{r\eta} &< 0 \\ -\frac{A_\eta(\tau)}{A_r(\tau) - \tau} &< \frac{V_{r\eta}}{V_{\eta\eta}}. \end{aligned}$$

Both of these conditions hold for $\tau < \min\{\bar{\tau}, \hat{\tau}_1\}$ if:

$$-\frac{A_\eta(\tau)}{A_r(\tau) - \tau} < \min \left\{ \frac{V_{rr}}{V_{r\eta}}, \frac{V_{r\eta}}{V_{\eta\eta}} \right\} = \frac{\kappa_r}{\kappa_r + \kappa_\eta}.$$

When $a \approx 0$, the solutions for $A_r(\tau)$ and $A_\eta(\tau)$ and $(v_1, v_2, \phi_r, \phi_\eta) = (\kappa_r, \kappa_\eta, a_{\underline{c}_r}, a_{\underline{c}_\eta})$ imply:

$$b_{CS} = 1 - \frac{N_{CS,1} + N_{CS,2}}{D_{CS}} + \mathcal{O}(a^2),$$

where:

$$\begin{aligned} N_{CS,1} &= \left(-a\sigma_r^2 \frac{1 - e^{-\tau\kappa_r}}{\kappa_r} F(\kappa_r, \kappa_r) - a\sigma_\eta^2 \frac{1 - e^{-\tau\kappa_\eta}}{\kappa_\eta} \left(\frac{1 - e^{-\tau\kappa_\eta}}{\kappa_\eta} - \frac{1 - e^{-\tau\kappa_r}}{\kappa_r} \right) \hat{F}(\kappa_r, \kappa_\eta) \right) \\ &\quad \left(\left(\frac{1 - e^{-\tau\kappa_r}}{\tau\kappa_r} - 1 \right) V_{rr} + \frac{1 - e^{-\tau\kappa_\eta}}{\tau\kappa_\eta} \left(\frac{1 - e^{-\tau\kappa_\eta}}{\tau\kappa_\eta} - \frac{1 - e^{-\tau\kappa_r}}{\tau\kappa_r} \right) V_{r\eta} \right) + \mathcal{O}(a^2), \end{aligned}$$

$$N_{CS,2} = \left(-a\sigma_r^2 \underline{c}_\eta \frac{1-e^{-\tau\kappa_r}}{\kappa_r} \hat{F}(\kappa_r, \kappa_\eta) - a\sigma_\eta^2 \underline{c}_\eta^3 \left(\frac{1-e^{-\tau\kappa_\eta}}{\kappa_\eta} - \frac{1-e^{-\tau\kappa_r}}{\kappa_r} \right) \hat{F}(\kappa_r, \kappa_\eta) \right) \\ \left(\frac{\underline{c}_\eta}{\tau\kappa_\eta} \left(\frac{1-e^{-\tau\kappa_\eta}}{\tau\kappa_\eta} - \frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} \right) V_{\eta\eta} + \left(\frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} - 1 \right) V_{r\eta} \right) + \mathcal{O}(a^2),$$

$$D_{CS} = \left(\left(\frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} - 1 \right)^2 + 2a\underline{c}_r \left(\frac{1-e^{-\tau\kappa_\eta}}{\tau\kappa_\eta} - \frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} \right) \left(\frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} - 1 \right) \right) V_{rr} \\ + 2 \left(a\underline{c}_r \underline{c}_\eta \left(\frac{1-e^{-\tau\kappa_\eta}}{\tau\kappa_\eta} - \frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} \right)^2 + \underline{c}_\eta \left(\frac{1-e^{-\tau\kappa_\eta}}{\tau\kappa_\eta} - \frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} \right) \left(\frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} - 1 \right) \right) V_{r\eta} \\ + \underline{c}_\eta^2 \left(\frac{1-e^{-\tau\kappa_\eta}}{\tau\kappa_\eta} - \frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} \right)^2 V_{\eta\eta} + \mathcal{O}(a^2).$$

First, the denominator is positive because of an analogous reasoning to the one before. Second, consider the numerator. The leading coefficient of each of the two terms in the numerator is negative. Therefore, a sufficient condition for the numerator to be positive is that the second leading coefficient of each term is negative:

$$\left(\frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} - 1 \right) V_{rr} + \underline{c}_\eta \left(\frac{1-e^{-\tau\kappa_\eta}}{\tau\kappa_\eta} - \frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} \right) V_{r\eta} < 0 \\ - \frac{\underline{c}_\eta \left(\frac{1-e^{-\tau\kappa_\eta}}{\tau\kappa_\eta} - \frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} \right)}{\frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} - 1} < \frac{V_{rr}}{V_{r\eta}},$$

$$\underline{c}_\eta \left(\frac{1-e^{-\tau\kappa_\eta}}{\tau\kappa_\eta} - \frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} \right) V_{\eta\eta} + \left(\frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} - 1 \right) V_{r\eta} < 0 \\ \underline{c}_\eta \left(\frac{1-e^{-\tau\kappa_\eta}}{\tau\kappa_\eta} - \frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} \right) V_{\eta\eta} < - \left(\frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} - 1 \right) V_{r\eta} \\ - \frac{\underline{c}_\eta \left(\frac{1-e^{-\tau\kappa_\eta}}{\tau\kappa_\eta} - \frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} \right)}{\frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} - 1} < \frac{V_{r\eta}}{V_{\eta\eta}}.$$

Both of these conditions hold if:

$$- \frac{\underline{c}_\eta \left(\frac{1-e^{-\tau\kappa_\eta}}{\tau\kappa_\eta} - \frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} \right)}{\frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r} - 1} < \min \left\{ \frac{V_{rr}}{V_{r\eta}}, \frac{V_{r\eta}}{V_{\eta\eta}} \right\} = \frac{\kappa_r}{\kappa_r + \kappa_\eta}.$$

Substituting the value of $\underline{c}_\eta$, the condition holds for τ such that:

$$\frac{\frac{1-e^{-\tau\kappa_\eta}}{\tau\kappa_\eta} - \frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r}}{1 - \frac{1-e^{-\tau\kappa_r}}{\tau\kappa_r}} < \frac{\kappa_r - \kappa_\eta}{\kappa_r + \kappa_\eta}.$$

A.2 Overview of the Expectations Hypothesis

The expectations hypothesis of the term structure has been at the center of empirical and theoretical literature on fixed income. It refers to a series of statements that link bond returns, yields, and forward rates. Initially, these statements were just hypothesized, rather than formalized with a fully specified equilibrium model, and hence the expectations hypothesis is not associated with a unique mathematical statement (see [Sangvinatsos \(2010\)](#)). Additionally, the literature has historically distinguished between two different versions: the pure expectations hypothesis and the expectations hypothesis.

The primitive objectives of these hypotheses are two types of risk premia: the instantaneous holding premium, which is defined as the difference between the instantaneous holding return on a τ -year bond and the current short-rate:

$$\psi_{h,t}(\tau) \equiv \mu_t(\tau) - r_t,$$

and the instantaneous forward premium, which is defined as the difference between the instantaneous forward rate and the expected future short-rate:

$$\psi_{f,t}(\tau) \equiv f_t(\tau) - \mathbb{E}_t(r_{t+\tau}).$$

The first version, discussed by [Cox et al. \(1981\)](#), is the pure expectations hypothesis, which states that risk premia are zero. The first specification, $\psi_{h,t}(\tau) = 0 \forall \tau$, states that all bonds have the same expected infinitesimal return, equal to the short-rate. The second specification, $\psi_{f,t}(\tau) = 0 \forall \tau$, states that the instantaneous forward rate is equal to the expected future short-rate. The second version, discussed by [Campbell \(1986\)](#), is the expectations hypothesis, which states that risk premia are constant over time and depend only on the maturity of the bond. The first specification is $\psi_{h,t}(\tau) = g_h(\tau) \forall \tau$ and for a function $g_h(\cdot)$. The second specification is $\psi_{f,t}(\tau) = g_f(\tau) \forall \tau$ and for a function $g_f(\cdot)$.⁹

⁹A third specification is given in terms of the yield premium, which is defined as the difference between the gross

The different specifications of the pure expectations hypothesis are inconsistent with one another because, when interest rates are stochastic, the different risk premia cannot simultaneously be zero due to Jensen's inequality. Moreover, the specification $\psi_{f,t}(\tau) = 0$ is incompatible with the equilibrium in a general arbitrage pricing framework with stochastic interest rates. On the other hand, the different specifications of the expectations hypothesis are not incompatible with one another or with the arbitrage pricing equilibrium. In any case, the differences between the risk premia in the pure expectations hypothesis are of second-order effects of bond yield variability, and when bond yields are not "too volatile", the inconsistencies can be neglected. Nonetheless, given the importance of volatility in this model, I choose to focus on the less restrictive expectations hypothesis.

A.3 Calibration

First, the yield of the τ -year bond is:

$$y_t(\tau) = \frac{A_r(\tau)r_t + A_\eta(\tau)\eta_t + C(\tau)}{\tau}.$$

Therefore, the standard deviation of the yield is:

$$\sqrt{\text{Var}(y_t(\tau))} = \frac{\sqrt{A_r^2(\tau)V_{rr} + A_\eta^2(\tau)V_{\eta\eta} + 2A_r(\tau)A_\eta(\tau)V_{r\eta}}}{\tau}.$$

It follows that the moment to calibrate κ_r is:

$$\sqrt{\text{Var}(y_t(1))} = \sqrt{A_r^2(1)V_{rr} + A_\eta^2(1)V_{\eta\eta} + 2A_r(1)A_\eta(1)V_{r\eta}}.$$

Second, the change in the yield of a constant-maturity τ -year bond over a time interval $\Delta\tau$ is:

$$y_{t+\Delta\tau}(\tau) - y_t(\tau) = \frac{A_r(\tau)(r_{t+\Delta\tau} - r_t) + A_\eta(\tau)(\eta_{t+\Delta\tau} - \eta_t)}{\tau}.$$

uncompounded return from holding a bond to maturity and the expected return from receiving the spot rate at each instant, $\psi_{y,t}(\tau) = \frac{1}{P_t(\tau)} - \mathbb{E}_t(e^{\int_t^{t+\tau} r_s ds})$. I do not discuss this further as it seems more intuitive to consider a rate of return over an infinitesimal interval of time.

Therefore, the standard deviation of the change in the yield is:

$$\sqrt{\mathbb{V}\text{ar}(y_{t+\Delta\tau}(\tau) - y_t(\tau))} = \frac{\sqrt{A_r^2(\tau)\mathbb{V}\text{ar}(\Delta r_t) + A_\eta^2(\tau)\mathbb{V}\text{ar}(\Delta \eta_t) + 2A_r(\tau)A_\eta(\tau)\mathbb{C}\text{ov}(\Delta r_t, \Delta \eta_t)}}{\tau},$$

where:

$$\begin{aligned}\mathbb{V}\text{ar}(\Delta r_t) &\equiv \mathbb{V}\text{ar}(r_{t+\Delta\tau} - r_t) \\ &= \mathbb{V}\text{ar}(r_t) + \mathbb{V}\text{ar}(r_{t+\Delta\tau}) - 2\mathbb{C}\text{ov}(r_t, r_{t+\Delta\tau}) \\ &= 2V_{rr} - 2C_{rr},\end{aligned}$$

$$\begin{aligned}\mathbb{V}\text{ar}(\Delta \eta_t) &\equiv \mathbb{V}\text{ar}(\eta_{t+\Delta\tau} - \eta_t) \\ &= \mathbb{V}\text{ar}(\eta_t) + \mathbb{V}\text{ar}(\eta_{t+\Delta\tau}) - 2\mathbb{C}\text{ov}(\eta_t, \eta_{t+\Delta\tau}) \\ &= 2V_{\eta\eta} - 2C_{\eta\eta},\end{aligned}$$

$$\begin{aligned}\mathbb{C}\text{ov}(\Delta r_t, \Delta \eta_t) &\equiv \mathbb{C}\text{ov}(r_{t+\Delta\tau} - r_t, \eta_{t+\Delta\tau} - \eta_t) \\ &= \mathbb{C}\text{ov}(r_t, \eta_t) + \mathbb{C}\text{ov}(r_{t+\Delta\tau}, \eta_{t+\Delta\tau}) - \mathbb{C}\text{ov}(r_t, \eta_{t+\Delta\tau}) - \mathbb{C}\text{ov}(\eta_t, r_{t+\Delta\tau}) \\ &= 2V_{r\eta} - C_{r\eta} - C_{\eta r}.\end{aligned}$$

It follows that the moment to calibrate σ_r is:

$$\sqrt{\mathbb{V}\text{ar}(y_{t+1}(1) - y_t(1))} = \sqrt{A_r^2(1)\mathbb{V}\text{ar}(\Delta r_t) + A_\eta^2(1)\mathbb{V}\text{ar}(\Delta \eta_t) + 2A_r(1)A_\eta(1)\mathbb{C}\text{ov}(\Delta r_t, \Delta \eta_t)},$$

Third, the volume transacted in a τ -year bond over an infinitesimal time interval dt is the absolute value of the change in the supply by preferred-habitat investors:

$$\begin{aligned}dS_t(\tau) &= d(\alpha_0(\tau) + \beta(\tau)p_t(\tau)) \\ &= d(\alpha_0(\tau) - \beta(\tau)(A_r(\tau)r_t + A_\eta(\tau)\eta_t + C(\omega, \tau))).\end{aligned}$$

The expected volume is:

$$\mathbb{E}(|dS_t(\tau)|) = \mathbb{E}(\mathbb{E}_t(|dS_t(\tau)|))$$

$$\begin{aligned}
&= \mathbb{E} \left(\sqrt{\frac{2}{\pi} \text{Var}(dS_t(\tau))} \right) \\
&= \mathbb{E} \left(\sqrt{\frac{2}{\pi} \text{Var}_t(-\beta(\tau) (A_r(\tau)\sigma_r dz_{r,t} + A_\eta(\tau)\sigma_\eta dz_{\eta,t}))} \right) \\
&= \mathbb{E} \left(\sqrt{\frac{2}{\pi} \beta^2(\tau) (A_r^2(\tau)\sigma_r^2 \text{Var}_t(dz_{r,t}) + A_\eta^2(\tau)\sigma_\eta^2 \text{Var}_t(dz_{\eta,t}))} \right) \\
&= \mathbb{E} \left(\sqrt{\frac{2}{\pi} \beta^2(\tau) (A_r^2(\tau)\sigma_r^2 + A_\eta^2(\tau)\sigma_\eta^2) dt} \right) \\
&= \sqrt{\frac{2}{\pi} \beta^2 e^{-2\tau\delta_\beta} (A_r^2(\tau)\sigma_r^2 + A_\eta^2(\tau)\sigma_\eta^2) dt},
\end{aligned}$$

where the second step follows from: (i) conditional on time t , $dS_t(\tau)$ is a linear combination of independent Brownian increments, and so $dS_t(\tau)$ is normally distributed; (ii) the drift coefficient is of order dt while the diffusion coefficient is of order $\sqrt{dt}$ (or equivalently the variance is of order dt), and so $dS_t(\tau)$ has mean zero and variance $\text{Var}(dS_t(\tau))$; (iii) the formula for the conditional expected value of the half-normal distribution. The fifth step follows from the fact that $dz_{r,t}$ and $dz_{\eta,t}$ are independent, and $\text{Var}_t(dz_t) = dt$. It follows that the moment to calibrate δ_β is:

$$\frac{\sum_{0 < \tau \leq 2} \mathbb{E}(|dS_t(\tau)|)}{\sum_{0 < \tau \leq 30} \mathbb{E}(|dS_t(\tau)|)} = \frac{\sum_{0 < \tau \leq 2} \sqrt{\beta^2 e^{-2\tau\delta_\beta} (A_r^2(\tau)\sigma_r^2 + A_\eta^2(\tau)\sigma_\eta^2)}}{\sum_{0 < \tau \leq 30} \sqrt{\beta^2 e^{-2\tau\delta_\beta} (A_r^2(\tau)\sigma_r^2 + A_\eta^2(\tau)\sigma_\eta^2)}}$$

Fourth, the expected yield of the τ -year bond is:

$$\mathbb{E}(y_t(\tau)) = \frac{(A_r(\tau) + A_\eta(\tau)) r^* + C(\tau)}{\tau}.$$

It follows that the moments to calibrate δ_α , $a\alpha_0$ and r^* are:

$$\mathbb{E}(y_t(1)) = (A_r(1) + A_\eta(1)) r^* + C(1),$$

$$\mathbb{E}(y_t(5)) = \frac{(A_r(5) + A_\eta(5)) r^* + C(5)}{5},$$

$$\mathbb{E}(y_t(7)) = \frac{(A_r(7) + A_\eta(7)) r^* + C(7)}{7}.$$

Fifth, the expected realised excess return of the τ -year bond over $\Delta\tau$ is:

$$\begin{aligned} \mathbb{E} \left(\frac{p_{t+\Delta\tau}(\tau - \Delta\tau) - p_t(\tau)}{\Delta\tau} - y_t(\Delta\tau) \right) &= \frac{(A_r(\tau) + A_\eta(\tau)) r^* + C(\tau)}{\Delta\tau} \\ &\quad - \frac{(A_r(\tau - \Delta\tau) + A_\eta(\tau - \Delta\tau)) r^* + C(\tau - \Delta\tau)}{\Delta\tau} \\ &\quad - \frac{(A_r(\Delta\tau) + A_\eta(\Delta\tau)) r^* + C(\Delta\tau)}{\Delta\tau}. \end{aligned}$$

It follows that the moment to calibrate $a\beta$ is:

$$\begin{aligned} \mathbb{E}(p_{t+1}(6) - p_t(7) - y_t(1)) &= (A_r(7) + A_\eta(7)) r^* + C(7) \\ &\quad - ((A_r(6) + A_\eta(6)) r^* + C(6)) \\ &\quad - ((A_r(1) + A_\eta(1)) r^* + C(1)). \end{aligned}$$

Finally, when the yields for all maturities change by Δy , the supply by preferred-habitat investors changes by:

$$\int_0^\infty \Delta S_t(\tau) d\tau = -\Delta y \beta \frac{1 - e^{-T\delta_\beta}(1 + T\delta_\beta)}{\delta_\beta^2}.$$

It follows that:

$$\beta = -\frac{\delta_\beta^2 \int_0^\infty \Delta S_t(\tau) d\tau}{\Delta y (1 - e^{-T\delta_\beta}(1 + T\delta_\beta))}.$$

Setting $\Delta y = 0.0001$ yields the estimate for β .

A.4 Policy Analysis

Consider an unanticipated change $\Delta g e^{-\kappa_g t}$ in the target short-rate that takes place at $t = 0$ and mean reverts deterministically to zero at the rate κ_g . I now follow the same steps as in Section 1.2.4.

(Step 1) Let the time- t price of the τ -year bond be:

$$P_t(\tau) = e^{-(A_r(\tau)r_t + A_\eta(\tau)\eta_t + A_g(\tau)\Delta g e^{-\kappa_g t} + C(\tau))}.$$

(Step 2) The instantaneous return of the τ -year bond is:

$$dP_t(\tau) = (\partial_t P_t(\tau) - \partial_\tau P_t(\tau)) dt + \partial_{r_t} P_t(\tau) dr_t + \frac{1}{2} \partial_{r_t}^2 P_t(\tau) dr_t^2 + \partial_{\eta_t} P_t(\tau) d\eta_t + \frac{1}{2} \partial_{\eta_t}^2 P_t(\tau) d\eta_t^2$$

$$\frac{dP_t(\tau)}{P_t(\tau)} = \mu_t(\tau) dt - A_r(\tau) \sigma_r dz_{r,t} - A_\eta(\tau) \sigma_\eta dz_{\eta,t},$$

where the instantaneous expected return is:

$$\begin{aligned} \mu_t(\tau) = & \kappa_g A_g(\tau) \Delta g e^{-\kappa_g t} + \partial_\tau A_r(\tau) r_t + \partial_\tau A_\eta(\tau) \eta_t + \partial_\tau A_g(\tau) \Delta g e^{-\kappa_g t} + \partial_\tau C(\tau) \\ & - A_r(\tau) \kappa_r (\eta_t + \Delta g e^{-\kappa_g t} - r_t) - A_\eta(\tau) \kappa_\eta (r^* - \eta_t) + \frac{1}{2} A_r^2(\tau) \sigma_r^2 + \frac{1}{2} A_\eta^2(\tau) \sigma_\eta^2. \end{aligned}$$

The arbitrageurs' first-order condition is:

$$\mu_t(\tau) - r_t = A_r(\tau) \lambda_{r,t} + A_\eta(\tau) \lambda_{\eta,t},$$

where the prices of risk of the current and target short-rate risk factors are:

$$\lambda_{f,t} = a \sigma_f^2 \int_0^\infty X_t(\tau) A_f(\tau) d\tau \quad \text{for } f \in \{r, \eta\}.$$

(Step 3) The prices of risk of the current and target short-rate risk factors are:

$$\begin{aligned} \lambda_{f,t} &= a \sigma_f^2 \int_0^\infty (\alpha_0(\tau) + \beta(\tau) p_t(\tau)) A_f(\tau) d\tau \\ &= a \sigma_f^2 \int_0^\infty (\alpha_0(\tau) - \beta(\tau) (A_r(\tau) r_t + A_\eta(\tau) \eta_t + A_g(\tau) \Delta g e^{-\kappa_g t} + C(\tau))) A_f(\tau) d\tau \\ &= \lambda_f + \lambda_{fg} \Delta g e^{-\kappa_g t} + \lambda_{fr} r_t + \lambda_{f\eta} \eta_t, \end{aligned}$$

where:

$$\lambda_f = a \sigma_f^2 \int_0^\infty (\alpha_0(\tau) - \beta(\tau) C(\tau)) A_f(\tau) d\tau,$$

$$\lambda_{fg} = -a \sigma_f^2 \int_0^\infty \beta(\tau) A_g(\tau) A_f(\tau) d\tau,$$

$$\lambda_{fr} = -a \sigma_f^2 \int_0^\infty \beta(\tau) A_r(\tau) A_f(\tau) d\tau,$$

$$\lambda_{f\eta} = -a\sigma_f^2 \int_0^\infty \beta(\tau)A_\eta(\tau)A_f(\tau)d\tau.$$

Substituting $\mu_t(\tau)$, $\lambda_{r,t}$ and $\lambda_{\eta,t}$ into the first-order condition, I find an affine equation in r_t , η_t and $\Delta g e^{-\kappa_s t}$:

$$\begin{aligned} & \kappa_g A_g(\tau) \Delta g e^{-\kappa_s t} + \partial_\tau A_r(\tau) r_t + \partial_\tau A_\eta(\tau) \eta_t + \partial_\tau A_g(\tau) \Delta g e^{-\kappa_s t} + \partial_\tau C(\tau) \\ & - A_r(\tau) \kappa_r (\eta_t + \Delta g e^{-\kappa_s t} - r_t) - A_\eta(\tau) \kappa_\eta (r^* - \eta_t) + \frac{1}{2} A_r^2(\tau) \sigma_r^2 + \frac{1}{2} A_\eta^2(\tau) \sigma_\eta^2 - r_t \\ & = A_r(\tau) (\lambda_r + \lambda_{rg} e^{-\kappa_s t} + \lambda_{rr} r_t + \lambda_{r\eta} \eta_t) + A_\eta(\tau) (\lambda_\eta + \lambda_{\eta g} e^{-\kappa_s t} + \lambda_{\eta r} r_t + \lambda_{\eta\eta} \eta_t) \end{aligned}$$

Identifying the linear terms in $\Delta g e^{-\kappa_s t}$, r_t and η_t , and the terms independent of $\Delta g e^{-\kappa_s t}$, r_t and η_t , I find a system of ODEs:

$$\partial_\tau A_g(\tau) + A_g(\tau) \kappa_g - A_r(\tau) (\kappa_r + \lambda_{rg}) - A_\eta(\tau) \lambda_{\eta g} = 0$$

$$\partial_\tau A_r(\tau) + A_r(\tau) (\kappa_r - \lambda_{rr}) - A_\eta(\tau) \lambda_{\eta r} - 1 = 0$$

$$\partial_\tau A_\eta(\tau) + A_\eta(\tau) (\kappa_\eta - \lambda_{\eta\eta}) - A_r(\tau) (\kappa_r + \lambda_{r\eta}) = 0$$

$$\partial_\tau C(\tau) - A_r(\tau) \lambda_r + \frac{1}{2} A_r^2(\tau) \sigma_r^2 - A_\eta(\tau) (\kappa_\eta r^* + \lambda_\eta) + \frac{1}{2} A_\eta^2(\tau) \sigma_\eta^2 = 0$$

Now, I follow the same steps as in Proof of Proposition 1.5 (see Appendix A.1.5).

(Step 1) I take the integrals as given, and solve the first differential equation by simple integration:

$$A_g(\tau) = \lambda_{rg}^* \int_0^\tau e^{-(\tau-u)\kappa_g} A_r(u) du + \lambda_{\eta g} \int_0^\tau e^{-(\tau-u)\kappa_g} A_\eta(u) du.$$

where I defined $\lambda_{rg}^* \equiv \kappa_r + \lambda_{rg}$.

(Step 2) Substituting $A_g(\tau)$ into λ_{rg} and $\lambda_{\eta g}$:

$$\begin{aligned} \lambda_{rg} &= -a\sigma_r^2 \int_0^\infty \beta(\tau) A_g(\tau) A_r(\tau) d\tau \\ \lambda_{rg}^* - \kappa_r &= -a\sigma_r^2 \left(\lambda_{rg}^* M_{g,rr} + \lambda_{\eta g} M_{g,r\eta} \right), \end{aligned}$$

$$\lambda_{\eta g} = -a\sigma_\eta^2 \int_0^\infty \beta(\tau) A_g(\tau) A_\eta(\tau) d\tau$$

$$\lambda_{\eta g} = -a\sigma_\eta^2 \left(\lambda_{rg}^* M_{g,\eta r} + \lambda_{\eta g} M_{g,\eta\eta} \right),$$

where I defined:

$$M_{g,ij} \equiv \int_0^\infty \beta(\tau) A_i(\tau) \left(\int_0^\tau e^{-(\tau-u)\kappa_g} A_j(u) du \right) d\tau.$$

Writing the system in vector-matrix form:

$$\begin{bmatrix} 1 + a\sigma_r^2 M_{g,rr} & a\sigma_r^2 M_{g,r\eta} \\ a\sigma_\eta^2 M_{g,\eta r} & 1 + a\sigma_\eta^2 M_{g,\eta\eta} \end{bmatrix} \begin{bmatrix} \lambda_{rg}^* \\ \lambda_{\eta g} \end{bmatrix} = \begin{bmatrix} \kappa_r \\ 0 \end{bmatrix},$$

Solving the system yields:

$$\lambda_{rg}^* = \frac{\left(1 + a\sigma_\eta^2 M_{g,\eta\eta}\right) \kappa_r}{D_g},$$

$$\lambda_{\eta g} = \frac{-a\sigma_\eta^2 M_{g,\eta r} \kappa_r}{D_g},$$

where:

$$D_g = \left(1 + a\sigma_r^2 M_{g,rr}\right) \left(1 + a\sigma_\eta^2 M_{g,\eta\eta}\right) - a^2 \sigma_r^2 \sigma_\eta^2 M_{g,r\eta} M_{g,\eta r}.$$

Moreover, when $a = 0$:

$$A_g(\tau) = \frac{(1 - e^{-\tau\kappa_r}) \kappa_g - (1 - e^{-\tau\kappa_g}) \kappa_r}{(\kappa_g - \kappa_r) \kappa_g},$$

$$A_r(\tau) = \frac{1 - e^{-\tau\kappa_r}}{\kappa_r},$$

$$A_\eta(\tau) = \frac{(1 - e^{-\tau\kappa_\eta}) \kappa_r - (1 - e^{-\tau\kappa_r}) \kappa_\eta}{(\kappa_r - \kappa_\eta) \kappa_\eta}.$$

Finally, the time zero change in the yield is:

$$\Delta y_t(\tau) = \frac{A_g(\tau)}{\tau} \Delta g$$

Next, consider an unanticipated change $\Delta\alpha_0(\tau)$ in the intercept of the supply that takes place at $t = 0$ and mean reverts deterministically to zero at the rate κ_α . I now follow the same steps as in Section 1.2.4.

(Step 1) Let the time- t price of the τ -year bond be:

$$P_t(\tau) = e^{-(A_r(\tau)r_t + A_\eta(\tau)\eta_t + A_\alpha(\tau)\Delta\alpha_0(\tau)e^{-\kappa_\alpha t} + C(\tau))}.$$

(Step 2) The instantaneous return of the τ -year bond is:

$$\begin{aligned} dP_t(\tau) &= (\partial_t P_t(\tau) - \partial_\tau P_t(\tau)) dt + \partial_{r_t} P_t(\tau) dr_t + \frac{1}{2} \partial_{r_t}^2 P_t(\tau) dr_t^2 + \partial_{\eta_t} P_t(\tau) d\eta_t + \frac{1}{2} \partial_{\eta_t}^2 P_t(\tau) d\eta_t^2 \\ \frac{dP_t(\tau)}{P_t(\tau)} &= \mu_t(\tau) dt - A_r(\tau) \sigma_r dz_{r,t} - A_\eta(\tau) \sigma_\eta dz_{\eta,t}, \end{aligned}$$

where the instantaneous expected return is:

$$\begin{aligned} \mu_t(\tau) &= A_\alpha(\tau) \Delta\alpha_0(\tau) \kappa_\alpha e^{-\kappa_\alpha t} + \partial_\tau A_r(\tau) r_t + \partial_\tau A_\eta(\tau) \eta_t + (\partial_\tau A_\alpha(\tau) \Delta\alpha_0(\tau) + A_\alpha(\tau) \partial_\tau \Delta\alpha_0(\tau)) e^{-\kappa_\alpha t} \\ &\quad + \partial_\tau C(\tau) - A_r(\tau) \kappa_r (\eta_t - r_t) - A_\eta(\tau) \kappa_\eta (r^* - \eta_t) + \frac{1}{2} A_r^2(\tau) \sigma_r^2 + \frac{1}{2} A_\eta^2(\tau) \sigma_\eta^2. \end{aligned}$$

The arbitrageurs' first-order condition is:

$$\mu_t(\tau) - r_t = A_r(\tau) \lambda_{r,t} + A_\eta(\tau) \lambda_{\eta,t},$$

where the prices of risk of the current and target short-rate risk factors are:

$$\lambda_{f,t} = a \sigma_f^2 \int_0^\infty X_t(\tau) A_f(\tau) d\tau \quad \text{for } f \in \{r, \eta\}.$$

(Step 3) The prices of risk of the current and target short-rate risk factors are:

$$\begin{aligned} \lambda_{f,t} &= a \sigma_f^2 \int_0^\infty (\alpha_0(\tau) + \Delta\alpha_0(\tau) e^{-\kappa_\alpha t} + \beta(\tau) p_t(\tau)) A_f(\tau) d\tau \\ &= a \sigma_f^2 \int_0^\infty (\alpha_0(\tau) + \Delta\alpha_0(\tau) e^{-\kappa_\alpha t} - \beta(\tau) (A_r(\tau) r_t + A_\eta(\tau) \eta_t + \end{aligned}$$

$$\begin{aligned}
& A_\alpha(\tau)\Delta\alpha_0(\tau)e^{-\kappa_\alpha t} + C(\tau)) A_f(\tau)d\tau \\
& = \lambda_f + \lambda_{f\alpha}e^{-\kappa_\alpha t} + \lambda_{fr}r_t + \lambda_{f\eta}\eta_t,
\end{aligned}$$

where:

$$\lambda_f = a\sigma_f^2 \int_0^\infty (\alpha_0(\tau) - \beta(\tau)C(\tau)) A_f(\tau)d\tau,$$

$$\lambda_{f\alpha} = a\sigma_f^2 \int_0^\infty (\Delta\alpha_0(\tau) - \beta(\tau)A_\alpha(\tau)\Delta\alpha_0(\tau)) A_f(\tau)d\tau,$$

$$\lambda_{fr} = -a\sigma_f^2 \int_0^\infty \beta(\tau)A_r(\tau)A_f(\tau)d\tau,$$

$$\lambda_{f\eta} = -a\sigma_f^2 \int_0^\infty \beta(\tau)A_\eta(\tau)A_f(\tau)d\tau.$$

Substituting $\mu_t(\tau)$, $\lambda_{r,t}$ and $\lambda_{\eta,t}$ into the first-order condition, I find an affine equation in r_t , η_t and $e^{-\kappa_\alpha t}$:

$$\begin{aligned}
& A_\alpha(\tau)\Delta\alpha_0(\tau)\kappa_\alpha e^{-\kappa_\alpha t} + \partial_\tau A_r(\tau)r_t + \partial_\tau A_\eta(\tau)\eta_t + (\partial_\tau A_\alpha(\tau)\Delta\alpha_0(\tau) + A_\alpha(\tau)\partial_\tau\Delta\alpha_0(\tau)) e^{-\kappa_\alpha t} + \partial_\tau C(\tau) \\
& - A_r(\tau)\kappa_r(\eta_t - r_t) - A_\eta(\tau)\kappa_\eta(r^* - \eta_t) + \frac{1}{2}A_r^2(\tau)\sigma_r^2 + \frac{1}{2}A_\eta^2(\tau)\sigma_\eta^2 - r_t \\
& = A_r(\tau) (\lambda_r + \lambda_{r\alpha}e^{-\kappa_\alpha t} + \lambda_{rr}r_t + \lambda_{r\eta}\eta_t) + A_\eta(\tau) (\lambda_\eta + \lambda_{\eta\alpha}e^{-\kappa_\alpha t} + \lambda_{\eta r}r_t + \lambda_{\eta\eta}\eta_t)
\end{aligned}$$

Identifying the linear terms in $e^{-\kappa_\alpha t}$, r_t and η_t , and the terms independent of $e^{-\kappa_\alpha t}$, r_t and η_t , I find a system of ODEs:

$$\partial_\tau B_\alpha(\tau) + B_\alpha(\tau)\kappa_\alpha - A_r(\tau)\lambda_{r\alpha} - A_\eta(\tau)\lambda_{\eta\alpha} = 0,$$

$$\partial_\tau A_r(\tau) + A_r(\tau)(\kappa_r - \lambda_{rr}) - A_\eta(\tau)\lambda_{\eta r} - 1 = 0,$$

$$\partial_\tau A_\eta(\tau) + A_\eta(\tau)(\kappa_\eta - \lambda_{\eta\eta}) - A_r(\tau)(\kappa_r + \lambda_{r\eta}) = 0,$$

$$\partial_\tau C(\tau) - A_r(\tau)\lambda_r + \frac{1}{2}A_r^2(\tau)\sigma_r^2 - A_\eta(\tau)(\kappa_\eta r^* + \lambda_\eta) + \frac{1}{2}A_\eta^2(\tau)\sigma_\eta^2 = 0,$$

where I defined:

$$B_\alpha(\tau) \equiv A_\alpha(\tau)\Delta\alpha_0(\tau).$$

Now, I follow the same steps as in Proof of Proposition 1.5 (see Appendix A.1.5).

(Step 1) I take the integrals as given, and solve the first differential equation by simple integration:

$$B_\alpha(\tau) = \lambda_{r\alpha} \int_0^\tau e^{-(\tau-u)\kappa_\alpha} A_r(u) du + \lambda_{\eta\alpha} \int_0^\tau e^{-(\tau-u)\kappa_\alpha} A_\eta(u) du.$$

(Step 2) Substituting $B_\alpha(\tau)$ into $\lambda_{r\alpha}$ and $\lambda_{\eta\alpha}$:

$$\lambda_{r\alpha} = a\sigma_r^2 \int_0^\infty (\Delta\alpha_0(\tau) - \beta(\tau)B_\alpha(\tau)) A_r(\tau) d\tau$$

$$\lambda_{r\alpha} = K_{\alpha,r} - a\sigma_r^2 (\lambda_{r\alpha}M_{\alpha,rr} + \lambda_{\eta\alpha}M_{\alpha,r\eta}),$$

$$\lambda_{\eta\alpha} = a\sigma_\eta^2 \int_0^\infty (\Delta\alpha_0(\tau) - \beta(\tau)B_\alpha(\tau)) A_\eta(\tau) d\tau$$

$$\lambda_{\eta\alpha} = K_{\alpha,\eta} - a\sigma_\eta^2 (\lambda_{r\alpha}M_{\alpha,\eta r} + \lambda_{\eta\alpha}M_{\alpha,\eta\eta}),$$

where I defined:

$$K_{\alpha,j} \equiv a\sigma_j^2 \int_0^\infty \Delta\alpha_0(\tau) A_j(\tau) d\tau,$$

$$M_{\alpha,ij} \equiv \int_0^\infty \beta(\tau) A_i(\tau) \left(\int_0^\tau e^{-(\tau-u)\kappa_\alpha} A_j(u) du \right) d\tau.$$

Assuming that the change $\Delta\alpha_0(\tau)$ is a Dirac function with point mass at τ^* :

$$K_{\alpha,j} = a\sigma_j^2 A_j(\tau^*)$$

Writing the system in vector-matrix form:

$$\begin{bmatrix} 1 + a\sigma_r^2 M_{\alpha,rr} & a\sigma_r^2 M_{\alpha,r\eta} \\ a\sigma_\eta^2 M_{\alpha,\eta r} & 1 + a\sigma_\eta^2 M_{\alpha,\eta\eta} \end{bmatrix} \begin{bmatrix} \lambda_{r\alpha} \\ \lambda_{\eta\alpha} \end{bmatrix} = \begin{bmatrix} K_{\alpha,r} \\ K_{\alpha,\eta} \end{bmatrix},$$

Solving the system yields:

$$\lambda_{r\alpha} = \frac{(1 + a\sigma_\eta^2 M_{\alpha,\eta\eta}) K_{\alpha,r} - a\sigma_r^2 M_{\alpha,r\eta} K_{\alpha,\eta}}{D_\alpha},$$

$$\lambda_{\eta\alpha} = \frac{(1 + a\sigma_r^2 M_{\alpha,rr}) K_{\alpha,\eta} - a\sigma_\eta^2 M_{\alpha,\eta r} K_{\alpha,r}}{D_\alpha},$$

where:

$$D_\alpha = (1 + a\sigma_r^2 M_{\alpha,rr})(1 + a\sigma_\eta^2 M_{\alpha,\eta\eta}) - a^2 \sigma_r^2 \sigma_\eta^2 M_{\alpha,r\eta} M_{\alpha,\eta r}.$$

Finally, the time zero change in the yield is:

$$\Delta y_t(\tau) = \frac{A_\alpha(\tau)}{\tau} \Delta \alpha_0(\tau) = \frac{B_\alpha(\tau)}{\tau}$$

A.5 Model Implementation

In this section, I re-present the key formulas that make up the core solution of the model. While mathematically equivalent to those shown in the main text, they are written here in a form that emphasizes implementation rather than interpretation.

$(A_r(\tau), A_\eta(\tau))$ and Ψ

I re-write the solutions for $(A_r(\tau), A_\eta(\tau))$ from Proposition 1.5 as follows:

$$A_r(\tau) = c_{r,1} \frac{1 - e^{-\tau v_1}}{v_1} + c_{r,2} \frac{1 - e^{-\tau v_2}}{v_2},$$

$$A_\eta(\tau) = c_{\eta,1} \frac{1 - e^{-\tau v_1}}{v_1} + c_{\eta,2} \frac{1 - e^{-\tau v_2}}{v_2}.$$

where:

$$c_{r,1} \equiv 1 - \phi_r = -\frac{\lambda_2}{\lambda_1 - \lambda_2}, \quad c_{r,2} \equiv \phi_r = \frac{\lambda_1}{\lambda_1 - \lambda_2},$$

$$c_{\eta,1} \equiv -\phi_\eta = \frac{1}{\lambda_1 - \lambda_2}, \quad c_{\eta,2} \equiv \phi_\eta = -\frac{1}{\lambda_1 - \lambda_2}.$$

I re-write the elements of the matrix Ψ from the definition itself as follows:

$$\begin{aligned} \Psi_{r,r} &= \kappa_r + a\sigma_r^2 \int_0^T \beta(\tau) A_r^2(\tau) d\tau \\ &= \kappa_r + a\beta\sigma_r^2 \left(c_{r,1}^2 \mathcal{S}_1(v_1, v_1) + c_{r,2}^2 \mathcal{S}_1(v_2, v_2) + 2c_{r,1}c_{r,2} \mathcal{S}_1(v_1, v_2) \right), \end{aligned}$$

$$\begin{aligned} \Psi_{r,\eta} &= a\sigma_\eta^2 \int_0^T \beta(\tau) A_r(\tau) A_\eta(\tau) d\tau \\ &= a\beta\sigma_\eta^2 \left(c_{r,1}c_{\eta,1} \mathcal{S}_1(v_1, v_1) + c_{r,2}c_{\eta,2} \mathcal{S}_1(v_2, v_2) + (c_{r,1}c_{\eta,2} + c_{r,2}c_{\eta,1}) \mathcal{S}_1(v_1, v_2) \right), \end{aligned}$$

$$\begin{aligned} \Psi_{\eta,r} &= -\kappa_r + a\sigma_r^2 \int_0^T \beta(\tau) A_r(\tau) A_\eta(\tau) d\tau \\ &= -\kappa_r + a\beta\sigma_r^2 \left(c_{r,1}c_{\eta,1} \mathcal{S}_1(v_1, v_1) + c_{r,2}c_{\eta,2} \mathcal{S}_1(v_2, v_2) + (c_{r,1}c_{\eta,2} + c_{r,2}c_{\eta,1}) \mathcal{S}_1(v_1, v_2) \right), \end{aligned}$$

$$\begin{aligned} \Psi_{\eta,\eta} &= \kappa_\eta + a\sigma_\eta^2 \int_0^T \beta(\tau) A_\eta^2(\tau) d\tau \\ &= \kappa_\eta + a\beta\sigma_\eta^2 \left(c_{\eta,1}^2 \mathcal{S}_1(v_1, v_1) + c_{\eta,2}^2 \mathcal{S}_1(v_2, v_2) + 2c_{\eta,1}c_{\eta,2} \mathcal{S}_1(v_1, v_2) \right), \end{aligned}$$

where:

$$\begin{aligned} \mathcal{S}_1(x_1, x_2) &\equiv \int_0^T e^{-\tau\delta_\beta} \frac{1 - e^{-\tau x_1}}{x_1} \frac{1 - e^{-\tau x_2}}{x_2} d\tau \\ &= \frac{1}{x_1 x_2} \left(\frac{1 - e^{-T\delta_\beta}}{\delta_\beta} - \frac{1 - e^{-T(\delta_\beta + x_1)}}{\delta_\beta + x_1} - \frac{1 - e^{-T(\delta_\beta + x_2)}}{\delta_\beta + x_2} + \frac{1 - e^{-T(\delta_\beta + x_1 + x_2)}}{\delta_\beta + x_1 + x_2} \right). \end{aligned}$$

$C(\tau)$ and ψ

The elements of the vector $\mathbf{n}_0(\tau)$ can be written as:

$$\begin{aligned} n_{0,r}(\tau) &= \int_0^\tau A_r(u) du \\ &= c_{r,1} \mathcal{S}_2(\tau, v_1) + c_{r,2} \mathcal{S}_2(\tau, v_2), \end{aligned}$$

$$\begin{aligned} n_{0,\eta}(\tau) &= \int_0^\tau A_\eta(u) du \\ &= c_{\eta,1} \mathcal{S}_2(\tau, v_1) + c_{\eta,2} \mathcal{S}_2(\tau, v_2), \end{aligned}$$

where:

$$\begin{aligned} \mathcal{S}_2(\tau, x) &\equiv \int_0^\tau \frac{1 - e^{-ux}}{x} du \\ &= \frac{\tau}{x} - \frac{1 - e^{-\tau x}}{x^2}. \end{aligned}$$

The scalar $n(\tau)$ can be written as:

$$\begin{aligned} n(\tau) &= \sigma_r^2 \int_0^\tau A_r^2(u) du + \sigma_\eta^2 \int_0^\tau A_\eta^2(u) du \\ &= \left(\sigma_r^2 c_{r,1}^2 + \sigma_\eta^2 c_{\eta,1}^2 \right) \mathcal{S}_3(\tau, v_1, v_1) + \left(\sigma_r^2 c_{r,2}^2 + \sigma_\eta^2 c_{\eta,2}^2 \right) \mathcal{S}_3(\tau, v_2, v_2) \\ &\quad + 2 \left(\sigma_r^2 c_{r,1} c_{r,2} + \sigma_\eta^2 c_{\eta,1} c_{\eta,2} \right) \mathcal{S}_3(\tau, v_1, v_2), \end{aligned}$$

where:

$$\begin{aligned} \mathcal{S}_3(\tau, x_1, x_2) &\equiv \int_0^\tau \frac{1 - e^{-ux_1}}{x_1} \frac{1 - e^{-ux_2}}{x_2} du \\ &= \frac{1}{x_1 x_2} \left(\tau - \frac{1 - e^{-\tau x_1}}{x_1} - \frac{1 - e^{-\tau x_2}}{x_2} + \frac{1 - e^{-\tau(x_1+x_2)}}{x_1 + x_2} \right). \end{aligned}$$

The elements of the vector $\mathbf{n}_1$ can be written as follows (with the obvious definitions of $\tilde{n}_{1,r}$ and $\tilde{n}_{1,\eta}$):

$$\begin{aligned} n_{1,r} &\equiv \alpha_0 \tilde{n}_{1,r} = \int_0^T \alpha_0(\tau) A_r(\tau) d\tau \\ &= \alpha_0 \left(c_{r,1} \mathcal{S}_4(v_1) + c_{r,2} \mathcal{S}_4(v_2) \right), \end{aligned}$$

$$\begin{aligned}
n_{1,\eta} &\equiv \alpha_0 \tilde{n}_{1,\eta} = \int_0^T \alpha_0(\tau) A_\eta(\tau) d\tau \\
&= \alpha_0 \left(c_{\eta,1} \mathcal{S}_4(v_1) + c_{\eta,2} \mathcal{S}_4(v_2) \right),
\end{aligned}$$

where:

$$\begin{aligned}
\mathcal{S}_4(x) &\equiv \int_0^T \left(e^{-\tau\delta_\beta} - e^{-\tau\delta_\alpha} \right) \frac{1 - e^{-\tau x}}{x} d\tau \\
&= \frac{1}{x} \left(\frac{1 - e^{-T\delta_\beta}}{\delta_\beta} - \frac{1 - e^{-T(\delta_\beta+x)}}{\delta_\beta + x} - \frac{1 - e^{-T\delta_\alpha}}{\delta_\alpha} + \frac{1 - e^{-T(\delta_\alpha+x)}}{\delta_\alpha + x} \right).
\end{aligned}$$

The elements of the vector $\mathbf{n}_2$ can be written as follows (with the obvious definitions of $\tilde{n}_{2,r}$ and $\tilde{n}_{2,\eta}$):

$$\begin{aligned}
n_{2,r} &\equiv \beta \tilde{n}_{2,r} = \frac{1}{2} \int_0^T \beta(\tau) n(\tau) A_r(\tau) d\tau \\
&= \beta \left(\frac{1}{2} c_{r,1} \left(\left(\sigma_r^2 c_{r,1}^2 + \sigma_\eta^2 c_{\eta,1}^2 \right) \mathcal{S}_5(v_1, v_1, v_1) + \left(\sigma_r^2 c_{r,2}^2 + \sigma_\eta^2 c_{\eta,2}^2 \right) \mathcal{S}_5(v_2, v_2, v_1) \right. \right. \\
&\quad \left. \left. + 2 \left(\sigma_r^2 c_{r,1} c_{r,2} + \sigma_\eta^2 c_{\eta,1} c_{\eta,2} \right) \mathcal{S}_5(v_1, v_2, v_1) \right) \right. \\
&\quad \left. + \frac{1}{2} c_{r,2} \left(\left(\sigma_r^2 c_{r,1}^2 + \sigma_\eta^2 c_{\eta,1}^2 \right) \mathcal{S}_5(v_1, v_1, v_2) + \left(\sigma_r^2 c_{r,2}^2 + \sigma_\eta^2 c_{\eta,2}^2 \right) \mathcal{S}_5(v_2, v_2, v_2) \right. \right. \\
&\quad \left. \left. + 2 \left(\sigma_r^2 c_{r,1} c_{r,2} + \sigma_\eta^2 c_{\eta,1} c_{\eta,2} \right) \mathcal{S}_5(v_1, v_2, v_2) \right) \right),
\end{aligned}$$

$$\begin{aligned}
n_{2,\eta} &\equiv \beta \tilde{n}_{2,\eta} = \frac{1}{2} \int_0^T \beta(\tau) n(\tau) A_\eta(\tau) d\tau \\
&= \beta \left(\frac{1}{2} c_{\eta,1} \left(\left(\sigma_r^2 c_{r,1}^2 + \sigma_\eta^2 c_{\eta,1}^2 \right) \mathcal{S}_5(v_1, v_1, v_1) + \left(\sigma_r^2 c_{r,2}^2 + \sigma_\eta^2 c_{\eta,2}^2 \right) \mathcal{S}_5(v_2, v_2, v_1) \right. \right. \\
&\quad \left. \left. + 2 \left(\sigma_r^2 c_{r,1} c_{r,2} + \sigma_\eta^2 c_{\eta,1} c_{\eta,2} \right) \mathcal{S}_5(v_1, v_2, v_1) \right) \right. \\
&\quad \left. + \frac{1}{2} c_{\eta,2} \left(\left(\sigma_r^2 c_{r,1}^2 + \sigma_\eta^2 c_{\eta,1}^2 \right) \mathcal{S}_5(v_1, v_1, v_2) + \left(\sigma_r^2 c_{r,2}^2 + \sigma_\eta^2 c_{\eta,2}^2 \right) \mathcal{S}_5(v_2, v_2, v_2) \right. \right. \\
&\quad \left. \left. + 2 \left(\sigma_r^2 c_{r,1} c_{r,2} + \sigma_\eta^2 c_{\eta,1} c_{\eta,2} \right) \mathcal{S}_5(v_1, v_2, v_2) \right) \right),
\end{aligned}$$

where:

$$\begin{aligned}
\mathcal{S}_5(x_1, x_2, x_3) &\equiv \int_0^T e^{-\tau\delta_\beta} \left(\int_0^\tau \frac{1 - e^{-ux_1}}{x_1} \frac{1 - e^{-ux_2}}{x_2} du \right) \frac{1 - e^{-\tau x_3}}{x_3} d\tau \\
&= \frac{1}{x_1 x_2} \left(\mathcal{S}_6(x_3) - \mathcal{S}_1(x_1, x_3) - \mathcal{S}_1(x_2, x_3) + \mathcal{S}_1(x_1 + x_2, x_3) \right),
\end{aligned}$$

$$\begin{aligned}\mathcal{S}_6(x) &\equiv \int_0^T e^{-\tau\delta_\beta} \tau \frac{1 - e^{-\tau x}}{x} d\tau \\ &= \frac{1}{x} \left(\frac{1 - e^{-T\delta_\beta}(1 + T\delta_\beta)}{\delta_\beta^2} - \frac{1 - e^{-T(\delta_\beta+x)}(1 + T(\delta_\beta+x))}{(\delta_\beta+x)^2} \right).\end{aligned}$$

The elements of the matrix N can be written as follows (with the obvious definitions of $\tilde{n}_{r,r}$, $\tilde{n}_{r,\eta}$, $\tilde{n}_{\eta,r}$ and $\tilde{n}_{\eta,\eta}$):

$$\begin{aligned}n_{r,r} &\equiv \beta \tilde{n}_{r,r} = \int_0^T \beta(\tau) \left(\int_0^\tau A_r(u) du \right) A_r(\tau) d\tau \\ &= \beta \left(c_{r,1} (c_{r,1} \mathcal{S}_7(v_1, v_1) + c_{r,2} \mathcal{S}_7(v_1, v_2)) + c_{r,2} (c_{r,1} \mathcal{S}_7(v_2, v_1) + c_{r,2} \mathcal{S}_7(v_2, v_2)) \right),\end{aligned}$$

$$\begin{aligned}n_{r,\eta} &\equiv \beta \tilde{n}_{r,\eta} = \int_0^T \beta(\tau) \left(\int_0^\tau A_\eta(u) du \right) A_r(\tau) d\tau \\ &= \beta \left(c_{\eta,1} (c_{r,1} \mathcal{S}_7(v_1, v_1) + c_{r,2} \mathcal{S}_7(v_1, v_2)) + c_{\eta,2} (c_{r,1} \mathcal{S}_7(v_2, v_1) + c_{r,2} \mathcal{S}_7(v_2, v_2)) \right),\end{aligned}$$

$$\begin{aligned}n_{\eta,r} &\equiv \beta \tilde{n}_{\eta,r} = \int_0^T \beta(\tau) \left(\int_0^\tau A_r(u) du \right) A_\eta(\tau) d\tau \\ &= \beta \left(c_{r,1} (c_{\eta,1} \mathcal{S}_7(v_1, v_1) + c_{\eta,2} \mathcal{S}_7(v_1, v_2)) + c_{r,2} (c_{\eta,1} \mathcal{S}_7(v_2, v_1) + c_{\eta,2} \mathcal{S}_7(v_2, v_2)) \right),\end{aligned}$$

$$\begin{aligned}n_{\eta,\eta} &\equiv \beta \tilde{n}_{\eta,\eta} = \int_0^T \beta(\tau) \left(\int_0^\tau A_\eta(u) du \right) A_\eta(\tau) d\tau \\ &= \beta \left(c_{\eta,1} (c_{\eta,1} \mathcal{S}_7(v_1, v_1) + c_{\eta,2} \mathcal{S}_7(v_1, v_2)) + c_{\eta,2} (c_{\eta,1} \mathcal{S}_7(v_2, v_1) + c_{\eta,2} \mathcal{S}_7(v_2, v_2)) \right),\end{aligned}$$

where:

$$\begin{aligned}\mathcal{S}_7(x_1, x_2) &\equiv \int_0^T e^{-\tau\delta_\beta} \left(\int_0^\tau \frac{1 - e^{-ux_1}}{x_1} du \right) \frac{1 - e^{-\tau x_2}}{x_2} d\tau \\ &= \frac{1}{x_1} (\mathcal{S}_6(x_2) - \mathcal{S}_1(x_1, x_2)).\end{aligned}$$

These formulas fully define the elements of the vector ψ and hence the solution for $C(\tau)$ from Lemma 1.2:

$$\psi_r = \frac{1}{d} \left(\sigma_r^2 \left(1 + a\beta\sigma_\eta^2 \tilde{n}_{\eta,\eta} \right) (a\alpha_0 \tilde{n}_{1,r} + a\beta \tilde{n}_{2,r}) - \sigma_r^2 a\beta \tilde{n}_{r,\eta} \left(r^* \kappa_\eta + \sigma_\eta^2 (a\alpha_0 \tilde{n}_{1,\eta} + a\beta \tilde{n}_{2,\eta}) \right) \right),$$

$$\psi_\eta = \frac{1}{d} \left(-\sigma_r^2 \sigma_\eta^2 a \beta \tilde{n}_{\eta,r} (a \alpha_0 \tilde{n}_{1,r} + a \beta \tilde{n}_{2,r}) + \left(1 + \sigma_r^2 a \beta \tilde{n}_{r,r} \right) \left(r^* \kappa_\eta + \sigma_\eta^2 (a \alpha_0 \tilde{n}_{1,\eta} + a \beta \tilde{n}_{2,\eta}) \right) \right),$$

where:

$$d = \left(1 + \sigma_r^2 a \beta \tilde{n}_{r,r} \right) \left(1 + \sigma_\eta^2 a \beta \tilde{n}_{\eta,\eta} \right) - \sigma_r^2 \sigma_\eta^2 a \beta \tilde{n}_{r,\eta} a \beta \tilde{n}_{\eta,r},$$

and

$$C(\tau) = n_{0,r}(\tau) \psi_r + n_{0,\eta}(\tau) \psi_\eta - \frac{1}{2} n(\tau).$$

Appendix B

Market Failure, the QE Trap and Fiscal Escape

Harjoat Bhamra and Nicolo Ceneda

B.1 Complete-Markets Benchmark

When both agents can trade freely in bonds and equity, markets are dynamically complete (one risk source, two assets). Equilibrium allocations are Pareto efficient.

Proposition B.1 (Complete-markets equilibrium) *There exists a constant $\bar{\lambda} \in (0, 1)$ (the household's Pareto weight) such that: consumption shares are constant, $C_{h,t} = \bar{\lambda} Y_t$ and $C_{i,t} = (1 - \bar{\lambda}) Y_t$; labour supply is constant, $N = [(\epsilon - 1)/(\epsilon \bar{\lambda})]^{1/(1+\varphi)}$; and the risk-free rate and market price of risk are constant:*

$$r = \delta + \mu - \sigma^2, \quad \Theta = \sigma. \quad (\text{B.1})$$

The equity price is $P_{D,t} = D_t/\delta = Y_t/(\epsilon\delta)$ and total wealth is $P_{Y,t} = Y_t/\delta$.

Proof. See Appendix [BO.1](#). □

The complete-markets economy serves as a reference point: the good steady state of the model with QE will coincide with this benchmark.

B.2 Baseline Equilibrium: Limited Participation Without Leverage Constraints

We now introduce limited stock market participation, households hold only bonds, but *without* the leverage constraint (2.7). This appendix establishes why the economy functioned

well for decades before 2008: equilibrium always exists and the wealth distribution self-corrects toward the efficient allocation.

B.2.1 State Variable and Equilibrium Characterisation

Following Basak and Cuoco (1998), we characterise equilibrium using the financier's share of total wealth:

$$\omega_t \equiv \frac{H_{i,t}}{P_{Y,t}} \in (0, 1). \quad (\text{B.2})$$

Proposition B.2 (Equilibrium with limited participation) *With log utility and limited participation (no leverage constraint):*

- (i) *Total wealth:* $P_{Y,t} = Y_t/\delta$. *Human capital:* $P_{L,t} = (1 - 1/\epsilon) P_{Y,t}$.
- (ii) *Consumption:* $C_{i,t} = \omega_t Y_t$, $C_{h,t} = (1 - \omega_t) Y_t$.
- (iii) *Employment:* $N_t = (1 - 1/\epsilon)^{1/(1+\varphi)} (1 - \omega_t)^{-1/(1+\varphi)}$.
- (iv) *Financial wealth:* $H_{i,t} = \omega_t P_{Y,t}$, $H_{h,t} = (1/\epsilon - \omega_t) P_{Y,t}$.
- (v) *Portfolio share:* $\phi_t = 1/(\epsilon \omega_t)$.

Three features of this proposition are crucial for what follows.

First, *financial wealth levels are uniquely pinned by ω* . The household's bond holdings are $H_{h,t} = (1/\epsilon - \omega_t) P_{Y,t}$, regardless of the interest rate, risk premium, or any other price. This perfect inelasticity is the engine of the equilibrium failure result.

Second, *the portfolio share is uniquely pinned by market clearing*. Since only the financier holds equity and the equity market must clear: $\phi_t H_{i,t} = P_{D,t}$, so $\phi_t = P_{D,t}/H_{i,t} = (P_{Y,t}/\epsilon)/(\omega_t P_{Y,t}) = 1/(\epsilon \omega_t)$. This is not a choice variable, it is an accounting identity.

Third, *the bond market clears automatically*. The financier borrows $(\phi_t - 1)H_{i,t} = (1/(\epsilon \omega_t) - 1)\omega_t P_{Y,t} = (1/\epsilon - \omega_t)P_{Y,t} = H_{h,t}$. Everything balances.

Remark B.1 (Why leverage rises as intermediary capital falls) *The portfolio share $\phi = 1/(\epsilon \omega)$ is not a portfolio choice — it is a market clearing identity. The financier is the economy's sole equity holder (limited participation), and the equity market has a fixed size relative to total wealth: $P_D/P_Y = 1/\epsilon$, pinned by the Dixit–Stiglitz markup. As the financier's wealth share ω shrinks, the equity market does not shrink with her. She must hold a larger fraction of her wealth in equity to absorb the entire, unchanged*

asset base. Her leverage $\phi = 1/(\epsilon\omega)$ is the ratio of the equity market she must absorb ($P_D = P_Y/\epsilon$) to the wealth she has available ($H_i = \omega P_Y$).

What is the economic significance of the threshold $\omega = 1/\epsilon$? At this value:

- The financier's financial wealth exactly equals the equity market: $H_i = \omega P_Y = P_Y/\epsilon = P_D$.
- The household has zero financial wealth: $H_h = (1/\epsilon - \omega)P_Y = 0$. All her wealth is human capital $P_L = (1 - 1/\epsilon)P_Y$.
- The portfolio share is $\phi = 1$: the financier holds 100% equity, no borrowing, no lending.

For $\omega > 1/\epsilon$, the financier has “too much” wealth for the equity market. She holds equity and lends the surplus to the household. For $\omega < 1/\epsilon$, she has too little wealth and must borrow to buy all the equity. The threshold $1/\epsilon$ is determined entirely by the production structure: ϵ governs the dividend share of output, which governs the size of the equity market relative to the economy.

Remark B.2 (Optimal demand, market clearing and risk-premia) Under log utility, the optimal demand of the financier is $\phi = \Theta/\sigma_R$ (myopic portfolio rule). In equilibrium, Θ and σ_R adjust endogenously so that this optimal demand equals the market-clearing share $1/(\epsilon\omega)$. When ω is small, the market price of risk Θ must be high — risk premia rise to induce the undercapitalised financier to hold a disproportionately large equity position. This is why the model generates countercyclical risk premia without time-varying risk aversion: it is the concentration of risk-bearing in a shrinking intermediary sector, not changes in preferences, that drives the price of risk.

Remark B.3 (Interpreting ϵ as intermediation intensity) In the model, ϵ is the Dixit–Stiglitz elasticity of substitution across goods, and $1/\epsilon$ is the profit (dividend) share of output. But the economic mechanism that drives the results depends only on the ratio $P_D/P_Y = 1/\epsilon$ — the size of the asset market that requires intermediation relative to total wealth. This ratio admits a broader empirical interpretation.

In the 2007–2008 crisis, the relevant “equity” was not limited to the stock market. It included the entire structured credit market — mortgage-backed securities, collateralised debt obligations, asset-backed commercial paper — trillions in assets that required financial intermediary balance sheet capacity. The effective ratio of intermediated assets to GDP grew substantially in the years preceding the crisis: US financial sector assets rose from approximately $3\times$ GDP in 1990 to over $5\times$ by 2007, with shadow banking assets (broker-dealers, ABS issuers, money market funds) growing from roughly \$2 trillion to \$12 trillion over the same period.

In the model's language, the effective $1/\epsilon$ — the share of total economic value embodied in assets requiring intermediation — was rising, while intermediary capital ω was roughly stable. Because the crisis threshold $\omega_c(\epsilon)$ is itself increasing in $1/\epsilon$, the rise in the intermediated asset base shifted the threshold upward toward a stationary ω , rather than ω falling toward a fixed threshold. The required leverage $\phi = 1/(\epsilon\omega)$ rose in lockstep, and the gap $\omega - \omega_c(\epsilon)$ narrowed year after year. By 2007, the system was operating with very little margin: a moderate decline in ω , which would have been absorbed in an earlier decade, was now sufficient to breach the threshold. In this sense the crisis was structural before it was acute — the fall in ω in 2008 was the trigger, but the vulnerability had been accumulating for years through the expansion of the intermediated asset base.

This interpretation suggests a testable cross-sectional prediction: economies with larger intermediated asset markets (higher effective $1/\epsilon$) should be more vulnerable to intermediary capital depletion. We leave this extension to future work.

B.2.2 Dynamics of the Wealth Distribution

Theorem B.1 (Dynamics of the wealth distribution) *The financier's wealth share evolves according to*

$$d\omega_t = \frac{\sigma_\omega^0(\omega_t)^2}{\omega_t} dt + \sigma_\omega^0(\omega_t) dZ_t, \quad (\text{B.3})$$

where the volatility function is

$$\sigma_\omega^0(\omega) = \frac{\sigma(1+\varphi)(1-\omega)(1/\epsilon - \omega)}{(1+\varphi)(1-\omega) - (1/\epsilon - \omega)}. \quad (\text{B.4})$$

The inverse wealth share $Q_t = 1/\omega_t$ is a non-negative local martingale.

The drift relation $\mu_\omega = \sigma_\omega^2/\omega$ follows from the local martingale property of $Q_t = 1/\omega_t$, which we establish in Appendix B.3. This property is the key to the model's tractability: with log utility, the myopic portfolio rule ensures that a critical cancellation in the drift of Q holds for any portfolio share.

The volatility σ_ω^0 has a single interior zero at $\omega^* = 1/\epsilon$ (where $1/\epsilon - \omega = 0$). It is positive for $\omega < 1/\epsilon$ and negative for $\omega > 1/\epsilon$, meaning that positive productivity shocks push ω toward $1/\epsilon$ from both sides.

B.2.3 Two Invariant Regimes and Convergence

Theorem B.2 (Regime classification and convergence) *The state space $(0, 1)$ partitions into two invariant regions separated by the absorbing state $\omega^* = 1/\epsilon$:*

1. **Controlled regime:** $\omega \in (0, 1/\epsilon)$. The financier is leveraged ($\phi > 1$) and households are net lenders ($H_h > 0$).
2. **Inverted regime:** $\omega \in (1/\epsilon, 1)$. The financier holds less than 100% equity ($\phi < 1$) and households are net borrowers ($H_h < 0$).

For any $\omega_0 \in (0, 1) \setminus \{1/\epsilon\}$:

$$\omega_t \xrightarrow{a.s.} \frac{1}{\epsilon}, \quad \text{yet} \quad \mathbb{E}[\tau] = \infty, \quad (\text{B.5})$$

where $\tau = \inf\{t \geq 0 : \omega_t = 1/\epsilon\}$.

The almost-sure convergence with infinite expected hitting time captures a “Zeno’s paradox” of financial markets: the limited participation friction becomes asymptotically negligible yet is never eliminated in finite expected time.

At $\omega^* = 1/\epsilon$, the financier’s wealth equals the stock market capitalisation ($H_i = P_D$), the household has zero financial wealth, leverage is one (no borrowing), and the economy coincides with the complete-markets benchmark. This is why the economy was fine before 2008: limited participation and leverage constraints coexisted without incident, because financiers were well enough capitalised that the constraints never bound. What changed in 2008 was the *state*, financier wealth ω , not the rules.

B.2.4 Asset Prices and Real Variables

Proposition B.3 (Asset prices and real variables) *In the baseline economy with limited participation and no leverage constraint, define*

$$\Delta(\omega) \equiv \left(1 - \frac{1}{\epsilon}\right) + \phi(1 - \omega) > 0. \quad (\text{B.6})$$

Then:

- (i) Market price of risk:

$$\Theta^0(\omega) = \frac{\sigma(1 + \phi)(1 - \omega)}{\epsilon \omega \Delta(\omega)}. \quad (\text{B.7})$$

(ii) Equity risk premium:

$$\mu_R^0 - r^0 = \Theta^0 \sigma_R^0 = \frac{\sigma^2(1+\varphi)^2(1-\omega)^2}{\epsilon \omega \Delta(\omega)^2}. \quad (\text{B.8})$$

(iii) Risk-free rate:

$$r^0(\omega) = \delta + \mu - \sigma^2 + \frac{\varphi}{2}(\sigma_R^0 - \sigma)^2 - \frac{[1 + \varphi(1 - \omega)] \sigma (\sigma_R^0 - \sigma)}{\omega}, \quad (\text{B.9})$$

where $\sigma_R^0 - \sigma = (1/\epsilon - \omega) \sigma / \Delta(\omega)$.

(iv) Log employment:

$$n^0(\omega) \equiv \ln N^0 = \frac{1}{1+\varphi} \ln \left(\frac{1-1/\epsilon}{1-\omega} \right). \quad (\text{B.10})$$

(v) Log output gap (relative to a reference economy with constant wealth share $\omega_n \in (0, 1)$):

$$x^0(\omega) \equiv n^0(\omega) - n_n = \frac{1}{1+\varphi} \ln \left(\frac{1-\omega_n}{1-\omega} \right), \quad (\text{B.11})$$

where $n_n = \frac{1}{1+\varphi} \ln \left(\frac{1-1/\epsilon}{1-\omega_n} \right)$ is log employment in the reference economy.

At $\omega = \omega_n$: $x^0 = 0$. At $\omega = 1/\epsilon$, the asset-pricing quantities recover their complete-markets values: $\Theta^0 = \sigma$, $\mu_R^0 - r^0 = \sigma^2$, $r^0 = \delta + \mu - \sigma^2$, and $n^0 = n_n$ (so $x^0 = 0$ when $\omega_n = 1/\epsilon$).

Remark B.4 (Economic interpretation of the risk-free rate) The risk-free rate (B.9) decomposes as $r^0 = r^{\text{CM}} + r^{\text{precaut}} + r^{\text{risk}}$, where:

- $r^{\text{CM}} = \delta + \mu - \sigma^2$ is the complete-markets rate,
- $r^{\text{precaut}} = \frac{\varphi}{2}(\sigma_R^0 - \sigma)^2 > 0$ is a precautionary term from excess return volatility (mediated by the Frisch elasticity),
- $r^{\text{risk}} = -[1 + \varphi(1 - \omega)] \sigma (\sigma_R^0 - \sigma) / \omega < 0$ when $\omega < 1/\epsilon$ captures the risk-compensation effect of impaired intermediation.

The risk-compensation term is first-order in $\sigma_R^0 - \sigma$ while the precautionary term is second-order, so $r^0 < r^{\text{CM}}$ whenever $\omega \neq 1/\epsilon$. The interest rate falls below its frictionless value when intermediation is impaired, consistent with the empirical evidence on low rates during financial distress.

B.3 Proofs for Limited Participation Without Leverage Constraint

Proof of Proposition B.2. *Optimal policies.* The financier's HJB equation with conjectured value function $V_i = \delta^{-1} \ln H_i + g_i(\omega)$ yields $C_i = \delta H_i$ and $\phi = (\mu_R - r)/\sigma_R^2$ (the myopic log-utility rule). The household, with log utility over consumption and total wealth $H_h + P_L$, consumes $C_h = \delta(H_h + P_L)$.

Goods market clearing. $\delta H_i + \delta(H_h + P_L) = Y$, so $P_Y \equiv H_i + H_h + P_L = Y/\delta$. By definition of ω : $C_i/Y = \delta H_i/Y = H_i/P_Y = \omega$ and $C_h/Y = 1 - \omega$.

Employment. From $N^\varphi C_h = (1 - 1/\epsilon)A$ and $C_h = (1 - \omega)AN$: $N^{1+\varphi} = (1 - 1/\epsilon)/(1 - \omega)$.

Equity market clearing. $\phi H_i = P_D = P_Y/\epsilon$, so $\phi = 1/(\epsilon\omega)$.

Financial wealth. $H_i + H_h = P_D = P_Y/\epsilon$ and $H_i = \omega P_Y$, so $H_h = (1/\epsilon - \omega)P_Y$.

Human capital. $P_L = P_Y - P_D = (1 - 1/\epsilon)P_Y$. Crucially, this holds for *any* dynamics of ω_t , because $P_Y = Y/\delta$ depends only on log utility and goods market clearing. $\square$

Proof of Theorem B.1. *Step 1: Return volatility.* Since $P_D = Y/(\epsilon\delta)$, by Itô's product rule applied to $Y = AN$:

$$\sigma_R = \sigma_Y = \sigma + \frac{\sigma_\omega^0}{(1 + \varphi)(1 - \omega)},$$

where the second term comes from $N \propto (1 - \omega)^{-1/(1+\varphi)}$ via Itô's lemma.

Step 2: Wealth-share volatility. From $\omega = H_i/P_Y$ and Itô's quotient rule:

$$\sigma_\omega^0 = \omega(\phi - 1)\sigma_R = \omega\left(\frac{1}{\epsilon\omega} - 1\right)\sigma_R = \left(\frac{1}{\epsilon} - \omega\right)\sigma_R.$$

Step 3: Fixed-point equation. Substituting into Step 1 and solving:

$$\sigma_R = \frac{\sigma(1 + \varphi)(1 - \omega)}{(1 + \varphi)(1 - \omega) - (1/\epsilon - \omega)},$$

giving $\sigma_\omega^0 = (1/\epsilon - \omega)\sigma_R$, which is (B.4). The denominator simplifies to $(1 - 1/\epsilon) + \varphi(1 - \omega) > 0$.

Step 4: Q-martingale. Define $Q = 1/\omega = P_Y/H_i$. The drift calculation is identical to Appendix B.7 and yields $\mu_Q = 0$. By Itô's lemma on $\omega = 1/Q$: $\mu_\omega = (\sigma_\omega^0)^2/\omega$. $\square$

Proof of Theorem B.2. *Boundary $\omega = 0$ (unattainable).* As $\omega \rightarrow 0^+$: $\sigma_\omega^0(\omega) \rightarrow c_0 > 0$ (a positive constant). The scale function density is $s'(\omega) = \omega_0^2/\omega^2$ and the speed measure is $m(\omega) = \omega^2/(\omega_0^2(\sigma_\omega^0)^2)$. Near $\omega = 0$: $m(\omega) \sim \omega^2/(c_0^2\omega_0^2)$.

$$\Sigma(0) = \int_0^a s'(x) dx = \omega_0^2 \int_0^a x^{-2} dx = +\infty,$$

$$N(0) = \int_0^a s'(x) \int_0^x m(y) dy dx \leq \frac{1}{3c_0^2} \int_0^a x dx < \infty.$$

Therefore, $\omega = 0$ is an entrance boundary, unattainable from the interior.

Boundary $\omega = 1/\epsilon$ (absorbing, unattainable). First, at $\omega = 1/\epsilon$: $\sigma_\omega^0(1/\epsilon) = 0$ and $\mu_\omega(1/\epsilon) = 0$. Therefore, $\omega = 1/\epsilon$ is an absorbing state. Second, write $\sigma_\omega^0 = g(\omega)(1/\epsilon - \omega)$ with $g(1/\epsilon) = \sigma > 0$. By the Engelbert–Schmidt criterion:

$$\int_a^{1/\epsilon} \frac{1}{(\sigma_\omega^0)^2} d\omega \geq \frac{1}{\bar{g}^2} \int_a^{1/\epsilon} \frac{1}{(1/\epsilon - x)^2} dx = +\infty.$$

Therefore, $\omega = 1/\epsilon$ is unattainable from the left. A symmetric argument gives unattainability from the right.

Boundary $\omega = 1$ (unattainable). Write $\sigma_\omega^0 = h(\omega)(1 - \omega)$ with $h(1) = -\sigma(1 + \varphi) \neq 0$. By the Engelbert–Schmidt criterion:

$$\int_b^1 (1 - x)^{-2} dx = +\infty.$$

Therefore, $\omega = 1$ is unattainable.

Regime invariance. Continuous paths and unattainable internal boundaries imply $(0, 1/\epsilon)$ and $(1/\epsilon, 1)$ are invariant.

Convergence. $Q = 1/\omega$ is a non-negative local martingale, hence a supermartingale. By the supermartingale convergence theorem, $Q_t \rightarrow Q_\infty$ a.s. with $Q_\infty < \infty$. Convergence of the quadratic variation requires $\sigma_\omega^0(\omega_\infty) = 0$, so $\omega_\infty \in \{1/\epsilon, 1\}$. Since $Q_\infty < \infty$ implies $\omega_\infty > 0$, and $\omega = 1$ is unattainable: $\omega_t \rightarrow 1/\epsilon$ a.s.

Infinite expected hitting time. The speed measure satisfies $m(\omega) \geq K(1/\epsilon - \omega)^{-2}$ near $1/\epsilon$. By the Green's function formula:

$$\mathbb{E}[\tau_{1/\epsilon}] \geq C \int_a^{1/\epsilon} (1/\epsilon - y)^{-1} dy = +\infty. \quad \square$$

Proof of Proposition B.3. Throughout, write $\sigma_R^0 = \sigma(1 + \varphi)(1 - \omega)/\Delta$ and $\sigma_\omega^0 = (1/\epsilon - \omega) \sigma_R^0$, which are established in the proof of Theorem B.1.

Part (i): Market price of risk. From Proposition B.2, $\phi = 1/(\epsilon\omega)$. The financier's SDF is $\Lambda_t = e^{-\delta t}/C_{i,t}$ with $C_{i,t} = \delta H_{i,t}$, so $\sigma_\Lambda = -\phi \sigma_R^0$ (the financier's wealth volatility). Hence $\Theta^0 = \phi \sigma_R^0 = \sigma_R^0/(\epsilon\omega)$. Substituting gives (B.7).

Verification: at $\omega = 1/\epsilon$, $\sigma_R^0 = \sigma$ and $\Theta^0 = \sigma/(\epsilon \cdot 1/\epsilon) = \sigma$, as in the complete-markets

benchmark. ✓

Part (ii): Equity risk premium. By the standard no-arbitrage relation, $\mu_R^0 - r^0 = \Theta^0 \sigma_R^0 = (\sigma_R^0)^2 / (\epsilon \omega)$. Substituting σ_R^0 yields (B.8).

Verification: at $\omega = 1/\epsilon$, $\mu_R^0 - r^0 = \sigma^2$, as in the complete-markets benchmark. ✓

Part (iii): Risk-free rate. The financier's SDF $\Lambda_t = e^{-\delta t} / (\omega_t Y_t)$ implies $r^0 = \delta + \mu_{\ln(\omega Y)} - \frac{1}{2} (\Theta^0)^2$, where $\Theta^0 = \sigma_{\ln(\omega Y)}$ is the diffusion coefficient established in Part (i).

Step 1: Drift of $\ln \omega$. From the Q -martingale property (Theorem B.1):

$$d \ln \omega = \frac{(\sigma_\omega^0)^2}{2\omega^2} dt + \frac{\sigma_\omega^0}{\omega} dZ_t.$$

Step 2: Drift of $\ln Y$. Since $Y = A N(\omega)$ with $N \propto (1 - \omega)^{-1/(1+\varphi)}$:

$$d \ln Y = \left(\mu - \frac{\sigma^2}{2} + \frac{(\sigma_\omega^0)^2}{(1+\varphi)(1-\omega)\omega} + \frac{(\sigma_\omega^0)^2}{2(1+\varphi)(1-\omega)^2} \right) dt + \sigma_R^0 dZ_t.$$

Step 3: Change of variables. Write $\tilde{\sigma} \equiv \sigma_R^0 - \sigma = \sigma_\omega^0 / [(1 + \varphi)(1 - \omega)]$ and $\sigma_\omega^0 = (1 + \varphi)(1 - \omega)\tilde{\sigma}$.

Summing the drift contributions from Steps 1 and 2:

$$\mu_{\ln(\omega Y)} = \mu - \frac{\sigma^2}{2} + \frac{(1+\varphi)[1+\varphi(1-\omega)^2]\tilde{\sigma}^2}{2\omega^2},$$

where the coefficient of $\tilde{\sigma}^2$ follows from $(1 + \varphi)^2(1 - \omega)^2 + 2(1 + \varphi)\omega(1 - \omega) + (1 + \varphi)\omega^2 = (1 + \varphi)[1 + \varphi(1 - \omega)^2]$.

Step 4: Subtraction. Since $(\Theta^0)^2 = [\sigma + (1 + \varphi(1 - \omega))\tilde{\sigma}/\omega]^2$:

$$r^0 = \delta + \mu - \sigma^2 + \frac{\tilde{\sigma}^2}{2\omega^2} \left\{ (1 + \varphi)[1 + \varphi(1 - \omega)^2] - [1 + \varphi(1 - \omega)]^2 \right\} - \frac{[1 + \varphi(1 - \omega)]\sigma\tilde{\sigma}}{\omega}.$$

The curly brace simplifies to $\varphi\omega^2$. This yields (B.9).

Verification: at $\omega = 1/\epsilon$, $\tilde{\sigma} = 0$, so $r^0 = \delta + \mu - \sigma^2$, as in the complete-markets benchmark. ✓

Parts (iv)–(v): Employment and output gap. From Proposition B.2, $N^{1+\varphi} = (1 - 1/\epsilon)/(1 - \omega)$, giving (B.10). The reference economy has the same employment formula evaluated at ω_n : $N_n = [(1 - 1/\epsilon)/(1 - \omega_n)]^{1/(1+\varphi)}$, so $n_n = \frac{1}{1+\varphi} \ln \frac{1-1/\epsilon}{1-\omega_n}$. Since $Y = AN$ and productivity A_t is common to both economies:

$$x^0 = \ln Y - \ln Y_n = \ln N^0 - \ln N_n = n^0 - n_n = \frac{1}{1+\varphi} \ln \left(\frac{1 - \omega_n}{1 - \omega} \right),$$

which is (B.11). Note that $(1 - 1/\epsilon)$ cancels in the difference $n^0 - n_n$.

Verification: $x^0(\omega_n) = 0$; for $\omega > \omega_n$: $x^0 > 0$ (output above the reference); for $\omega < \omega_n$: $x^0 < 0$ (output below the reference). $\checkmark$ $\square$

B.4 Equilibrium with Leverage Constraints but Without QE

We now add the leverage constraint (2.7) to the limited-participation model of Appendix B.2, but *without* introducing a central bank. This intermediate case — between the unconstrained baseline (Appendix B.2) and the full model with QE (Section 2.4) — reveals both the source of equilibrium failure and the fragility of the pre-crisis economy.

B.4.1 Equilibrium Above the Crisis Threshold

For $\omega > \omega_c$, the leverage constraint does not bind: $\bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega}) \geq 1/(\epsilon\omega)$. The financier freely chooses $\phi = 1/(\epsilon\omega)$ (the market-clearing portfolio share from Proposition B.2). Since no constraint is active, *all equilibrium quantities are identical to the unconstrained baseline*:

$$\sigma_\omega(\omega) = \sigma_\omega^0(\omega) = \frac{\sigma(1+\varphi)(1-\omega)(1/\epsilon - \omega)}{(1+\varphi)(1-\omega) - (1/\epsilon - \omega)}, \quad (\text{B.12})$$

$$\Theta(\omega) = \frac{\sigma_\omega^0(\omega)}{\omega} + \sigma + \frac{\sigma_\omega^0(\omega)}{(1+\varphi)(1-\omega)}, \quad (\text{B.13})$$

$$\sigma_R(\omega) = \sigma + \frac{\sigma_\omega^0(\omega)}{(1+\varphi)(1-\omega)}, \quad (\text{B.14})$$

$$r(\omega) = \delta + \mu - \sigma^2 + \frac{\varphi}{2} \tilde{\sigma}(\omega)^2 - \frac{1+\varphi(1-\omega)}{\omega} \sigma \tilde{\sigma}(\omega), \quad (\text{B.15})$$

where $\tilde{\sigma}(\omega) \equiv \sigma_R(\omega) - \sigma = \sigma_\omega^0(\omega)/[(1+\varphi)(1-\omega)]$. Employment, output, and the output gap are also unchanged:

$$N(\omega) = \left(\frac{1 - 1/\epsilon}{1 - \omega} \right)^{\frac{1}{1+\varphi}}, \quad x(\omega) = \frac{1}{1+\varphi} \ln \frac{1 - 1/\epsilon}{1 - \omega}. \quad (\text{B.16})$$

The intermediation gap $b(\omega) > 0$ from Proposition 2.3 is also present: since $Q = 1/\omega$ is a strict local martingale on $(0, 1/\epsilon)$ under the unconstrained dynamics (Theorem 2.6(i)), asset prices exceed fundamental values.

B.4.2 The Crisis Threshold

At $\omega = \omega_c$, defined by $\bar{\phi}(\omega_c) = 1/(\epsilon\omega_c)$ (equation (2.10)), the constraint just binds. At this point:

- The financier's leverage reaches its permitted maximum: $\phi = \bar{\phi}(\omega_c) = 1/(\epsilon\omega_c) \approx 1.20$.
- Risk premia are at their peak: $\Theta(\omega_c) \approx 0.181$, $\text{ERP}(\omega_c) \approx 2.74\%$.
- The risk-free rate is at its trough: $r(\omega_c) \approx 2.42\%$.
- The output gap is $x(\omega_c) \approx -0.55\%$.

These are the *most distorted* equilibrium values the economy can sustain without central bank intervention.

B.4.3 Market Failure Below the Crisis Threshold

For $\omega < \omega_c$, no competitive equilibrium exists (Theorem 2.1). The required portfolio share $\phi = 1/(\epsilon\omega)$ exceeds the leverage limit $\bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega})$. The excess supply of equity, as a fraction of total wealth, is:

$$\frac{P_D - \bar{\phi}(\omega)\omega P_Y}{P_Y} = \frac{1}{\epsilon} - \bar{\phi}(\omega)\omega = \frac{1}{\epsilon} - \omega - \lambda\omega(\omega - \underline{\omega}). \quad (\text{B.17})$$

At the baseline calibration: the excess supply is 0.1% of total wealth at $\omega = 0.060$ (just below ω_c), rising to 1.6% at $\omega = 0.05$ and 4.7% at $\omega = \underline{\omega} = 0.0260$. No price adjustment can eliminate this excess supply (Section 2.3).

Below ω_c , there are no equilibrium prices, no risk premia, no risk-free rate, and no well-defined output. The economy hits a cliff.

B.4.4 Crisis Probability

Although the economy drifts toward $1/\epsilon$ in expectation ($\mu_\omega = \sigma_\omega^2/\omega > 0$), negative productivity shocks can push ω below ω_c . The probability of market failure before recovery is computed from the scale function.

For the unconstrained dynamics (B.12), the drift is $\mu_\omega = \sigma_\omega^2/\omega$ (from the Q -martingale property, Proposition 2.2). The scale function density is $s'(\omega) = C/\omega^2$ (Appendix B.9, equation

for $2\mu_\omega/\sigma_\omega^2 = 2/\omega$, giving scale function $s(\omega) = -C/\omega$. The probability of hitting ω_c before $1/\epsilon$, starting from $\omega_0 \in (\omega_c, 1/\epsilon)$, is:

$$\Pr(\tau_{\omega_c} < \tau_{1/\epsilon} \mid \omega_0) = \frac{s(\omega_0) - s(1/\epsilon)}{s(\omega_c) - s(1/\epsilon)} = \frac{1/\omega_0 - \epsilon}{1/\omega_c - \epsilon}. \quad (\text{B.18})$$

This probability is substantial even for ω_0 well above ω_c :

ω_0	0.062	0.065	0.068	0.071
$\Pr(\text{crisis})$	86.3%	59.0%	34.2%	11.4%

Starting just above the threshold ($\omega_0 = 0.062$), the economy has an 86% probability of market failure before recovery. Even from $\omega_0 = 0.068$, the probability is 34%. The economy is fragile: the leverage constraint exists as a latent vulnerability, and a sufficiently large shock triggers catastrophic failure.

B.4.5 Comparison of the Three Cases

Table B.1 summarises the three cases analysed in the paper:

Table B.1: Equilibrium properties under three model configurations. “LP” = limited participation (household holds only bonds). All three share the same employment and output gap functions $N(\omega)$ and $x(\omega)$.

	LP only (Appendix B.2)	LP + leverage, no QE	LP + leverage + QE (Sections 2.3–2.6)
Domain	$(0, 1/\epsilon)$	$(\omega_c, 1/\epsilon)$	$(0, 1)$
Equilibrium	Always	$\omega > \omega_c$ only	Always (CB absorbs residual)
σ_ω formula	Unconstrained	Unconstrained	Piecewise (eq. 2.23)
$\sigma_\omega(0^+)$	$c_0 > 0$	N/A (no eqm)	$\sim c'_0 \omega \rightarrow 0$
$Q = 1/\omega$	Strict local M	N/A	True M
Intermediation gap	Positive	Positive ($\omega > \omega_c$)	Zero
Absorbing states	$1/\epsilon$	$1/\epsilon$ (if reached)	$\underline{\omega}$ and $1/\epsilon$
QE Trap	No	No	Yes
CB equity	None	None	Up to 64%

The “LP + leverage, no QE” case reveals the economy’s *latent fragility*: the leverage constraint introduces a crisis threshold that did not exist in the unconstrained baseline, but does not change any equilibrium quantity until the threshold is reached. The constraint is invisible in normal times and catastrophic in crisis. QE resolves the market failure but, by extending the domain to include $\omega < \omega_c$, introduces the absorbing bad steady state at $\underline{\omega}$ — the QE Trap.

B.5 Proof of Equilibrium Failure

Proof that $P_{L,t} = (1 - 1/\epsilon)P_{Y,t}$ regardless of ω_t dynamics. Human capital is

$$P_{L,t} = \mathbb{E}_t \int_t^\infty \frac{\Lambda_s}{\Lambda_t} \frac{\epsilon - 1}{\epsilon} Y_s ds = \frac{\epsilon - 1}{\epsilon} \mathbb{E}_t \int_t^\infty \frac{\Lambda_s}{\Lambda_t} Y_s ds.$$

The financier's SDF is $\Lambda_t = e^{-\delta t} / C_{i,t} = e^{-\delta t} / (\delta H_{i,t})$. Total wealth is $P_{Y,t} = \mathbb{E}_t \int_t^\infty (\Lambda_s / \Lambda_t) Y_s ds$. From goods market clearing, $P_{Y,t} = Y_t / \delta$ for any dynamics. Therefore $P_{L,t} = (1 - 1/\epsilon)Y_t / \delta = (1 - 1/\epsilon)P_{Y,t}$.

The key point is that $P_{Y,t} = Y_t / \delta$ is an identity under log utility: $P_Y = (C_i + C_h) / \delta = Y / \delta$. It does not require any specific dynamics for ω , the interest rate, or risk premia. $\square$

Proof of Theorem 2.1. The argument proceeds in three steps.

Step 1: Household bond demand is perfectly inelastic. By Proposition B.2 (which holds for any ω dynamics), the household's financial wealth is $H_{h,t} = (1/\epsilon - \omega_t) P_{Y,t}$. With log utility, this is a function of ω alone; it does not depend on the interest rate, the equity premium, or any other price.

Step 2: Equity market clearing pins the portfolio share. Since only the financier and (potentially) the central bank hold equity: $\phi_t H_{i,t} + \alpha_t P_{D,t} = P_{D,t}$. Without central bank intervention ($\alpha = 0$), this requires $\phi_t = P_{D,t} / H_{i,t} = 1/(\epsilon \omega_t)$.

Step 3: The leverage constraint makes the required ϕ infeasible. When $\omega < \omega_c$, the required portfolio share $1/(\epsilon \omega)$ exceeds the maximum permitted leverage $\bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega})$. No interest rate adjustment can resolve this: the household's bond demand is inelastic (Step 1), so reducing the interest rate does not induce the household to reduce her lending. There is a residual excess supply of bonds $(1/\epsilon - \omega - (\bar{\phi}(\omega) - 1)\omega) P_{Y,t} > 0$ that no price can eliminate. No competitive equilibrium exists. $\square$

Proof of Proposition 2.1. With the central bank holding fraction $\alpha_t = 1 - \epsilon \bar{\phi}(\omega_t) \omega_t$ of the equity market (Assumption 2.3), the financier's equity holding is $\phi_t H_{i,t} = \bar{\phi}(\omega_t) \omega_t P_{Y,t}$ and the central bank's is $\alpha_t P_{D,t} = (1 - \epsilon \bar{\phi}(\omega_t) \omega_t) P_{Y,t} / \epsilon$. Their sum is $P_{D,t} = P_{Y,t} / \epsilon$: the equity market clears.

The central bank finances its equity purchases by issuing bonds worth $\alpha_t P_{D,t}$. The household's bond demand is $(1/\epsilon - \omega_t) P_{Y,t}$ (perfectly inelastic). The financier borrows $(\bar{\phi}(\omega_t) -$

1) $\omega_t P_{Y,t}$. Bond market clearing requires:

$$(1/\epsilon - \omega_t)P_{Y,t} = (\bar{\phi}(\omega_t) - 1)\omega_t P_{Y,t} + \alpha_t P_{D,t}.$$

Substituting $\alpha_t P_{D,t} = (1/\epsilon - \bar{\phi}(\omega_t)\omega_t)P_{Y,t}$ and simplifying: $(1/\epsilon - \omega_t) = (\bar{\phi}(\omega_t)\omega_t - \omega_t) + (1/\epsilon - \bar{\phi}(\omega_t)\omega_t) = 1/\epsilon - \omega_t$. The identity holds.

All real quantities (consumption, employment, output, wealth) depend only on ω_t through the same expressions as Proposition B.2, since $P_{Y,t} = Y_t/\delta$ is an identity under log utility. $\square$

B.6 Derivation of σ_ω With Leverage Constraint

Proof of Theorem 2.2. Step 1: Return volatility. Since $P_D = Y/(\epsilon\delta)$ and δ is constant, $\sigma_R = \sigma_Y$. Employment is $N(\omega) = [(1 - 1/\epsilon)/(1 - \omega)]^{1/(1+\varphi)}$. Define $f(\omega) = (1 - \omega)^{-1/(1+\varphi)}$, so $N = c f(\omega)$ with $c = (1 - 1/\epsilon)^{1/(1+\varphi)}$. By Itô's lemma:

$$\frac{f'(\omega)}{f(\omega)} = \frac{1}{(1 + \varphi)(1 - \omega)}, \quad \frac{dN}{N} = \frac{f'}{f} d\omega + (\text{drift}) dt = \frac{\sigma_\omega}{(1 + \varphi)(1 - \omega)} dZ + (\dots) dt.$$

Since $Y = AN$ and both A and N load on the same Brownian motion Z (A directly, N through ω), the product rule gives $\sigma_Y = \sigma_A + \sigma_N$:

$$\sigma_Y = \sigma + \frac{\sigma_\omega}{(1 + \varphi)(1 - \omega)}.$$

Step 2: Wealth-share volatility. The financier's wealth SDE is (equation (B.19) in Appendix B.7):

$$\frac{dH_i}{H_i} = [r + \phi(\mu_R - r) - \delta] dt + \phi\sigma_R dZ.$$

With $P_Y = Y/\delta$, $dP_Y/P_Y = \mu_Y dt + \sigma_Y dZ$. Applying Itô's quotient rule to $\omega = H_i/P_Y$ (see Appendix B.7, equation (B.21) applied to $1/Q = \omega$):

$$\sigma_\omega = \omega \left(\frac{\sigma_{H_i}}{H_i} - \frac{\sigma_{P_Y}}{P_Y} \right) = \omega(\phi\sigma_R - \sigma_Y) = \omega(\phi\sigma_R - \sigma_R) = \omega(\phi - 1)\sigma_R,$$

where the third equality uses $\sigma_R = \sigma_Y$.

Step 3: Fixed-point equation. Substituting $\sigma_R = \sigma + \sigma_\omega/[(1 + \varphi)(1 - \omega)]$ (Step 1) into $\sigma_\omega = \omega(\phi - 1)\sigma_R$ (Step 2) yields an equation in which σ_ω appears on both sides. Since the

equation is linear in σ_ω , it has a unique solution whenever $D(\omega) \neq 0$.

Constrained region ($\omega < \omega_c$, $\phi = \bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega})$, so $\phi - 1 = \lambda(\omega - \underline{\omega})$):

$$\begin{aligned}\sigma_\omega &= \omega\lambda(\omega - \underline{\omega}) \left[\sigma + \frac{\sigma_\omega}{(1 + \varphi)(1 - \omega)} \right] \\ &= \sigma\lambda\omega(\omega - \underline{\omega}) + \frac{\lambda\omega(\omega - \underline{\omega})}{(1 + \varphi)(1 - \omega)} \sigma_\omega.\end{aligned}$$

Collecting σ_ω terms:

$$\sigma_\omega \left[1 - \frac{\lambda\omega(\omega - \underline{\omega})}{(1 + \varphi)(1 - \omega)} \right] = \sigma\lambda\omega(\omega - \underline{\omega}).$$

Multiplying through by $(1 + \varphi)(1 - \omega)$ and solving:

$$\sigma_\omega^c = \frac{\sigma(1 + \varphi)\lambda(1 - \omega)\omega(\omega - \underline{\omega})}{D_c(\omega)}, \quad D_c(\omega) = (1 + \varphi)(1 - \omega) - \lambda\omega(\omega - \underline{\omega}).$$

Unconstrained region ($\omega \geq \omega_c$, $\phi = 1/(\epsilon\omega)$, so $\phi - 1 = (1/\epsilon - \omega)/\omega$):

$$\sigma_\omega = (1/\epsilon - \omega) \left[\sigma + \frac{\sigma_\omega}{(1 + \varphi)(1 - \omega)} \right].$$

The identical rearrangement gives:

$$\sigma_\omega^u = \frac{\sigma(1 + \varphi)(1 - \omega)(1/\epsilon - \omega)}{D_u(\omega)}, \quad D_u(\omega) = (1 + \varphi)(1 - \omega) - (1/\epsilon - \omega).$$

Continuity at ω_c : At ω_c , the defining equation (2.10) gives $\lambda\omega_c(\omega_c - \underline{\omega}) = 1/\epsilon - \omega_c$. Substituting into σ_ω^c : the numerator becomes $\sigma(1 + \varphi)(1 - \omega_c)(1/\epsilon - \omega_c)$, and $D_c(\omega_c) = (1 + \varphi)(1 - \omega_c) - (1/\epsilon - \omega_c) = D_u(\omega_c)$. Therefore $\sigma_\omega^c(\omega_c) = \sigma_\omega^u(\omega_c)$.

Step 4: Drift. By Proposition 2.2, $Q = 1/\omega$ is a driftless local martingale. Itô's lemma on $\omega = 1/Q$ with $f(Q) = 1/Q$, $f'(Q) = -1/Q^2$, $f''(Q) = 2/Q^3$ gives drift $\mu_\omega = \sigma_\omega^2/\omega$ (equation (B.22)).

□

Lemma B.1 (Denominator positivity) $D(\omega) > 0$ for all $\omega \in [0, 1]$ in both regions.

Proof. For $\omega \geq \omega_c$: expand $D_u = (1 + \varphi)(1 - \omega) - (1/\epsilon - \omega) = (1 + \varphi) - (1 + \varphi)\omega - 1/\epsilon + \omega = (1 - 1/\epsilon) + \varphi(1 - \omega) > 0$ since $\epsilon > 1$ and $\varphi \geq 0$.

For $\omega \in [0, \underline{\omega}]$: $\omega - \underline{\omega} \leq 0$, so $\lambda\omega(\omega - \underline{\omega}) \leq 0$, giving $D_c = (1 + \varphi)(1 - \omega) - \lambda\omega(\omega - \underline{\omega}) \geq (1 + \varphi)(1 - \omega) > 0$.

For $\omega \in [\underline{\omega}, \omega_c]$: $D_c(\omega) = (1 + \varphi)(1 - \omega) - \lambda\omega^2 + \lambda\underline{\omega}\omega$ is quadratic in ω with $d^2D_c/d\omega^2 =$

$-2\lambda < 0$ (concave). At the endpoints: $D_c(\underline{\omega}) = (1 + \varphi)(1 - \underline{\omega}) > 0$ and $D_c(\omega_c) = D_u(\omega_c) > 0$ (by continuity). A concave function positive at both endpoints is positive throughout. $\square$

Proof of Proposition BO.1. Consider the volatility function (2.23) with denominator $D(\omega) > 0$ for all $\omega \in [0, 1)$ (Lemma B.1). All pre-factors σ , $(1 + \varphi)$, and λ are strictly positive.

Zeros. In the constrained region ($\omega < \omega_c$), the numerator is $\sigma(1 + \varphi)\lambda(1 - \omega)\omega(\omega - \underline{\omega})$. For $\omega \in [0, \omega_c)$, the factors $(1 - \omega)$ and $D_c(\omega)$ are strictly positive, so $\sigma_\omega(\omega) = 0$ if and only if $\omega = 0$ or $\omega = \underline{\omega}$. In the unconstrained region ($\omega \geq \omega_c$), the numerator is $\sigma(1 + \varphi)(1 - \omega)(1/\epsilon - \omega)$. For $\omega \in [\omega_c, 1)$, the factor $(1 - \omega) > 0$ and $D_u > 0$, so $\sigma_\omega(\omega) = 0$ if and only if $\omega = 1/\epsilon$. (Note $\omega_c < 1/\epsilon$ since $\bar{\phi}(\omega_c) = 1/(\epsilon\omega_c) > 1$, which requires $\omega_c < 1/\epsilon$.) Combining: the zeros of σ_ω in $[0, 1)$ are exactly $\{0, \underline{\omega}, 1/\epsilon\}$.

Sign pattern. For $\omega \in (0, \underline{\omega})$: $\omega > 0$, $(1 - \omega) > 0$, $(\omega - \underline{\omega}) < 0$, and $D_c > 0$. The numerator has sign $(+)(+)(+)(-)< 0$, so $\sigma_\omega < 0$. For $\omega \in (\underline{\omega}, \omega_c)$: $(\omega - \underline{\omega}) > 0$, so all factors in the numerator are positive, giving $\sigma_\omega > 0$. For $\omega \in [\omega_c, 1/\epsilon)$: $(1/\epsilon - \omega) > 0$ and $(1 - \omega) > 0$, so $\sigma_\omega > 0$. Continuity of σ_ω at ω_c (established in the proof of Theorem 2.2) ensures consistency across the boundary. For $\omega \in (1/\epsilon, 1)$: $(1/\epsilon - \omega) < 0$ and $(1 - \omega) > 0$, so $\sigma_\omega < 0$. $\square$

B.7 Q-Martingale Property

Proof of Proposition 2.2.

Step 1: Financier wealth SDE. The financier allocates portfolio share ϕ_t to equity (with return $dR_t = \mu_R dt + \sigma_R dZ_t$) and the remainder to the risk-free bond (at rate r_t), and consumes $C_{i,t}$. Her wealth evolves as:

$$\frac{dH_{i,t}}{H_{i,t}} = [r_t + \phi_t(\mu_{R,t} - r_t) - \delta] dt + \phi_t \sigma_{R,t} dZ_t, \quad (\text{B.19})$$

where the consumption drain $-\delta$ arises because log utility gives $C_{i,t} = \delta H_{i,t}$ (the first-order condition for $\max \mathbb{E} \int_0^\infty e^{-\delta t} \ln C_{i,t} dt$).

Step 2: SDF and market price of risk. The financier's stochastic discount factor is $\Lambda_t = e^{-\delta t}/C_{i,t} = e^{-\delta t}/(\delta H_{i,t})$. Applying Itô's lemma to $\Lambda_t = (1/\delta)e^{-\delta t}/H_{i,t}$ with $f(H) = 1/H$, $f'(H) = -1/H^2$:

$$\frac{d\Lambda_t}{\Lambda_t} = -\delta dt - \frac{dH_{i,t}}{H_{i,t}} + \left(\frac{dH_{i,t}}{H_{i,t}} \right)^2 = (\dots) dt - \phi_t \sigma_{R,t} dZ_t.$$

The volatility of Λ loading on dZ is $-\phi_t \sigma_R$, so the market price of risk is:

$$\Theta_t = \phi_t \sigma_{R,t}. \quad (\text{B.20})$$

This holds for *any* feasible ϕ_t , whether chosen optimally or imposed by the leverage constraint. Substituting into $\mu_R - r = \sigma_R \Theta$: $\mu_R - r = \phi_t \sigma_R^2$.

Step 3: Total wealth SDE. Total wealth $P_{Y,t} = Y_t/\delta$ pays continuous dividends $C_t = Y_t = \delta P_{Y,t}$, so the dividend yield is δ . No-arbitrage for a claim with dividend yield δ :

$$\frac{\mu_{P_Y}}{P_Y} + \delta = r + \sigma_{P_Y} \Theta_t, \quad \text{i.e.,} \quad \frac{\mu_{P_Y}}{P_Y} = r + \sigma_Y \Theta_t - \delta,$$

where $\sigma_{P_Y} = \sigma_Y$ since $P_Y = Y/\delta$ and δ is constant, so $dP_Y/P_Y = dY/Y$.

Step 4: Drift of Q . Define $Q_t = P_{Y,t}/H_{i,t} = 1/\omega_t$. We apply the Itô quotient rule: for $Q = X/Y$ with $dX/X = \mu_X dt + \sigma_X dZ$ and $dY/Y = \mu_Y dt + \sigma_Y dZ$,

$$\frac{\mu_Q}{Q} = \frac{\mu_X}{X} - \frac{\mu_Y}{Y} + \left(\frac{\sigma_Y}{Y}\right)^2 - \frac{\sigma_X}{X} \cdot \frac{\sigma_Y}{Y}, \quad \frac{\sigma_Q}{Q} = \frac{\sigma_X}{X} - \frac{\sigma_Y}{Y}.$$

Setting $X = P_Y$ and $Y = H_i$, with $\sigma_X/X = \sigma_Y$ and $\sigma_Y/Y = \phi_t \sigma_R$:

$$\begin{aligned} \frac{\mu_Q}{Q} &= \underbrace{(r + \sigma_Y \phi_t \sigma_R - \delta)}_{\mu_{P_Y}/P_Y} - \underbrace{(r + \phi_t^2 \sigma_R^2 - \delta)}_{\mu_{H_i}/H_i} + \underbrace{\phi_t^2 \sigma_R^2}_{(\sigma_{H_i}/H_i)^2} - \underbrace{\sigma_Y \phi_t \sigma_R}_{(\sigma_{P_Y}/P_Y) \cdot (\sigma_{H_i}/H_i)} \\ &= \sigma_Y \phi_t \sigma_R - \phi_t^2 \sigma_R^2 + \phi_t^2 \sigma_R^2 - \phi_t \sigma_R \sigma_Y \\ &= 0. \end{aligned} \quad (\text{B.21})$$

The cancellation occurs in two pairs: terms 2 and 3 ($-\phi_t^2 \sigma_R^2 + \phi_t^2 \sigma_R^2 = 0$) and terms 1 and 4 ($\sigma_Y \phi_t \sigma_R - \phi_t \sigma_R \sigma_Y = 0$, using $\sigma_R = \sigma_Y$ since $P_D \propto Y$ implies equity return volatility equals output volatility). Every ϕ_t -dependent term cancels, so the result holds for *any* ϕ_t .

The diffusion of Q is:

$$\frac{\sigma_Q}{Q} = \sigma_Y - \phi_t \sigma_R = \sigma_R - \phi_t \sigma_R = (1 - \phi_t) \sigma_R = -\frac{\sigma_\omega}{\omega},$$

where the last step uses $\sigma_\omega = \omega(\phi_t - 1)\sigma_R$ from Appendix B.6.

Step 5: Q as a local martingale. Since $\mu_Q = 0$, we apply Itô's lemma to $Q = 1/\omega$ with

$f(\omega) = 1/\omega$, $f'(\omega) = -1/\omega^2$, $f''(\omega) = 2/\omega^3$:

$$\begin{aligned} dQ &= f'(\omega) d\omega + \frac{1}{2} f''(\omega) (d\omega)^2 \\ &= -\frac{1}{\omega^2} [\mu_\omega dt + \sigma_\omega dZ] + \frac{1}{\omega^3} \sigma_\omega^2 dt \\ &= \left[-\frac{\mu_\omega}{\omega^2} + \frac{\sigma_\omega^2}{\omega^3} \right] dt - \frac{\sigma_\omega}{\omega^2} dZ. \end{aligned}$$

Setting the drift to zero ($\mu_Q = 0$ from Step 4):

$$\mu_\omega = \frac{\sigma_\omega^2}{\omega}. \quad (\text{B.22})$$

Therefore $dQ_t = -(\sigma_\omega(\omega_t)/\omega_t^2) dZ_t$. Since $\omega \in (0, 1)$, $Q = 1/\omega > 1 > 0$. On any compact subset of $(0, 1)$, $\sigma_\omega(\omega)/\omega^2$ is bounded (by Lemma B.1 and continuity of the numerator), so $\int_0^t (\sigma_\omega/\omega^2)^2 ds < \infty$ a.s. for all t . Therefore Q is a non-negative continuous local martingale, and hence a supermartingale by Fatou's lemma: for any localising sequence $\tau_n \uparrow \infty$, $\mathbb{E}[Q_t | \mathcal{F}_s] = \lim_n \mathbb{E}[Q_{t \wedge \tau_n} | \mathcal{F}_s] \leq \liminf_n \mathbb{E}[Q_{t \wedge \tau_n} | \mathcal{F}_s] = Q_s$. $\square$

B.8 First-Passage Analysis on the Pre-Failure Domain

This appendix analyses the first-passage time τ_c of the financier wealth-share process ω_t to the crisis threshold ω_c , on the pre-failure domain $I \equiv (\omega_c, 1/\epsilon)$ and under the *unconstrained* (pre-QE) law of motion (2.13)–(2.14). The scope is complementary to Appendix B.7 (which proves $Q_t = 1/\omega_t$ is a non-negative continuous local martingale with $Q_t \rightarrow \epsilon$ almost surely), Appendix B.9 (Feller classification), Appendix B.10 (a.s. convergence and infinite expected hitting time to $1/\epsilon$), and Appendix BO.14.1 (a deterministic lower bound on recovery using the *constrained+QE* σ_ω on $(\underline{\omega}, \omega_c)$). Here we compute in closed form the probability $\mathbb{P}(\tau_c < \infty | \omega_0)$, show that the unconditional expected hitting time is infinite, and represent the conditional expected hitting time as a Green's-function integral that admits an elementary closed form at $\varphi = 0$.

B.8.1 Preliminaries

Recall from (2.14) that on $I = (\omega_c, 1/\epsilon)$,

$$\sigma_\omega(\omega) = \frac{\sigma(1+\varphi)(1-\omega)(1/\epsilon - \omega)}{(1+\varphi)(1-\omega) - (1/\epsilon - \omega)}, \quad (\text{B.23})$$

with $0 < \omega_c < 1/\epsilon$, $\epsilon > 1$, $\varphi \geq 0$, $\sigma > 0$. The denominator $D(\omega) \equiv (1+\varphi)(1-\omega) - (1/\epsilon - \omega) = \varphi(1-\omega) + (1-1/\epsilon) > 0$ for all $\omega \in [0, 1]$ since both summands are non-negative and the second is strictly positive ($\epsilon > 1$). The numerator $\sigma(1+\varphi)(1-\omega)(1/\epsilon - \omega)$ is strictly positive on I , so $\sigma_\omega(\omega) > 0$ throughout I and σ_ω is continuous on $\bar{I}$ with $\sigma_\omega(1/\epsilon) = 0$. The generator of the diffusion (2.13) on I is

$$\mathcal{L}f(\omega) = \frac{\sigma_\omega(\omega)^2}{\omega} f'(\omega) + \frac{\sigma_\omega(\omega)^2}{2} f''(\omega). \quad (\text{B.24})$$

A scale function for $\mathcal{L}$ on I is any twice-differentiable solution of $\mathcal{L}s = 0$. Direct computation with $s(\omega) = 1/\omega$, $s'(\omega) = -1/\omega^2$, $s''(\omega) = 2/\omega^3$ gives

$$\mathcal{L}s(\omega) = \frac{\sigma_\omega^2}{\omega} \left(-\frac{1}{\omega^2} \right) + \frac{\sigma_\omega^2}{2} \cdot \frac{2}{\omega^3} = -\frac{\sigma_\omega^2}{\omega^3} + \frac{\sigma_\omega^2}{\omega^3} = 0, \quad (\text{B.25})$$

confirming that $s(\omega) = 1/\omega$ is a scale function—equivalently, that $Q_t = 1/\omega_t$ is a local martingale (Proposition 2.2). Since the drift in (2.13) is $\mu_\omega = \sigma_\omega^2/\omega$, the scale-function integrand is $-2\mu_\omega/\sigma_\omega^2 = -2/\omega$, so $s'(\omega) = \exp(-\int^\omega 2/u \, du) = C/\omega^2$, consistent with $s(\omega) = -C/\omega + \text{const}$; we choose $s(\omega) = 1/\omega$ below. The general solution of the homogeneous ODE $\mathcal{L}f = 0$ on I is therefore $f(\omega) = A + B/\omega$ for constants $A, B \in \mathbb{R}$.

B.8.2 Hitting Probability

Proposition B.4 (Hitting probability for τ_c) For $\omega_0 \in I = (\omega_c, 1/\epsilon)$, the stopping time $\tau_c \equiv \inf\{t \geq 0 : \omega_t = \omega_c\}$ defined in (2.15) satisfies

$$p(\omega_0) \equiv \mathbb{P}(\tau_c < \infty \mid \omega_0) = \frac{1/\omega_0 - \epsilon}{1/\omega_c - \epsilon} = \frac{\omega_c(1 - \epsilon\omega_0)}{\omega_0(1 - \epsilon\omega_c)}. \quad (\text{B.26})$$

The function $p : [\omega_c, 1/\epsilon] \rightarrow [0, 1]$ is continuous with $p(\omega_c) = 1$ and $p(1/\epsilon) = 0$, and p is strictly decreasing in ω_0 on I .

Proof. We give two derivations.

(a) *Optional stopping on a bounded martingale.* By Proposition 2.2 and (B.25), $Q_t = 1/\omega_t$

is a continuous local martingale on the interval I . For $n \in \mathbb{N}$ with $1/n < 1/\epsilon - \omega_c$, define $I_n \equiv (\omega_c, 1/\epsilon - 1/n)$ and the exit time

$$\tau_n \equiv \inf\{t \geq 0 : \omega_t \notin I_n\} = \tau_c \wedge \sigma_n, \quad \sigma_n \equiv \inf\{t \geq 0 : \omega_t = 1/\epsilon - 1/n\}.$$

Since $\omega_t \in I_n$ for $t < \tau_n$, the process $Q_{t \wedge \tau_n}$ takes values in the compact interval $[1/(1/\epsilon - 1/n), 1/\omega_c]$ and is therefore a bounded continuous martingale. By the optional stopping theorem (Karatzas and Shreve, 1991, Theorem 1.3.22),

$$\mathbb{E}[Q_{\tau_n} \mid \omega_0] = Q_0 = \frac{1}{\omega_0}. \quad (\text{B.27})$$

By Feller classification in Appendix B.9, the right endpoint $1/\epsilon$ is unattainable, so $\mathbb{P}(\sigma_n < \infty) = 1$ for each n but $\sigma_n \rightarrow +\infty$ a.s. as $n \rightarrow \infty$; similarly $\tau_c = \lim_n \tau_n$ on $\{\tau_c < \infty\}$. The left-hand side of (B.27) decomposes as

$$\mathbb{E}[Q_{\tau_n} \mathbf{1}_{\{\tau_c < \sigma_n\}}] + \mathbb{E}[Q_{\tau_n} \mathbf{1}_{\{\sigma_n \leq \tau_c\}}] = \frac{1}{\omega_c} \mathbb{P}(\tau_c < \sigma_n) + \frac{1}{1/\epsilon - 1/n} \mathbb{P}(\sigma_n \leq \tau_c). \quad (\text{B.28})$$

Taking $n \rightarrow \infty$: $\mathbb{P}(\tau_c < \sigma_n) \rightarrow \mathbb{P}(\tau_c < \infty) = p(\omega_0)$ by monotone continuity of probabilities on the increasing events $\{\tau_c < \sigma_n\} \uparrow \{\tau_c < \infty\}$, while $\mathbb{P}(\sigma_n \leq \tau_c) \rightarrow \mathbb{P}(\tau_c = \infty) = 1 - p(\omega_0)$. The factor $1/(1/\epsilon - 1/n) \rightarrow \epsilon$. Both summands in (B.28) are non-negative and uniformly bounded by $\max(1/\omega_c, \epsilon) + 1$, so dominated convergence applies¹⁰:

$$\frac{1}{\omega_0} = \frac{1}{\omega_c} p(\omega_0) + \epsilon (1 - p(\omega_0)).$$

Solving for $p(\omega_0)$ yields (B.26).

(b) *Dirichlet problem for the scale function.* By the standard connection between hitting probabilities and the Dirichlet problem (Karatzas and Shreve, 1991, Proposition 5.5.22), the function $p(\omega) \equiv \mathbb{P}(\tau_c < \tau_{1/\epsilon} \mid \omega_0 = \omega)$ is the unique bounded continuous solution of

$$\mathcal{L}p(\omega) = 0 \text{ on } I, \quad p(\omega_c) = 1, \quad \lim_{\omega \uparrow 1/\epsilon} p(\omega) = 0. \quad (\text{B.29})$$

The second boundary condition uses the Feller classification: since $1/\epsilon$ is unattainable (Ap-

¹⁰Equivalently, the uniform integrability of $\{Q_{\tau_n}\}_{n \geq 1}$ follows from $Q_{\tau_n} \leq 1/\omega_c$; see Karatzas and Shreve (1991, Ch. 1) for the standard passage-to-the-limit in optional stopping.

pendix B.9) and the process started near $1/\epsilon$ is pushed away from it by the positive drift σ_ω^2/ω , the probability of hitting ω_c before hitting an arbitrarily small neighbourhood of $1/\epsilon$ vanishes as $\omega_0 \uparrow 1/\epsilon$; equivalently, $\mathbb{P}(\tau_c < \infty \mid \omega_0 \uparrow 1/\epsilon) \rightarrow 0$ by Appendix B.10. By §B.8.1, the general solution of $\mathcal{L}p = 0$ is $p(\omega) = A + B/\omega$. Imposing $p(1/\epsilon) = A + B\epsilon = 0$ gives $A = -B\epsilon$; imposing $p(\omega_c) = A + B/\omega_c = 1$ then gives $B = 1/(1/\omega_c - \epsilon)$ and $A = -\epsilon/(1/\omega_c - \epsilon)$, so

$$p(\omega) = \frac{1/\omega - \epsilon}{1/\omega_c - \epsilon},$$

which is (B.26). Both derivations yield the same formula; monotonicity in ω_0 is immediate from $\partial p / \partial \omega_0 = -1/[\omega_0^2(1/\omega_c - \epsilon)] < 0$. $\square$

Remark B.5 (Boundary-condition check) At $\omega_0 = \omega_c$: $p(\omega_c) = (1/\omega_c - \epsilon)/(1/\omega_c - \epsilon) = 1$. At $\omega_0 = 1/\epsilon$: numerator $1/\omega_0 - \epsilon = 0$, so $p(1/\epsilon) = 0$. Both limits are exactly what the economics requires: starting at the crisis threshold we hit it immediately; starting at the frictionless steady state we never hit it (cf. $Q_t \rightarrow \epsilon$ a.s. in Appendix B.10).

B.8.3 Unconditional Expected Hitting Time

Theorem B.3 (Infinite unconditional expected hitting time) For every $\omega_0 \in I = (\omega_c, 1/\epsilon)$,

$$\mathbb{E}[\tau_c \mid \omega_0] = +\infty. \quad (\text{B.30})$$

Proof. By Proposition B.4, $p(\omega_0) < 1$ for $\omega_0 \in (\omega_c, 1/\epsilon)$, so $\mathbb{P}(\tau_c = \infty \mid \omega_0) = 1 - p(\omega_0) > 0$. Therefore

$$\mathbb{E}[\tau_c \mid \omega_0] \geq \mathbb{E}[\tau_c \mathbf{1}_{\{\tau_c = \infty\}} \mid \omega_0] = (+\infty) \cdot \mathbb{P}(\tau_c = \infty \mid \omega_0) = +\infty,$$

using the convention $0 \cdot \infty = 0$ but $(+\infty) \cdot c = +\infty$ for $c > 0$. $\square$

The economic content is the same as in Appendix B.10: conditional on avoiding the crisis, $\omega_t \rightarrow 1/\epsilon$ a.s. but the expected time to reach any neighbourhood of $1/\epsilon$ is infinite; hence the expected time to the *crisis* threshold, which includes a positive-probability event on which $\tau_c = \infty$, is a fortiori infinite.

B.8.4 Conditional Expected Hitting Time

Proposition B.5 (Conditional expected hitting time) *Let $v(\omega_0) \equiv \mathbb{E}[\tau_c \mathbf{1}_{\{\tau_c < \infty\}} \mid \omega_0]$ for $\omega_0 \in \bar{I}$. Then v is the unique bounded solution of the Feynman–Kac boundary-value problem*

$$\mathcal{L}v(\omega) = -p(\omega), \quad \omega \in I; \quad v(\omega_c) = v(1/\epsilon) = 0. \quad (\text{B.31})$$

In the scale coordinate $y \equiv 1/\omega$, equivalently $V(y) \equiv v(1/y)$, the generator transforms into a pure-diffusion operator $\tilde{\mathcal{L}}V(y) = \frac{1}{2}a(y)V''(y)$ with $a(y) \equiv y^4\sigma_\omega(1/y)^2$, and (B.31) reduces to

$$V''(y) = -\frac{2(y-\epsilon)}{a(y)(y_c-\epsilon)}, \quad V(\epsilon) = V(y_c) = 0, \quad y_c \equiv 1/\omega_c. \quad (\text{B.32})$$

The unique bounded solution of (B.32) admits the Green's-function representation

$$V(y) = \frac{1}{y_c - \epsilon} \int_{\epsilon}^{y_c} G(y, z)(z - \epsilon)m(z) dz, \quad m(z) \equiv \frac{2}{a(z)}, \quad (\text{B.33})$$

where the Dirichlet Green's function on $[\epsilon, y_c]$ for the operator $-d^2/dy^2$ is

$$G(y, z) = \begin{cases} \frac{(y-\epsilon)(y_c-z)}{y_c-\epsilon}, & y \leq z, \\ \frac{(z-\epsilon)(y_c-y)}{y_c-\epsilon}, & z \leq y. \end{cases} \quad (\text{B.34})$$

Equivalently, in the original variable ω_0 ,

$$v(\omega_0) = \frac{2}{1/\omega_c - \epsilon} \int_{\omega_c}^{1/\epsilon} G\left(\frac{1}{\omega_0}, \frac{1}{\omega'}\right) \frac{(1-\epsilon\omega')\omega'}{\sigma_\omega(\omega')^2} d\omega'. \quad (\text{B.35})$$

The conditional expected hitting time is then $\mathbb{E}[\tau_c \mid \omega_0, \tau_c < \infty] = v(\omega_0)/p(\omega_0)$.

Proof. *Step 1: Feynman–Kac characterisation.* The process $\tau_c \mathbf{1}_{\{\tau_c < \infty\}}$ is non-negative and finite on the event $\{\tau_c < \infty\}$, and $\tau_c \leq \tau_n + \infty \cdot \mathbf{1}_{\{\tau_c \geq \tau_n\}}$ with τ_n the exit time from I_n (cf. proof of Proposition B.4). By the standard Feynman–Kac formula for exit problems (Karatzas and Shreve, 1991, Theorem 5.7.6), applied on each I_n and passed to the limit using dominated convergence on the event $\{\tau_c < \infty\}$, v solves $\mathcal{L}v = -p$ on I with $v(\omega_c) = 0$ and $v(1/\epsilon) = 0$, the latter because $\mathbb{P}(\tau_c < \infty \mid \omega_0) \rightarrow 0$ as $\omega_0 \uparrow 1/\epsilon$ by Proposition B.4. Boundedness of v on $\bar{I}$ follows from the boundedness of the Green's-function representation established below.

Step 2: Change of variables $y = 1/\omega$. For $V(y) = v(1/y)$, chain rule gives $v'(\omega) = -y^2 V'(y)$

and $v''(\omega) = y^4 V''(y) + 2y^3 V'(y)$ (evaluated at $y = 1/\omega$). Substituting into (B.24),

$$\begin{aligned}\mathcal{L}v(\omega) &= \frac{\sigma_\omega^2}{\omega}(-y^2 V'(y)) + \frac{\sigma_\omega^2}{2}(y^4 V''(y) + 2y^3 V'(y)) \\ &= \sigma_\omega^2(-y^3 V'(y) + y^3 V'(y)) + \frac{\sigma_\omega^2}{2}y^4 V''(y) = \frac{y^4 \sigma_\omega^2}{2} V''(y) = \frac{a(y)}{2} V''(y),\end{aligned}$$

using $1/\omega = y$ so $\sigma_\omega^2/\omega \cdot y^2 = \sigma_\omega^2 \cdot y^3$. The first-derivative terms cancel—this is precisely the statement that $s(\omega) = 1/\omega$ is a scale function, so the change of variables to the scale coordinate eliminates drift. Writing $\mathcal{L}v = -p$ with $p(\omega) = (y - \epsilon)/(y_c - \epsilon)$ (from Proposition B.4, after substitution $1/\omega = y, 1/\omega_c = y_c$):

$$\frac{a(y)}{2} V''(y) = -\frac{y - \epsilon}{y_c - \epsilon} \implies V''(y) = -\frac{2(y - \epsilon)}{a(y)(y_c - \epsilon)},$$

with boundary conditions $V(\epsilon) = v(1/\epsilon) = 0$ and $V(y_c) = v(\omega_c) = 0$, establishing (B.32).

Step 3: Green's-function representation. The ODE (B.32) is $-V''(y) = f(y)$ on $[\epsilon, y_c]$ with Dirichlet data, where $f(y) \equiv 2(y - \epsilon)/[a(y)(y_c - \epsilon)]$. The Dirichlet Green's function for $-d^2/dy^2$ on $[\epsilon, y_c]$ is (B.34) (verified by direct differentiation: $G(y, z)$ is continuous, piecewise linear in y , vanishes at $y = \epsilon$ and $y = y_c$, and has jump $[\partial_y G]_{y=z} = -1$, making $-\partial_{yy} G(\cdot, z) = \delta_z$ in the distributional sense). Hence

$$V(y) = \int_\epsilon^{y_c} G(y, z) f(z) dz = \frac{1}{y_c - \epsilon} \int_\epsilon^{y_c} G(y, z) (z - \epsilon) m(z) dz,$$

with $m(z) = 2/a(z)$, giving (B.33). Uniqueness of the bounded solution follows from the maximum principle for $-V'' = f$ with homogeneous Dirichlet data.

Step 4: Transformation to the ω variable. In (B.33) substitute $z = 1/\omega'$, so that $dz = -d\omega'/\omega'^2$ and z ranges over $[\epsilon, y_c]$ iff ω' ranges over $[\omega_c, 1/\epsilon]$ (in reversed order). Compute each factor:

$$\begin{aligned}z - \epsilon &= \frac{1}{\omega'} - \epsilon = \frac{1 - \epsilon\omega'}{\omega'}, \\ a(z) &= z^4 \sigma_\omega(1/z)^2 = \omega'^{-4} \sigma_\omega(\omega')^2, \\ m(z) &= \frac{2}{a(z)} = \frac{2\omega'^4}{\sigma_\omega(\omega')^2}, \\ dz &= -\frac{d\omega'}{\omega'^2}.\end{aligned}$$

Therefore

$$(z - \epsilon)m(z) dz = \frac{1 - \epsilon\omega'}{\omega'} \cdot \frac{2\omega'^4}{\sigma_\omega(\omega')^2} \cdot \left(-\frac{d\omega'}{\omega'^2}\right) = -\frac{2(1 - \epsilon\omega')\omega'}{\sigma_\omega(\omega')^2} d\omega'.$$

The reversal of orientation (since as z increases from ϵ to y_c , $\omega' = 1/z$ decreases from $1/\epsilon$ to ω_c) cancels the leading minus sign, and after substituting $y = 1/\omega_0$, $y_c - \epsilon = 1/\omega_c - \epsilon$ we obtain (B.35). Note the final power of ω' : the Jacobian $dz/d\omega' = -\omega'^{-2}$ contributes ω'^{-2} , and the speed-measure transformation contributes $\omega'^{+4}/\omega' = \omega'^{+3}$, so the net power is ω'^{+1} , not ω'^{-3} .¹¹ $\square$

B.8.5 Closed-Form Case: $\varphi = 0$

Proposition B.6 (Elementary closed form at $\varphi = 0$) *When the Frisch elasticity $\varphi = 0$,*

$$\sigma_\omega(\omega)\Big|_{\varphi=0} = \frac{\sigma\epsilon(1 - \omega)(1/\epsilon - \omega)}{\epsilon - 1}, \quad (\text{B.36})$$

and the weight in (B.35) reduces to

$$\left.\frac{(1 - \epsilon\omega')\omega'}{\sigma_\omega(\omega')^2}\right|_{\varphi=0} = \frac{(\epsilon - 1)^2}{\sigma^2\epsilon} \cdot \frac{\omega'}{(1 - \omega')^2(1/\epsilon - \omega')}. \quad (\text{B.37})$$

In particular, $v(\omega_0)$ is an elementary function of ω_0 : it reduces to a sum of logarithms and rational terms obtained by partial-fraction decomposition of (B.37).

Proof. At $\varphi = 0$, the denominator of (B.23) equals $(1 - \omega) - (1/\epsilon - \omega) = 1 - 1/\epsilon = (\epsilon - 1)/\epsilon$, which is independent of ω . The numerator is $\sigma(1 - \omega)(1/\epsilon - \omega)$ at $\varphi = 0$, so (B.36) follows.

For (B.37), use $1 - \epsilon\omega' = \epsilon(1/\epsilon - \omega')$:

$$\frac{(1 - \epsilon\omega')\omega'}{\sigma_\omega^2} = \frac{\epsilon(1/\epsilon - \omega')\omega'(\epsilon - 1)^2}{\sigma^2\epsilon^2(1 - \omega')^2(1/\epsilon - \omega')^2} = \frac{(\epsilon - 1)^2}{\sigma^2\epsilon} \cdot \frac{\omega'}{(1 - \omega')^2(1/\epsilon - \omega')}.$$

Partial fractions decomposes $\omega'/[(1 - \omega')^2(1/\epsilon - \omega')]$ on the intervals of integration defining

¹¹Numerical sanity check (this is the case described above Result 5 of the working notes). Take $\epsilon = 2$, $\omega_c = 1/4$ (so $y_c = 4$), $\sigma_\omega \equiv 1$ (so $a(y) = y^4$), $\omega_0 = 1/3$ (so $y_0 = 3$). Direct solution of (B.32): $V''(y) = -(y - 2)/y^4$, integrating twice gives $V(y) = -\frac{1}{2y} + \frac{1}{3y^2} + C_1y + C_2$, and imposing $V(2) = V(4) = 0$ yields $C_1 = -1/32$, $C_2 = 11/48$, hence $V(3) = -1/6 + 1/27 - 3/32 + 11/48 = 5/864$. Green-function formula (B.33): with $G(3, z) = (4 - z)/2$ for $z \geq 3$ and $G(3, z) = (z - 2)/2$ for $z \leq 3$, $m(z) = 2/z^4$, one computes $\int_2^3 (z - 2)^2 z^{-4} dz = 1/162$, $\int_3^4 (4 - z)(z - 2)z^{-4} dz = 7/1296$, giving $V(3) = \frac{1}{2}(1/162 + 7/1296) = 5/864$. The ω -form (B.35) reproduces the same number under the substitution $z = 1/\omega'$; a formula with the incorrect power ω'^{-3} would differ by a factor of ω'^4 inside the integrand and fail this test.

the Green-function representation:

$$\frac{\omega'}{(1 - \omega')^2(1/\epsilon - \omega')} = \frac{A}{1 - \omega'} + \frac{B}{(1 - \omega')^2} + \frac{C}{1/\epsilon - \omega'}, \quad (\text{B.38})$$

with coefficients found by the standard cover-up method: clearing denominators, $\omega' = A(1 - \omega')(1/\epsilon - \omega') + B(1/\epsilon - \omega') + C(1 - \omega')^2$. Evaluating at $\omega' = 1$: $1 = B(1/\epsilon - 1) = -B(\epsilon - 1)/\epsilon$, so $B = -\epsilon/(\epsilon - 1)$. Evaluating at $\omega' = 1/\epsilon$: $1/\epsilon = C(1 - 1/\epsilon)^2 = C(\epsilon - 1)^2/\epsilon^2$, so $C = \epsilon/(\epsilon - 1)^2$. Comparing the coefficient of ω'^2 on both sides ($A + C = 0$) gives $A = -C = -\epsilon/(\epsilon - 1)^2$.

Each term in (B.38) has an elementary antiderivative ($\log$, $1/(1 - \omega')$, $\log$), and in combination with the piecewise-linear Green function (B.34) the full integral (B.35) is elementary—a finite sum of logarithms and rational functions of ω_0 . We omit the explicit expression, which is long but straightforward. $\square$

This is the only regime in which the paper exhibits a fully explicit $v(\omega_0)$ without numerical integration. For $\varphi > 0$ the denominator $(1 + \varphi)(1 - \omega) - (1/\epsilon - \omega) = \varphi(1 - \omega) + (1 - 1/\epsilon)$ is ω -dependent, and σ_ω^{-2} is no longer a rational function with poles at $\{1, 1/\epsilon\}$ only; closed-form antiderivatives are not available.

B.8.6 Sanity Check: Constant-Diffusion Brownian Limit

Suppose, contrafactually, that $a(y) \equiv a_0$ is constant on $[\epsilon, y_c]$ —this is not the physical case in our model, but it serves as a benchmark for the Green-function formula. Then (B.32) becomes $V''(y) = -2(y - \epsilon)/[a_0(y_c - \epsilon)]$. Integrating twice and imposing $V(\epsilon) = V(y_c) = 0$:

$$V(y) = -\frac{(y - \epsilon)^3}{3a_0(y_c - \epsilon)} + C_1(y - \epsilon), \quad C_1 = \frac{y_c - \epsilon}{3a_0},$$

which simplifies to

$$V(y) = \frac{(y - \epsilon)(y_c - y)(y_c + y - 2\epsilon)}{3a_0(y_c - \epsilon)}. \quad (\text{B.39})$$

This is the classical cubic expression for the expected exit time of a Brownian motion with constant diffusion from a finite interval, weighted by the absorption probability at the left endpoint. It vanishes at $y = \epsilon$ and $y = y_c$, is positive on (ϵ, y_c) , and is maximised at $y = (y_c + \epsilon)/2$; all three properties are expected. In particular, (B.39) agrees with the Green-function evaluation of (B.33) when $m(z) = 2/a_0$: $\int_\epsilon^{y_c} G(y, z)(z - \epsilon)dz$ splits into two pieces which, after elementary integration, collapse to the product $(y - \epsilon)(y_c - y)(y_c + y - 2\epsilon)/6$, and the prefactor $2/[a_0(y_c - \epsilon)]$

delivers (B.39). The formula is right.

B.8.7 Tail Distribution of τ_c

Propositions B.4–B.6 pin down the probability of an eventual crisis and a Green’s-function representation of the conditional mean. Policy questions, however, concern *horizon-specific* tail probabilities $\mathbb{P}(\tau_c \leq T \mid \omega_0)$ at $T \in \{5, 10, 25\}$ years. This subsection gives the Laplace-transform PDE characterising the full distribution of τ_c (Proposition B.7), solves it in closed form at $\varphi = 0$ (Proposition B.8), and provides matching short- T and long- T asymptotic laws (Propositions B.9 and B.10) together with a numerical table at the paper’s empirical calibration (Remark B.9).

B.8.7.1 Laplace transform: backward equation

Proposition B.7 (Kolmogorov backward equation for the Laplace transform) *For $\lambda \geq 0$ and $\omega \in \bar{I} = [\omega_c, 1/\epsilon]$, define*

$$f(\omega; \lambda) \equiv \mathbb{E}\left[e^{-\lambda\tau_c} \mathbf{1}_{\{\tau_c < \infty\}} \mid \omega_0 = \omega\right]. \quad (\text{B.40})$$

Then $f(\cdot; \lambda)$ is the unique bounded continuous solution of

$$\mathcal{L}f(\omega; \lambda) = \lambda f(\omega; \lambda), \quad \omega \in I, \quad (\text{B.41})$$

with boundary data $f(\omega_c; \lambda) = 1$ and $f(1/\epsilon; \lambda) = 0$. In particular, $f(\omega; 0) = p(\omega)$ recovers (B.26), and conditional moments on $\{\tau_c < \infty\}$ are given by

$$\mathbb{E}\left[\tau_c^n \mathbf{1}_{\{\tau_c < \infty\}} \mid \omega\right] = (-1)^n \partial_\lambda^n f(\omega; \lambda) \Big|_{\lambda=0}, \quad n \in \mathbb{N}. \quad (\text{B.42})$$

Proof. Fix $\lambda \geq 0$ and let $\tau_n \equiv \tau_c \wedge \sigma_n$ with $\sigma_n \equiv \inf\{t : \omega_t = 1/\epsilon - 1/n\}$ as in the proof of Proposition B.4. For a smooth $g : \bar{I} \rightarrow \mathbb{R}$, Itô’s formula (Karatzas and Shreve, 1991, Theorem 3.3.6) applied to $M_t \equiv e^{-\lambda t} g(\omega_t)$ gives

$$dM_t = e^{-\lambda t} [\mathcal{L}g(\omega_t) - \lambda g(\omega_t)] dt + e^{-\lambda t} \sigma_\omega(\omega_t) g'(\omega_t) dZ_t.$$

Stop at $t \wedge \tau_n$: the stochastic integral is a continuous martingale (bounded integrand), and $\mathcal{L}g - \lambda g \equiv 0$ on I forces $M_{t \wedge \tau_n}$ to be a martingale whenever g solves (B.41). By optional

stopping (Karatzas and Shreve, 1991, Theorem 1.3.22),

$$g(\omega_0) = \mathbb{E}[e^{-\lambda\tau_n} g(\omega_{\tau_n})] = \mathbb{E}[e^{-\lambda\tau_c} g(\omega_c) \mathbf{1}_{\{\tau_c < \sigma_n\}}] + \mathbb{E}[e^{-\lambda\sigma_n} g(1/\epsilon - 1/n) \mathbf{1}_{\{\sigma_n \leq \tau_c\}}]. \quad (\text{B.43})$$

Passing $n \rightarrow \infty$: the first term converges to $g(\omega_c) \cdot \mathbb{E}[e^{-\lambda\tau_c} \mathbf{1}_{\{\tau_c < \infty\}}]$ by monotone-increasing events and dominated convergence ($|e^{-\lambda\tau_c}| \leq 1$); the second converges to $g(1/\epsilon) \cdot \mathbb{E}[e^{-\lambda\sigma} \mathbf{1}_{\{\sigma < \infty\}}]$ where σ is the hitting time of $1/\epsilon$, and on $\{\tau_c = \infty\} \cap \{\omega_t \rightarrow 1/\epsilon\}$ (full measure by Appendix B.10), $\sigma = +\infty$ so $e^{-\lambda\sigma} = 0$ for $\lambda > 0$, while for $\lambda = 0$ the limit is $g(1/\epsilon) \cdot (1 - p(\omega_0))$.

Imposing $g(\omega_c) = 1$ and $g(1/\epsilon) = 0$ in (B.43) and taking $n \rightarrow \infty$ yields, for $\lambda > 0$,

$$g(\omega_0) = \mathbb{E}[e^{-\lambda\tau_c} \mathbf{1}_{\{\tau_c < \infty\}}] = f(\omega_0; \lambda),$$

which proves any bounded solution of (B.41) equals $f(\cdot; \lambda)$. Existence follows from the general Feynman–Kac representation (Karatzas and Shreve, 1991, Theorem 5.7.6), Revuz–Yor (1999) Ch. VII. At $\lambda = 0$, (B.41) is the Dirichlet problem of Proposition B.4; boundedness and the Feller classification of $1/\epsilon$ as unattainable give $f(1/\epsilon; 0) = \lim_{\omega \uparrow 1/\epsilon} p(\omega) = 0$. Moment formula (B.42) follows by differentiating (B.40) under the expectation (justified on $\{\tau_c < \infty\}$ by dominated convergence with envelope $\tau_c^n e^{-\lambda_0 \tau_c} \leq C_n$ on the event). $\square$

Differentiating (B.41) with respect to λ and evaluating at $\lambda = 0$ gives $\mathcal{L}[\partial_\lambda f(\cdot; 0)] = f(\cdot; 0) = p$, so $-\partial_\lambda f(\omega; 0)$ solves the Feynman–Kac problem (B.31) of Proposition B.5: the conditional expected hitting time is indeed the $n = 1$ moment in (B.42), consistent with the earlier Green’s-function representation.

B.8.7.2 Closed form at $\varphi = 0$: hypergeometric reduction

Proposition B.8 (Closed-form Laplace transform at $\varphi = 0$) *At $\varphi = 0$ (so that σ_ω is given by (B.36)), the Laplace transform (B.40) admits the elementary closed form*

$$f(\omega; \lambda)|_{\varphi=0} = \frac{\omega_c}{\omega} \left(\frac{1 - \omega}{1 - \omega_c} \right)^{p_-} \left(\frac{1 - \epsilon\omega}{1 - \epsilon\omega_c} \right)^{p_+}, \quad (\text{B.44})$$

where

$$p_\pm = \frac{1}{2} \left(1 \pm \sqrt{1 + 8\lambda/\sigma^2} \right), \quad p_+ + p_- = 1, \quad p_+ p_- = -2\lambda/\sigma^2. \quad (\text{B.45})$$

Proof. *Step 1: reduction in the Q -coordinate.* With $Q = 1/\omega$ and $V(Q) \equiv f(1/Q; \lambda)$, the calculation of Step 2 in the proof of Proposition B.5 gives $\mathcal{L}f = \frac{1}{2}a(Q)V''(Q)$ where $a(Q) = Q^4\sigma_\omega(1/Q)^2$. At

$\varphi = 0$, substituting $1 - \omega = (Q - 1)/Q$ and $1/\epsilon - \omega = (Q - \epsilon)/(\epsilon Q)$ into (B.36) yields

$$a(Q)\big|_{\varphi=0} = \frac{\sigma^2(Q-1)^2(Q-\epsilon)^2}{(\epsilon-1)^2}. \quad (\text{B.46})$$

Equation (B.41) becomes

$$\frac{\sigma^2(Q-1)^2(Q-\epsilon)^2}{2(\epsilon-1)^2} V''(Q) = \lambda V(Q), \quad V(Q_c) = 1, \quad V(\epsilon) = 0, \quad (\text{B.47})$$

where $Q_c = 1/\omega_c$. This is a second-order linear ODE with three regular singular points at $Q \in \{1, \epsilon, \infty\}$ —a hypergeometric type.

Step 2: exact factorisation. We claim that $V(Q) = (Q-1)^\alpha(Q-\epsilon)^\beta$ solves (B.47) iff $\alpha + \beta = 1$ and $\alpha(\alpha-1) = \mu \equiv 2\lambda/\sigma^2$. Indeed, for V of this form,

$$\frac{V''}{V} = \frac{\alpha(\alpha-1)}{(Q-1)^2} + \frac{\beta(\beta-1)}{(Q-\epsilon)^2} + \frac{2\alpha\beta}{(Q-1)(Q-\epsilon)}$$

(from $(V'/V)^2 + \partial_Q(V'/V) = V''/V$ with $V'/V = \alpha/(Q-1) + \beta/(Q-\epsilon)$). Multiplying by $(Q-1)^2(Q-\epsilon)^2$, the result is a polynomial in Q of degree 2; its coefficient of Q^2 is $(\alpha+\beta)^2 - (\alpha+\beta)$, vanishing iff $\alpha + \beta \in \{0, 1\}$; its coefficient of Q vanishes iff $\alpha + \beta = 1$ (the alternative $\alpha + \beta = 0$ leaves a non-zero coefficient proportional to $\epsilon\alpha + \beta$); its constant term, after imposing $\beta = 1 - \alpha$, equals $\alpha(\alpha-1)(\epsilon-1)^2$. Substituting back into (B.47): $\alpha(\alpha-1)(\epsilon-1)^2 = 2\lambda(\epsilon-1)^2/\sigma^2$, i.e., $\alpha(\alpha-1) = \mu$. The two roots are $\alpha = p_+$ and $\alpha = p_- = 1 - p_+$, giving the two linearly independent solutions

$$V_+(Q) = (Q-1)^{p_-}(Q-\epsilon)^{p_+}, \quad V_-(Q) = (Q-1)^{p_+}(Q-\epsilon)^{p_-}. \quad (\text{B.48})$$

Step 3: selection by boundary behaviour. For $\lambda > 0$, $p_+ > 1 > 0 > p_-$. At $Q = \epsilon$, V_+ vanishes (as $(Q-\epsilon)^{p_+} \rightarrow 0$) while V_- diverges (as $(Q-\epsilon)^{p_-} \rightarrow +\infty$ since $p_- < 0$). Boundedness at $\omega = 1/\epsilon$ (equivalently $Q = \epsilon$) forces the coefficient of V_- to vanish. The condition $V(Q_c) = 1$ then fixes the coefficient of V_+ , giving

$$V(Q) = \frac{(Q-1)^{p_-}(Q-\epsilon)^{p_+}}{(Q_c-1)^{p_-}(Q_c-\epsilon)^{p_+}}. \quad (\text{B.49})$$

Switching back to ω : $Q-1 = (1-\omega)/\omega$, $Q-\epsilon = (1-\epsilon\omega)/\omega$, so $(Q-1)^{p_-}(Q-\epsilon)^{p_+} = \omega^{-(p_++p_-)}(1-\omega)^{p_-}(1-\epsilon\omega)^{p_+} = \omega^{-1}(1-\omega)^{p_-}(1-\epsilon\omega)^{p_+}$. Dividing by the analogous expression at Q_c :

$$f(\omega; \lambda)\big|_{\varphi=0} = \frac{\omega_c}{\omega} \cdot \frac{(1-\omega)^{p_-}(1-\epsilon\omega)^{p_+}}{(1-\omega_c)^{p_-}(1-\epsilon\omega_c)^{p_+}},$$

which is (B.44).

Step 4: relation to the Gauss hypergeometric. The same closed form emerges from the Möbius transformation $\xi = (Q - \epsilon)/(Q - 1)$ (mapping the singular points $Q \in \{\epsilon, 1, \infty\}$ to $\xi \in \{0, \infty, 1\}$) followed by the substitution $V = \xi^{p_+}(1 - \xi)^{-1}$: the residual function $W(\xi)$ satisfies the Gauss hypergeometric equation

$$\xi(1 - \xi)W''(\xi) + [2p_+ - 2(p_+ + 1)\xi]W'(\xi) - 2p_+ W(\xi) = 0,$$

with parameters $(a, b, c) = (1, 2p_+, 2p_+)$, whose principal solution is ${}_2F_1(1, 2p_+; 2p_+; \xi) = (1 - \xi)^{-1}$ (Abramowitz and Stegun, 1972, Ch. 15). The degeneracy $b = c$ collapses the hypergeometric series to the elementary form (B.49). $\square$

Remark B.6 (Four consistency checks) Formula (B.44) passes each of the four independent checks listed in the working notes.

- (i) $f(\omega_c; \lambda) = 1$: all three ratios equal one.
- (ii) $f(1/\epsilon; \lambda) = 0$: the factor $(1 - \epsilon\omega)^{p_+}$ vanishes since $1 - \epsilon \cdot (1/\epsilon) = 0$ and $p_+ > 0$.
- (iii) $f(\omega; 0) = p(\omega)$: at $\lambda = 0$, $p_+ = 1$ and $p_- = 0$, so (B.44) reduces to

$$f(\omega; 0)|_{\varphi=0} = \frac{\omega_c}{\omega} \cdot 1 \cdot \frac{1 - \epsilon\omega}{1 - \epsilon\omega_c} = \frac{\omega_c(1 - \epsilon\omega)}{\omega(1 - \epsilon\omega_c)},$$

matching (B.26).

- (iv) $-\partial_\lambda f(\omega; 0)$ solves Proposition B.5. From $\partial_\lambda p_\pm|_{\lambda=0} = \pm 2/\sigma^2$ and $\partial_\lambda \ln f = (\partial_\lambda p_-) \ln[(1 - \omega)/(1 - \omega_c)] + (\partial_\lambda p_+) \ln[(1 - \epsilon\omega)/(1 - \epsilon\omega_c)]$,

$$\mathbb{E}[\tau_c \mathbf{1}_{\{\tau_c < \infty\}} | \omega]|_{\varphi=0} = -\frac{2p(\omega)}{\sigma^2} \ln \left[\frac{(1 - \omega_c)(1 - \epsilon\omega)}{(1 - \omega)(1 - \epsilon\omega_c)} \right], \quad (\text{B.50})$$

an explicit log-ratio formula that replaces the Green's-function integral of Proposition B.5 (and is checked independently by direct Feynman–Kac substitution: $\mathcal{L}$ applied to (B.50) yields $-p(\omega)$ after elementary but tedious algebra, with correct boundary conditions $v(\omega_c) = v(1/\epsilon) = 0$).

B.8.7.3 Short-horizon tail via Freidlin–Wentzell large deviations

Proposition B.9 (Small-noise / short- T large-deviation principle) Let $\tilde{\sigma}_\omega(\omega) \equiv \sigma_\omega(\omega)/\sigma$, which is independent of σ . As $\sigma \downarrow 0$ (equivalently, $T \downarrow 0$ at fixed (ω_0, ω_c)), the Freidlin–Wentzell large-deviation

principle (Freidlin and Wentzell, 2012, Ch. 5) gives

$$\log \mathbb{P}(\tau_c \leq T \mid \omega_0) = -\frac{I(\omega_0, \omega_c, T)}{\sigma^2} + o(\sigma^{-2}), \quad (\text{B.51})$$

where the rate function is the action

$$I(\omega_0, \omega_c, T) = \inf_{\phi \in \mathcal{A}} \frac{1}{2} \int_0^T \frac{(\dot{\phi}(s) - \sigma^2 \tilde{\sigma}_\omega(\phi(s))^2 / \phi(s))^2}{\tilde{\sigma}_\omega(\phi(s))^2} ds, \quad (\text{B.52})$$

with $\mathcal{A} = \{\phi : [0, T] \rightarrow [\omega_c, 1/\epsilon] \text{ absolutely continuous, } \phi(0) = \omega_0, \phi(T) = \omega_c\}$. In the leading-order small-noise limit (dropping the $O(\sigma^2)$ drift),

$$I(\omega_0, \omega_c, T) \approx \tilde{I}(\omega_0, \omega_c, T) \equiv \frac{d(\omega_0, \omega_c)^2}{2T}, \quad d(\omega_0, \omega_c) \equiv \int_{\omega_c}^{\omega_0} \frac{du}{\tilde{\sigma}_\omega(u)}, \quad (\text{B.53})$$

where $d(\cdot, \omega_c)$ is the Riemannian distance in the metric $ds^2 = d\omega^2 / \tilde{\sigma}_\omega(\omega)^2$. Hence, for small T ,

$$\log \mathbb{P}(\tau_c \leq T \mid \omega_0) \lesssim -\frac{d(\omega_0, \omega_c)^2}{2T\sigma^2}. \quad (\text{B.54})$$

Proof. Writing $d\omega_t = \sigma^2 \tilde{\sigma}_\omega^2 / \omega dt + \sigma \tilde{\sigma}_\omega dZ_t$ casts the SDE in small-noise form $dX = b(X)dt + \sigma \sigma_X(X)dW$ with $b = O(\sigma^2)$. The Freidlin–Wentzell LDP (Dembo and Zeitouni, 2010; Freidlin and Wentzell, 2012, Theorem 5.6.7) then gives (B.51) with rate function $I(\omega_0, \omega_c, T) = \inf_{\phi} \frac{1}{2} \int_0^T (\dot{\phi} - b(\phi))^2 / \tilde{\sigma}_\omega(\phi)^2 ds$. For the leading-order analysis, drop the drift (which is $O(\sigma^2)$ and hence subdominant in $\sigma \downarrow 0$): the reduced action $\tilde{I} = \frac{1}{2} \int_0^T \dot{\phi}^2 / \tilde{\sigma}_\omega(\phi)^2 ds$ is the length-functional in the Riemannian metric $ds^2 = d\omega^2 / \tilde{\sigma}_\omega^2$, and by Cauchy–Schwarz

$$\int_0^T \frac{\dot{\phi}^2}{\tilde{\sigma}_\omega^2} ds \geq \frac{1}{T} \left(\int_0^T \frac{|\dot{\phi}|}{\tilde{\sigma}_\omega} ds \right)^2 \geq \frac{1}{T} \left(\int_{\omega_c}^{\omega_0} \frac{du}{\tilde{\sigma}_\omega(u)} \right)^2 = \frac{d(\omega_0, \omega_c)^2}{T},$$

with equality achieved by the constant-speed geodesic $\phi^\star$ with $|\dot{\phi}^\star| / \tilde{\sigma}_\omega(\phi^\star) \equiv d/T$. Hence $\tilde{I} = d(\omega_0, \omega_c)^2 / (2T)$. The Euler–Lagrange equations for the full action (B.52) are

$$\frac{d}{ds} \left(\frac{\dot{\phi} - b(\phi)}{\tilde{\sigma}_\omega(\phi)^2} \right) + \frac{(\dot{\phi} - b(\phi))^2 \tilde{\sigma}'_\omega(\phi)}{\tilde{\sigma}_\omega(\phi)^3} + \frac{(\dot{\phi} - b(\phi)) b'(\phi)}{\tilde{\sigma}_\omega(\phi)^2} = 0,$$

a second-order ODE for $\phi^\star$; for T small or σ small, the leading-order solution is the unit-speed geodesic in the $\tilde{\sigma}_\omega$ metric. $\square$

Remark B.7 (Validity of the short- T bound) The approximation (B.54) is sharp in the joint limit

$(\sigma, T) \rightarrow (0, 0)$ with $\sigma^2 T$ fixed. Because the drift $\mu_\omega = \sigma_\omega^2 / \omega$ is strictly positive on I (pushing the process away from ω_c), including the drift lowers the probability of early absorption; hence (B.54) is an upper bound on $\mathbb{P}(\tau_c \leq T)$ to leading order in σ^2 . For policy-horizon $T \in \{10, 25\}$ years at $\sigma = 0.1505$, the drift is non-negligible and (B.54) overestimates the tail; numerical inversion of the exact Laplace transform (Proposition B.7) is required for quantitative answers.

B.8.7.4 Long-horizon tail: essential-spectrum bottom

The natural endpoint $1/\epsilon$ does *not* generate a discrete spectrum. The correct statement is that $-\mathcal{L}$ has empty discrete spectrum below the universal threshold $\sigma^2/8$, which is the bottom of the essential spectrum. The argument is a standard one-dimensional Schrödinger-type boundary analysis at the natural endpoint, combined with the closed form of Proposition B.8 at $\varphi = 0$ and Weyl's theorem on essential spectrum for general φ .

Proposition B.10 (Essential-spectrum bottom is $\sigma^2/8$, universal in φ) Let $m(d\omega) = \frac{2\omega^2}{\sigma_\omega(\omega)^2} d\omega$ be the speed measure on $I = (\omega_c, 1/\epsilon)$ with respect to the scale $s(\omega) = 1/\omega$, and let $L^2(m)$ denote the associated weighted space. The operator $-\mathcal{L}$ with domain

$$\mathcal{D} = \{f \in C^2(I) \cap L^2(m) : f(\omega_c) = 0, \text{ and } f \text{ lies in } L^2(m) \text{ in a neighbourhood of } 1/\epsilon\}$$

is self-adjoint and non-negative in $L^2(m)$ (the right endpoint $1/\epsilon$ is natural-unattainable by the Feller classification of Appendix B.9, so requires no separate boundary condition beyond $L^2(m)$ -limit-point selection; see *Karlin and Taylor, 1981, Ch. 15*). Then, for every $\varphi \geq 0$,

(i) the essential spectrum satisfies $\inf \sigma_{\text{ess}}(-\mathcal{L}) = \sigma^2/8$;

(ii) the discrete spectrum $\sigma_{\text{disc}}(-\mathcal{L}) \cap (0, \sigma^2/8)$ is empty.

Consequently, conditional on $\{\tau_c < \infty\}$,

$$\mathbb{P}(\tau_c > T \mid \omega_0, \tau_c < \infty) = O(T^{-\alpha} e^{-\sigma^2 T/8}) \quad \text{as } T \rightarrow \infty, \quad (\text{B.55})$$

for some algebraic prefactor exponent $\alpha > 0$ fixed by the density of states at the spectrum edge (generically $\alpha = 3/2$ for a Schrödinger-type boundary (*Demuth and van Casteren, 2000, Ch. 3*)); the exponential rate $\sigma^2/8$ is universal (independent of $\varphi, \omega_c, \epsilon$).

Proof. The proof has four steps: a φ -universal boundary expansion of σ_ω , the indicial analysis at $1/\epsilon$, the $L^2(m)$ filter, and non-vanishing of the bounded fundamental solution at $\varphi = 0$. Universality in φ then follows from Weyl's theorem on invariance of essential spectrum under relatively compact perturbations.

Step 1 (φ -universal boundary behaviour). Write $\omega = 1/\epsilon - \xi$ with $\xi \downarrow 0$. The numerator of (B.23) is $\sigma(1+\varphi)(1-\omega)(1/\epsilon - \omega) = \sigma(1+\varphi)[(1-1/\epsilon) + \xi]\xi$; the denominator is $D(\omega) = \varphi(1-\omega) + (1-1/\epsilon) = (1+\varphi)(1-1/\epsilon) + \varphi\xi$ after rearranging $\varphi(1-\omega) = \varphi(1-1/\epsilon) + \varphi\xi$ and adding $(1-1/\epsilon)$. The ratio is

$$\sigma_\omega(1/\epsilon - \xi) = \sigma\xi \cdot \frac{(1+\varphi)[(1-1/\epsilon) + \xi]}{(1+\varphi)(1-1/\epsilon) + \varphi\xi} = \sigma\xi[1 + O(\xi)], \quad (\text{B.56})$$

since the leading order of the ratio is $(1+\varphi)(1-1/\epsilon)/[(1+\varphi)(1-1/\epsilon)] = 1$. The $(1+\varphi)(1-1/\epsilon)$ factors cancel. In particular,

$$\sigma_\omega(\omega)^2 = \sigma^2\xi^2[1 + O(\xi)] \quad (\xi = 1/\epsilon - \omega \downarrow 0), \quad (\text{B.57})$$

universally in $\varphi \geq 0$.

Step 2 (indicial equation at $1/\epsilon$). The eigenvalue equation $\mathcal{L}\psi + \lambda\psi = 0$ on I reads

$$\frac{\sigma_\omega(\omega)^2}{2}\psi''(\omega) + \frac{\sigma_\omega(\omega)^2}{\omega}\psi'(\omega) + \lambda\psi(\omega) = 0.$$

Substitute $\xi = 1/\epsilon - \omega$, so $d/d\omega = -d/d\xi$ and $\omega = 1/\epsilon + O(\xi)$. Using (B.57),

$$\frac{\sigma^2\xi^2}{2}\psi_{\xi\xi} - \sigma^2\epsilon\xi^2\psi_\xi + \lambda\psi = O(\xi^3\psi_{\xi\xi}) + O(\xi^3\psi_\xi). \quad (\text{B.58})$$

Try $\psi(\xi) = \xi^\alpha$. Then $\psi_\xi = \alpha\xi^{\alpha-1}$, $\psi_{\xi\xi} = \alpha(\alpha-1)\xi^{\alpha-2}$. The first term on the left of (B.58) is $\frac{1}{2}\sigma^2\alpha(\alpha-1)\xi^\alpha$; the second is $-\sigma^2\epsilon\alpha\xi^{\alpha+1} = O(\xi^{\alpha+1})$, subleading; the third is $\lambda\xi^\alpha$. Balancing ξ^α ,

$$\alpha(\alpha-1) = -\frac{2\lambda}{\sigma^2}, \quad \alpha_\pm(\lambda) = \frac{1}{2}\left(1 \pm \sqrt{1 - 8\lambda/\sigma^2}\right). \quad (\text{B.59})$$

Write $\nu(\lambda) \equiv \sqrt{1 - 8\lambda/\sigma^2}$. For $\lambda \in (0, \sigma^2/8)$, $\nu \in (0, 1)$ is real, and $\alpha_- < 1/2 < \alpha_+$. At $\lambda = \sigma^2/8$, $\nu = 0$, $\alpha_\pm = 1/2$ (coincident roots); the second fundamental solution acquires a $\log \xi$ factor. For $\lambda > \sigma^2/8$, ν is purely imaginary; $\alpha_\pm = 1/2 \pm i|\nu|/2$ are complex-conjugate, and $\xi^{\alpha_\pm}$ oscillates in $\log \xi$.

Step 3 ($L^2(m)$ filter at the natural endpoint). The speed measure near $1/\epsilon$ satisfies, using (B.57) and $\omega \rightarrow 1/\epsilon$,

$$m(d\omega) = \frac{2\omega^2}{\sigma_\omega^2} d\omega \sim \frac{2/\epsilon^2}{\sigma^2 \xi^2} d\xi \quad (\xi \downarrow 0). \quad (\text{B.60})$$

For $\psi \sim \xi^\alpha$,

$$\int_0^\xi |\psi(\xi)|^2 m(d\omega) \sim \frac{2}{\epsilon^2 \sigma^2} \int_0^\xi \xi^{2\alpha-2} d\xi,$$

which converges at $\xi = 0$ iff $2\alpha - 2 > -1$, i.e., iff $\alpha > 1/2$.

Combining Steps 2–3:

- For $\lambda \in (0, \sigma^2/8)$: $\alpha_+ > 1/2 > \alpha_-$, so exactly one fundamental solution (the α_+ -branch) is $L^2(m)$ -admissible at $1/\epsilon$. The natural endpoint is in the *limit-point* case (Karlin and Taylor, 1981, §15.7).
- For $\lambda = \sigma^2/8$: $\alpha = 1/2$; $|\psi|^2 \sim \xi^{-1}$, $\int \xi^{-1} d\xi$ diverges logarithmically. Neither fundamental solution is $L^2(m)$ -admissible.
- For $\lambda > \sigma^2/8$: $|\psi|^2 \sim \xi^{-1}$ with oscillation, again non- $L^2(m)$ for either branch.

Step 4 (no discrete eigenvalue below $\sigma^2/8$ at $\varphi = 0$). By Proposition B.8 and (B.48), the $L^2(m)$ -admissible fundamental solution of $\mathcal{L}\psi + \lambda\psi = 0$ at $\varphi = 0$, written in the $Q = 1/\omega$ coordinate, is

$$\psi_{\text{adm}}(Q; \lambda) \propto V_+(Q) = (Q-1)^{\alpha_-(\lambda)}(Q-\epsilon)^{\alpha_+(\lambda)}, \quad 0 < \lambda < \sigma^2/8. \quad (\text{B.61})$$

(Steps 2–3 above identified α_+ as the boundary-admissible exponent at $Q = \epsilon$; this matches p_+ in (B.48) once one checks the sign of the argument $1 - 8\lambda/\sigma^2$ versus $1 + 8\lambda/\sigma^2$: the eigenvalue equation has opposite sign for λ compared to the Laplace-transform equation (B.41), so the index roots are $\alpha_\pm(\lambda) = p_\mp(-\lambda)$ in the notation of (B.45), restoring the sign in (B.59) under $\lambda \mapsto -\lambda$.)

A discrete eigenvalue $\lambda \in (0, \sigma^2/8)$ would require $\psi_{\text{adm}}(Q_c; \lambda) = 0$. But for $\lambda \in (0, \sigma^2/8)$, $\alpha_\pm$ are real, and $Q_c - 1 \geq \epsilon - 1 > 0$ (since $\omega_c < 1/\epsilon$ implies $Q_c > \epsilon > 1$) and $Q_c - \epsilon > 0$. Hence

$$\psi_{\text{adm}}(Q_c; \lambda) = (Q_c - 1)^{\alpha_-(\lambda)}(Q_c - \epsilon)^{\alpha_+(\lambda)} > 0,$$

a strictly positive real number. The Dirichlet condition cannot be met, so no discrete eigenvalue of $-\mathcal{L}$ lies in $(0, \sigma^2/8)$ at $\varphi = 0$. Combined with Step 3, which shows that $\sigma^2/8$ itself is

the threshold above which no $L^2(m)$ -admissible fundamental solution exists at $1/\epsilon$, this gives $\inf \sigma(-\mathcal{L}) = \sigma^2/8$ and identifies $\sigma^2/8$ as the bottom of the essential spectrum at $\varphi = 0$.

Step 5 (universality in φ). Step 1 showed that the leading $\sigma_\omega \sim \sigma\xi$ behaviour is φ -universal. Hence the indicial exponents (B.59) and the $L^2(m)$ -threshold $\alpha > 1/2$ are identical for every $\varphi \geq 0$; the critical eigenvalue at which the two admissible fundamental solutions degenerate is $\sigma^2/8$ regardless of φ . This establishes claim (i) at general φ : $\inf \sigma_{\text{ess}}(-\mathcal{L}_\varphi) = \sigma^2/8$.

To extend the “no eigenvalue below $\sigma^2/8$ ” statement from $\varphi = 0$ to general φ , invoke *Weyl’s theorem* (Reed and Simon, 1978, Theorem XIII.14 and Corollary 2): the essential spectra of two self-adjoint operators agree if their difference is relatively compact. The operator family $\{-\mathcal{L}_\varphi\}_{\varphi \geq 0}$ differs only through the coefficient $\sigma_\omega(\omega; \varphi)$, which converges *uniformly* on any compact subset of I as $\varphi \rightarrow 0$, and whose boundary behaviour at both endpoints of I is φ -universal by Step 1 (and by direct inspection at ω_c : $\sigma_\omega(\omega_c; \varphi) = \sigma(1+\varphi)(1-\omega_c)(1/\epsilon - \omega_c)/D(\omega_c)$, bounded and smooth). The difference $-\mathcal{L}_\varphi - (-\mathcal{L}_0)$ is a second-order differential operator with coefficients bounded on any closed subinterval of I and with boundary asymptotics matching to leading order at both endpoints; its resolvent-difference is compact by the compact-embedding argument of Reed and Simon (1978, Theorem XIII.68). Weyl’s theorem then gives $\sigma_{\text{ess}}(-\mathcal{L}_\varphi) = \sigma_{\text{ess}}(-\mathcal{L}_0) = [\sigma^2/8, \infty)$.

For claim (ii) at general φ , we do not have a closed-form fundamental solution. However, the φ -universal boundary analysis of Steps 1–3 shows that if a discrete eigenvalue $\lambda \in (0, \sigma^2/8)$ existed, its eigenfunction would satisfy $\psi(\omega_c) = 0$ and decay as $\xi^{\alpha+(\lambda)}$ at $1/\epsilon$. A second-order ODE with two boundary conditions admits at most a discrete set of such λ , and by self-adjointness and Weyl’s theorem this discrete set is contained in $[0, \sigma^2/8) \cup (\text{accumulation points of } \sigma_{\text{disc}})$. For the specific one-dimensional diffusion at hand, the generator is of regular-boundary/natural-boundary type with a single limit-point endpoint; standard Sturm–Liouville theory for such operators (Karlin and Taylor, 1981, §15.7; Borodin and Salminen, 2002, Ch. VII.3) implies that the resolvent is Hilbert–Schmidt on the discrete-spectrum component, and that the discrete spectrum below $\inf \sigma_{\text{ess}}$ is either empty or a finite set accumulating only at $\inf \sigma_{\text{ess}} = \sigma^2/8$. At $\varphi = 0$ Step 4 showed the former. By relative compactness of the $\varphi \mapsto \varphi$ perturbation and continuity of discrete eigenvalues under compact perturbation (Reed and Simon, 1978, Theorem XIII.16), the discrete spectrum below $\sigma^2/8$ remains empty on an open neighbourhood of $\varphi = 0$ in the coefficient topology above; and the φ -universal boundary analysis precludes appearance of new discrete eigenvalues via bifurcation from the essential-spectrum edge because

the boundary exponents $\alpha_{\pm}(\lambda)$ do not depend on φ . Hence (ii) holds for all $\varphi \geq 0$.

Step 6 (decay of the conditional survival probability). Because $\sigma(-\mathcal{L}) \subset [\sigma^2/8, \infty)$ with empty discrete part below $\sigma^2/8$, the spectral theorem gives the semigroup bound $\|e^{\mathcal{L}T}\| \leq e^{-\sigma^2 T/8}$ in $L^2(m)$. The conditional survival probability is the ratio $\mathbb{P}(\tau_c > T \mid \omega_0)/p(\omega_0)$, and standard semigroup-kernel estimates for one-dimensional diffusions with natural-endpoint essential spectrum (Borodin and Salminen, 2002, Ch. II §2) yield the polynomial-prefactor correction

$$\mathbb{P}(\tau_c > T \mid \omega_0) = \int_I p_T(\omega_0, \omega') m(d\omega') = C(\omega_0) T^{-\alpha} e^{-\sigma^2 T/8} [1 + o(1)],$$

for some $\alpha > 0$ (the density-of-states exponent at the spectrum edge; for one-dimensional Schrödinger-type boundaries the generic value is $\alpha = 3/2$, see Demuth and van Casteren, 2000, Ch. 3). Dividing by $p(\omega_0)$ gives (B.55). $\square$

Remark B.8 (What is delicate, what survives) *The exact value of the algebraic prefactor α depends on the density of states at $\lambda = \sigma^2/8$, which in turn depends on the sub-leading coefficients in (B.56). Determining α explicitly requires a full Liouville normal-form reduction of the eigenvalue equation near $1/\epsilon$ (not carried out here). The exponential rate $\sigma^2/8$ is however analytically pinned down and φ -universal; that is the content of the Proposition. The quantitatively relevant horizons of Table B.2 below ($T \leq 50$) have $\sigma^2 T/8 \leq 0.142$, so $e^{-\sigma^2 T/8} \in [0.868, 1]$: the long- T exponential tail is negligibly slow on policy-relevant horizons, and the tail probabilities at $T \leq 50$ are determined by the full Laplace inversion, not by the asymptotic rate.*

B.8.7.5 Calibrated numerical table

Remark B.9 (Tail probabilities at the empirical calibration) *At the paper's empirical calibration $\epsilon = 13.7741$, $\sigma = 0.1505$, $\varphi = 1.3333$, $\omega_c = 0.0606$ (so $1/\epsilon \approx 0.0726$ and the pre-failure interval has width 0.0120), Table B.2 reports $\mathbb{P}(\tau_c \leq T \mid \omega_0)$ computed by numerical inversion of the Laplace transform of Proposition B.7 via Gaver–Stehfest (Stehfest, 1970) ($N = 10$ terms), with the Kolmogorov backward equation solved at each Stehfest node by finite-difference BVP on a mesh of 1500 points with Dirichlet cutoff at $1/\epsilon - 10^{-5}$.*

The empirical pre-2008 U.S. path for the HKM capital ratio puts $\bar{\omega}_{1995-2006} = 0.095$, well above $1/\epsilon \approx 0.0726$, so the “pre-failure domain” is only a valid description from roughly 2008Q1 onward; the post-crisis trough places ω_t below ω_c , in the QE-constrained regime (not this appendix's domain). The table's first row— ω_0 one-quarter of the way from ω_c to $1/\epsilon$, hence $\omega_0 = 0.0636$ —is the closest

ω_0 position	ω_0	$p(\omega_0)$	$T = 1$	$T = 5$	$T = 10$	$T = 25$	$T = 50$	$T = 100$
$\omega_c + 0.25(1/\epsilon - \omega_c)$	0.0636	0.715	0.047	0.323	0.447	0.570	0.633	0.674
midpoint	0.0666	0.455	0.000	0.026	0.093	0.223	0.314	0.382
$\omega_c + 0.75(1/\epsilon - \omega_c)$	0.0696	0.218	0.000	0.000	0.002	0.027	0.077	0.136
$1/\epsilon - 0.002$	0.0706	0.143	0.000	0.000	0.000	0.006	0.030	0.070

Table B.2: Tail probability $\mathbb{P}(\tau_c \leq T \mid \omega_0)$ at the empirical calibration $\epsilon = 13.7741$, $\sigma = 0.1505$, $\varphi = 1.3333$, $\omega_c = 0.0606$. Column $p(\omega_0)$ is the asymptotic limit $\lim_{T \rightarrow \infty} \mathbb{P}(\tau_c \leq T \mid \omega_0)$ from Proposition B.4; individual entries approach this limit from below. The last column shows that even at $T = 100$ years the approach is incomplete for $\omega_0 \geq \text{midpoint}$, reflecting the small essential-spectrum bottom $\lambda_* = \sigma^2/8 \approx 0.00283$ and the ≈ 353 -year characteristic time (Proposition B.10; see below). Horizons: years.

empirically-relevant starting point on the post-crisis HKM series. At this ω_0 , the 5-year crisis probability is 32%, the 10-year is 45%, and the asymptotic (infinite-horizon) is 71%; the decade horizon captures more than half of the ultimate crisis probability.

Freidlin–Wentzell benchmark. The short- T bound (B.54), ignoring drift, gives for the same row (Riemann distance $d = 0.286$) $\exp(-d^2/(2T\sigma^2)) = \exp(-1.81/T)$, yielding 0.70 at $T = 5$ and 0.83 at $T = 10$ —larger than the exact values, consistent with the positive drift reducing early-crisis probability relative to the drift-free bound.

Long-horizon decay rate: analytical, not FD. Proposition B.10 replaces the earlier (incorrect) attempt to extract a “principal eigenvalue” by finite-difference discretisation of $-\mathcal{L}$: the spectrum below $\sigma^2/8$ is not discrete, so no cutoff-dependent $\lambda_1 \in [0.004, 0.008]$ exists as an isolated eigenvalue. The correct long-horizon exponential decay rate is the essential-spectrum bottom $\lambda_* = \sigma^2/8$, φ -universal. At the calibration $\sigma = 0.1505$,

$$\lambda_* = \frac{\sigma^2}{8} = \frac{0.02265}{8} \approx 2.83 \times 10^{-3}, \quad \text{long-time characteristic} \quad \frac{1}{\lambda_*} = \frac{8}{\sigma^2} \approx 353 \text{ years.} \quad (\text{B.62})$$

The associated e -fold decay is at $T \approx 353$ years; equivalently, a half-life of $\ln 2 \cdot 8/\sigma^2 \approx 245$ years. Crucially, the Table entries at $T \in \{1, 5, 10, 25, 50, 100\}$ are insensitive to this asymptotic rate: $\sigma^2 T/8$ equals 0.00283, 0.0142, 0.0283, 0.0708, 0.142, 0.283 at those horizons, so $e^{-\sigma^2 T/8} \in [0.753, 0.997]$. For $T \leq 50$, the exponential correction is well under 15%, and the tail-probability magnitudes are set by the short-to-intermediate-horizon dynamics of the Laplace inversion, not by the spectrum-edge rate. The Gaver–Stehfest inversion solves (B.41) at each Stehfest node directly and does not invoke any spectral-gap estimate; entries $T \in \{50, 100\}$ in Table B.2 are from this backward-equation inversion and are not based on the incorrect FD principal eigenvalue.

Honest caveats. (i) The empirical anchor $\omega_c = 0.0606$ is taken from the calibration; the conference-version calibration with $\omega_c = 0.078$, $1/\epsilon = 0.10$ at $\epsilon = 10$ yields a wider pre-failure interval and less

binding tail probabilities. The long-time rate $\sigma^2/8$ depends only on σ , so changing only ω_c and ϵ does not alter it. (ii) The algebraic prefactor α in (B.55) is not analytically determined here (Remark B.8); the exponential rate $\sigma^2/8$ is. (iii) The $\varphi = 0$ closed form (B.44) does not apply at the empirical $\varphi = 1.3333$; the numerical inversion via (B.41) is the correct route at $\varphi > 0$, and the analytical decay rate $\sigma^2/8$ is φ -universal by Proposition B.10.

Cross-references. Table B.2 is the tail-probability analogue of Proposition B.4 (column $p(\omega_0)$, $T \rightarrow \infty$ limit) and Proposition B.5 (first-moment $\int_0^\infty T dF_T$, infinite by Theorem B.3). It complements Appendix BO.14.1, which computes the *deterministic* drift time on the QE-recovery region $(\underline{\omega}, \omega_c)$; the present appendix is about stochastic hitting on the pre-failure region $(\omega_c, 1/\epsilon)$.

B.8.8 Economic Interpretation

Proposition B.4 says: the probability of an eventual crisis is exactly the fraction of the distance—measured in the scale coordinate $Q = 1/\omega$ —that the economy has already covered from the frictionless steady state $1/\epsilon$ toward the crisis threshold $1/\omega_c$. The scale coordinate is the natural metric because Q_t is a martingale. In the raw wealth-share coordinate, the probability is $\omega_c(1 - \epsilon\omega_0)/[\omega_0(1 - \epsilon\omega_c)]$: it falls as ω_0 rises and falls as ω_c falls. Theorem B.3 sharpens the probabilistic picture: although a crisis occurs with strictly positive probability for every $\omega_0 \in I$, it occurs in *expected* finite time only on the conditional event $\{\tau_c < \infty\}$; averaged unconditionally, the expected time is infinite because the drift pulls the economy toward $1/\epsilon$ where it spends ergodically long excursions. Proposition B.5 quantifies the conditional expectation via a Green’s-function integral in the speed-measure-weighted scale coordinate; Proposition B.6 specialises this to an elementary formula at $\varphi = 0$. Taken together, these results show that the main paper’s Theorem 2.1 delivers positive-probability failure with a precisely characterised timing distribution, consistent with the narrative in §2.3.1 that decades of pre-2008 observations with $\omega_t \gg \omega_c$ are fully compatible with a finite—but conditionally-distributed— τ_c .

The full tail distribution characterised in Propositions B.7–B.10 adds the horizon-specific information policy makers need. Hitting probability and conditional expected time are the $T \rightarrow \infty$ limit and first moment of the same object: the tail distribution $T \mapsto \mathbb{P}(\tau_c \leq T \mid \omega_0)$ governed by the backward equation (B.41). The tail sits between two asymptotic regimes. At short T , Proposition B.9 shows the log-probability is quadratic in the Riemannian distance to ω_c and linear in $1/T$ with rate $1/\sigma^2$ —small-noise logic: many standard deviations of adverse shocks

in a short window are exponentially rare. At long T , the conditional tail decays as $T^{-\alpha} e^{-\sigma^2 T/8}$ with the exponential rate $\sigma^2/8$ equal to the essential-spectrum bottom of $-\mathcal{L}$ (Proposition B.10), a φ -universal analytical constant. The long-time characteristic is $1/(\sigma^2/8) = 8/\sigma^2 \approx 353$ years at $\sigma = 0.1505$, so the policy-horizon tail probabilities in Table B.2 ($T \leq 50$) are determined not by the asymptotic rate but by the short-to-intermediate dynamics of the Laplace inversion: the 5-year crisis probability from $\omega_0 = 0.064$ is 32%, the 10-year is 45%, and the 25-year 57%, with the ultimate 71% limit approached on decadal-to-centennial timescales. The σ -scaling from Proposition B.9 then says that even modest increases in fundamental volatility move the policy-relevant medium-horizon tail substantially—more than the hitting probability or the conditional mean separately would suggest; the same σ -dependence appears, through the $\sigma^2/8$ rate, in the asymptotic tail at the opposite horizon limit.

B.9 Boundary Classification

Proof of Theorem 2.3. At $\omega = \underline{\omega}$: $\sigma_\omega(\underline{\omega}) = 0$ (Proposition BO.1), so the diffusion coefficient vanishes. The drift relation $\mu_\omega = \sigma_\omega^2/\omega$ then gives $\mu_\omega(\underline{\omega}) = 0$. Both drift and diffusion vanish: $\omega_t = \underline{\omega}$ is a constant solution to the SDE (2.22). By pathwise uniqueness (Karatzas and Shreve, 1991, Theorem 5.2.5), this is the unique solution starting at $\underline{\omega}$. An identical argument applies at $\omega = 1/\epsilon$.

For uniqueness: at any other point $\omega^* \in (0, 1) \setminus \{\underline{\omega}, 1/\epsilon\}$, Proposition BO.1 gives $\sigma_\omega(\omega^*) \neq 0$, so the diffusion does not vanish and ω^* is not absorbing. $\square$

Proof of Theorem BO.1 and Proposition BO.2.

We use the Feller boundary classification for one-dimensional diffusions (Karlin and Taylor, 1981, Ch. 15), (Borodin and Salminen, 2002, Ch. II). For a diffusion $dX = \mu(X)dt + \sigma(X)dZ$ on an interval (l, r) , the boundary classification depends on the scale function density $s'(x) = \exp(-\int^x 2\mu(y)/\sigma(y)^2 dy)$ and the speed measure density $m(x) = 1/(\sigma(x)^2 s'(x))$. A boundary is *unattainable* (entrance or natural) if $\Sigma = \int s'(x)dx$ diverges at the boundary, or if $N = \int s'(x) \int m(y)dy dx$ diverges. A boundary where both Σ and N diverge is called *natural*.

For our diffusion (2.22), $\mu_\omega = \sigma_\omega^2/\omega$, so $2\mu_\omega/\sigma_\omega^2 = 2/\omega$. Therefore $s'(\omega) = \exp(-2 \ln \omega) = C/\omega^2$ (normalised at ω_0 : $s'(\omega) = \omega_0^2/\omega^2$) and $m(\omega) = \omega^2/(\omega_0^2 \sigma_\omega(\omega)^2)$.

(i) $\omega = 0$ (**natural boundary**). Near $\omega = 0$, the constrained formula gives $\sigma_\omega^c = \sigma(1 + \varphi)\lambda(1 - \omega)\omega(\omega - \underline{\omega})/D_c(\omega)$. Substituting $\omega \rightarrow 0$: $(1 - \omega) \rightarrow 1$, $(\omega - \underline{\omega}) \rightarrow -\underline{\omega}$, $D_c(0) = (1 + \varphi)$. Therefore

$\sigma_\omega \approx \sigma(1+\varphi)\lambda \cdot 1 \cdot \omega \cdot (-\underline{\omega})/(1+\varphi) = -\sigma\lambda\underline{\omega} \cdot \omega \equiv -c_0\omega$ with $c_0 = \sigma\lambda\underline{\omega} > 0$. There exist $\underline{c}, \bar{c} > 0$ with $\underline{c}\omega \leq |\sigma_\omega| \leq \bar{c}\omega$ on $(0, a]$ for some $a > 0$.

$$\Sigma(0) = \int_0^a s'(x) dx = \omega_0^2 \int_0^a x^{-2} dx = +\infty.$$

$$m(\omega) \geq \omega^2/(\omega_0^2 \bar{c}^2 \omega^2) = 1/(\omega_0^2 \bar{c}^2) \equiv K > 0. \text{ Hence } \int_0^x m(y) dy \geq Kx \text{ and}$$

$$N(0) = \int_0^a s'(x) \int_0^x m(y) dy dx \geq K \int_0^a \frac{\omega_0^2}{x^2} \cdot x dx = K\omega_0^2 \int_0^a x^{-1} dx = +\infty.$$

Both Σ and N diverge: $\omega = 0$ is a natural boundary, unattainable from the interior.

(ii) $\omega = \underline{\omega}$ (**unattainable from both sides**). The constrained formula $\sigma_\omega^c = \sigma(1+\varphi)\lambda(1-\omega)\omega(\omega-\underline{\omega})/D_c(\omega)$ factors as $\sigma_\omega = g(\omega)(\omega-\underline{\omega})$ with $g(\omega) = \sigma(1+\varphi)\lambda(1-\omega)\omega/D_c(\omega)$. Evaluating at $\underline{\omega}$: $D_c(\underline{\omega}) = (1+\varphi)(1-\underline{\omega})$, so $g(\underline{\omega}) = \sigma(1+\varphi)\lambda(1-\underline{\omega})\underline{\omega}/[(1+\varphi)(1-\underline{\omega})] = \sigma\lambda\underline{\omega} > 0$. There exist $\underline{g}, \bar{g} > 0$ with $\underline{g} \leq |g(\omega)| \leq \bar{g}$ for ω near $\underline{\omega}$.

Unattainability from the right. We show that the speed measure integral diverges, which forces $N(\underline{\omega}^+) = +\infty$ in the Feller classification. Near $\underline{\omega}$ from the right: $m(\omega) = \omega^2/(\omega_0^2 g(\omega)^2 (\omega-\underline{\omega})^2) \geq \underline{\omega}^2/(\omega_0^2 \bar{g}^2 (\omega-\underline{\omega})^2)$. Therefore:

$$\int_{\underline{\omega}}^x m(y) dy \geq \frac{\underline{\omega}^2}{\omega_0^2 \bar{g}^2} \int_{\underline{\omega}}^x \frac{1}{(y-\underline{\omega})^2} dy = \frac{\underline{\omega}^2}{\omega_0^2 \bar{g}^2} \cdot \frac{1}{x-\underline{\omega}} \rightarrow +\infty \text{ as } x \rightarrow \underline{\omega}^+.$$

Since $s'(\omega) = \omega_0^2/\omega^2 \geq \omega_0^2/b^2 > 0$ for $\omega \in [\underline{\omega}, b]$, it follows that the Feller integral is $N(\underline{\omega}^+) = \int_{\underline{\omega}}^b s'(x) \int_{\underline{\omega}}^x m(y) dy dx = +\infty$. Hence $\underline{\omega}$ is unattainable from the right. An identical argument (replacing $(\omega-\underline{\omega})$ by $(\underline{\omega}-\omega)$ in the crisis regime where $\sigma_\omega = g(\omega)(\omega-\underline{\omega})$ with $g > 0$ and $\omega < \underline{\omega}$ so $|\sigma_\omega| = |g|(\underline{\omega}-\omega)$) shows unattainability from the left.

(iii) $\omega = 1/\epsilon$ (**unattainable from both sides**). The unconstrained formula $\sigma_\omega^u = \sigma(1+\varphi)(1-\omega)(1/\epsilon-\omega)/D_u(\omega)$ factors as $\sigma_\omega = g(\omega)(1/\epsilon-\omega)$ with $g(\omega) = \sigma(1+\varphi)(1-\omega)/D_u(\omega)$. At $\omega = 1/\epsilon$: $D_u(1/\epsilon) = (1+\varphi)(1-1/\epsilon)$, so $g(1/\epsilon) = \sigma(1+\varphi)(1-1/\epsilon)/[(1+\varphi)(1-1/\epsilon)] = \sigma > 0$. The analysis is identical to (ii): $m(\omega) \geq K(1/\epsilon-\omega)^{-2}$ near $1/\epsilon$ from the left, so $\int m(y) dy$ diverges and $N(1/\epsilon^-) = +\infty$. Symmetrically from the right.

(iv) $\omega = 1$ (**unattainable**). The unconstrained formula factors as $\sigma_\omega = h(\omega)(1-\omega)$ with $h(\omega) = \sigma(1+\varphi)(1/\epsilon-\omega)/D_u(\omega)$. At $\omega = 1$: $D_u(1) = (1-1/\epsilon)$, so $h(1) = \sigma(1+\varphi)(1/\epsilon-1)/(1-1/\epsilon) = -\sigma(1+\varphi) \neq 0$. Near $\omega = 1$: $m(\omega) \geq K(1-\omega)^{-2}$, giving $N(1^-) = +\infty$. Unattainable.

Regime invariance. The process ω_t has continuous paths (it is a strong solution of an SDE with continuous coefficients on the interior of each regime). If $\omega_0 \in (0, \underline{\omega})$ and $\omega_t \notin (0, \underline{\omega})$

for some $t > 0$, then by continuity there exists $s \in (0, t)$ with $\omega_s \in \{0, \underline{\omega}\}$, contradicting the unattainability of both boundaries. The same argument applies to $(\underline{\omega}, 1/\epsilon)$ and $(1/\epsilon, 1)$.

□

B.10 Convergence

Proof of Theorem 2.4.

Almost-sure convergence. $Q_t = 1/\omega_t$ is a non-negative local martingale (Proposition 2.2), hence a supermartingale by Fatou's lemma. By the supermartingale convergence theorem (Karatzas and Shreve, 1991, Ch. 1), $Q_t \rightarrow Q_\infty$ a.s. with $\mathbb{E}[Q_\infty] \leq Q_0 < \infty$, so $Q_\infty < \infty$ a.s. and $\omega_\infty \equiv 1/Q_\infty > 0$.

The limit is a zero of σ_ω . The quadratic variation of Q is $[Q]_t = \int_0^t (\sigma_\omega(\omega_s)/\omega_s^2)^2 ds$. We claim $[Q]_\infty < \infty$ a.s. By the Dambis–Dubins–Schwarz representation (Karatzas and Shreve, 1991, Ch. 3), there exists a Brownian motion B such that $Q_t = Q_0 + B_{[Q]_t}$ for all $t \geq 0$. If $[Q]_\infty = +\infty$, then $Q_t = Q_0 + B_{[Q]_t}$ where B is evaluated at a time tending to infinity; since Brownian motion does not converge (it is neighbourhood-recurrent on $\mathbb{R}$, Karatzas and Shreve, 1991, Theorem 2.9.23), this contradicts the a.s. convergence $Q_t \rightarrow Q_\infty$. Therefore $[Q]_\infty < \infty$ a.s.

Since $\omega_t \rightarrow \omega_\infty$ and the integrand $(\sigma_\omega(\omega)/\omega^2)^2$ is continuous, the convergence $[Q]_\infty < \infty$ requires $\sigma_\omega(\omega_\infty) = 0$: if $\sigma_\omega(\omega_\infty) \neq 0$, then by continuity the integrand would be bounded below by a positive constant for all large t , and the integral would diverge. By Proposition BO.1, the zeros of σ_ω in $[0, 1)$ are $\{0, \underline{\omega}, 1/\epsilon\}$. Since $\omega_\infty > 0$, we have $\omega_\infty \in \{\underline{\omega}, 1/\epsilon\}$.

Regime-specific limits.

Crisis regime $(0, \underline{\omega})$: ω_t remains in $(0, \underline{\omega})$ for all t (regime invariance, Proposition BO.2). The only zero of σ_ω accessible from this regime is $\underline{\omega}$ (since $\omega_\infty > 0$ and $1/\epsilon \notin [0, \underline{\omega})$). Hence $\omega_t \rightarrow \underline{\omega}$.

Recovery regime $(\underline{\omega}, 1/\epsilon)$: ω_t remains in $(\underline{\omega}, 1/\epsilon)$. Since $\underline{\omega}$ is an entrance boundary (Proposition BO.2), the process cannot approach it from the interior. Moreover, in this regime the drift $\mu_\omega = \sigma_\omega^2/\omega > 0$ pushes ω away from $\underline{\omega}$, so $\omega_\infty \neq \underline{\omega}$. Hence $\omega_t \rightarrow 1/\epsilon$.

Inverted regime $(1/\epsilon, 1)$: ω_t remains in $(1/\epsilon, 1)$. Since 1 is unattainable, $\omega_\infty \neq 1$. The only zero of σ_ω in $[1/\epsilon, 1)$ is $1/\epsilon$. Hence $\omega_t \rightarrow 1/\epsilon$.

Infinite expected hitting time. Let $\omega^* \in \{\underline{\omega}, 1/\epsilon\}$ be the relevant absorbing state. Near ω^* : $\sigma_\omega(\omega) = g(\omega)(\omega^* - \omega)$ with g continuous and bounded between $\underline{g}, \bar{g} > 0$ (Proposition BO.2). The speed measure density satisfies $m(\omega) = \omega^2/(\omega_0^2 \sigma_\omega^2) \geq (\omega^*)^2/(2\omega_0^2 \bar{g}^2 (\omega^* - \omega)^2)$ for ω close to

ω^* . By the Green's function formula for expected hitting times (Karlin and Taylor, 1981, Ch. 15):

$$\mathbb{E}_{\omega_0}[\tau_{\omega^*}] = \int_{\omega_0}^{\omega^*} \int_{\omega_0}^y 2 m(z) s'(z) dz [s(y) - s(\omega_0)] dy + \int_{\omega_0}^{\omega^*} [s(\omega^*) - s(y)] 2 m(y) dy.$$

The second integral satisfies $\int_a^{\omega^*} [s(\omega^*) - s(y)] m(y) dy \geq K \int_a^{\omega^*} (\omega^* - y)^{-1} dy = +\infty$, where we used $s(\omega^*) - s(y) \sim s'(\omega^*)(\omega^* - y)$ for y near ω^* and $m(y) \geq K'(\omega^* - y)^{-2}$, giving an integrand of order $(\omega^* - y)^{-1}$. Therefore $\mathbb{E}[\tau] = +\infty$. $\square$

B.11 Proof of Intermediation Gap Classification

Proof of Theorem 2.6.

By Proposition 2.2, $Q = 1/\omega$ is a non-negative continuous local martingale. By (2.32), the intermediation gap is positive if and only if Q is a strict local martingale. We classify Q using the integral test of Mijatović and Urusov (2012): a non-negative continuous local martingale Q with $dQ = -(\sigma_\omega/\omega^2) dZ$ is a strict local martingale if and only if

$$\int_c^\infty \frac{q}{\sigma_Q(q)^2} dq < \infty \quad (\text{B.63})$$

for some (hence all) $c > 0$, where $\sigma_Q(q)$ is the diffusion coefficient of Q expressed as a function of Q .

Since $\sigma_Q = \sigma_\omega/\omega^2$ and $Q = 1/\omega$, we have $\sigma_Q(Q) = Q^2 |\sigma_\omega(1/Q)|$. The integral becomes:

$$\int_c^\infty \frac{q}{q^4 \sigma_\omega(1/q)^2} dq = \int_c^\infty \frac{1}{q^3 \sigma_\omega(1/q)^2} dq.$$

(i) Baseline model. Near $\omega = 0$ (equivalently, $Q \rightarrow \infty$), the baseline volatility (B.4) satisfies $\sigma_\omega^0(\omega) \rightarrow c_0 > 0$ where $c_0 = \sigma(1 + \varphi)(1/\epsilon)/[(1 - 1/\epsilon) + \varphi]$. Therefore $\sigma_\omega^0(1/q) \rightarrow c_0$ as $q \rightarrow \infty$, and

$$\int_c^\infty \frac{1}{q^3 c_0^2} dq = \frac{1}{2c_0^2 c^2} < \infty.$$

By the integral test, Q is a strict local martingale. A positive intermediation gap exists.

(ii) Constrained model. Near $\omega = 0$, the constrained volatility satisfies $|\sigma_\omega(\omega)| \sim c'_0 \omega$ with $c'_0 = \sigma \lambda \underline{\omega} > 0$ (Proposition BO.1). Therefore $|\sigma_\omega(1/q)| \sim c'_0/q$ as $q \rightarrow \infty$, and

$$\int_c^\infty \frac{1}{q^3 (c'_0/q)^2} dq = \frac{1}{(c'_0)^2} \int_c^\infty \frac{1}{q} dq = +\infty.$$

By the integral test, Q is a true martingale. The intermediation gap is zero.

The mechanism is transparent: the leverage constraint forces $\phi \leq \bar{\phi}(\omega) \rightarrow 1 - \lambda \underline{\omega}$ as $\omega \rightarrow 0$, so the financier is de-leveraged when her capital is depleted. Since $\sigma_\omega = \omega(\phi - 1)\sigma_R$ and $\phi - 1$ remains bounded, the extra factor of ω survives, driving $\sigma_\omega \propto \omega \rightarrow 0$. In the baseline, $\phi = 1/(\epsilon\omega) \rightarrow \infty$: the ω cancels and σ_ω stays positive. One factor of ω is the difference between convergent and divergent integrals, between a positive intermediation gap and a zero gap.

For the three individual regimes of the constrained model: in the recovery and expansionary regimes, Q ranges over a bounded interval ($Q \in (\epsilon, 1/\underline{\omega})$ or $Q \in (1, \epsilon)$), so Q is bounded and hence a true martingale. In the crisis regime $\omega \in (0, \underline{\omega})$, Q is unbounded above, but the true martingale property follows from the divergent integral established above. $\square$

B.12 Empirical Appendix

B.12.1 Constructing the Central Bank Risk-Bearing Measure

This subsection details the construction of $RB_t^{CB,j}$, our measure of central-bank risk-bearing in asset class j . The construction is closely related to the SOMA duration-weighting procedure in Appendix A of [Gulati and Smith \(2022\)](#), with three key differences. First, durations are computed using the contemporaneous Wu yield curve of [Liu and Wu \(2021\)](#) rather than a fixed July 30, 2014 curve. Second, inflation compensation on TIPS is added back to holdings. Third, we do not divide the resulting duration-weighted holdings by the duration of a 10-year Treasury; instead, we retain the raw duration-weighted dollar amount and then scale it by nominal GDP. The resulting series is therefore a GDP-scaled duration-weighted stock rather than a measure in 10-year equivalents.

We begin with weekly Wednesday CUSIP-level SOMA holdings from the Federal Reserve Bank of New York. For each security i on date w , the code first constructs the principal amount held by the Federal Reserve by summing the reported face-value and par-value fields after replacing missing values with zero. It then adds reported inflation compensation for TIPS holdings. Denote the resulting holding amount by H_{iw} . Securities are assigned to three asset classes: Treasury bills, notes and bonds, TIPS, and FRNs are grouped under Treasuries; agency debts are grouped under agency debt; and agency MBS together with CMBS are grouped under MBS.

For each security, we compute duration D_{iw} as follows.

- Treasury bills: duration equals remaining maturity in years.
- Floating-rate notes: duration is set to $\min\{T_{iw}, 6/365\}$, where T_{iw} is remaining maturity in years, reflecting the weekly reset frequency of FRNs.
- Agency MBS and CMBS: duration is fixed at 5.1 years.

For notes and bonds, TIPS, and agency debts, the code assumes semiannual coupon payments and constructs the cash-flow schedule by stepping backward from the maturity date in six-month increments, preserving month-end dates when the maturity date falls at month-end. If $t_{ik,w}$ denotes the time in years from date w to cash flow k , and if $CF_{ik,w}$ denotes the corresponding coupon or principal cash flow normalized to par, then duration is computed as:

$$D_{iw} = \frac{\sum_k t_{ik,w} CF_{ik,w} \exp[-y_w(t_{ik,w})t_{ik,w}]}{\sum_k CF_{ik,w} \exp[-y_w(t_{ik,w})t_{ik,w}]}, \quad (\text{B.64})$$

where $y_w(t)$ is obtained from the daily Wu zero-coupon Treasury curve matched to date w (using the last available observation when necessary) and linearly interpolated across maturities. If the cash-flow schedule is empty or the implied bond price is zero, the code falls back to remaining maturity.

Weekly central-bank risk-bearing in asset class j is then:

$$RB_w^{CB,j,\text{raw}} = \sum_{i \in j} H_{iw} D_{iw}. \quad (\text{B.65})$$

We compute these weekly series for Treasuries, agency debt, MBS, and the total. To align them with the quarterly macro data, we take the last weekly observation in each quarter and scale it by nominal GDP in dollars:

$$RB_t^{CB,j} = \frac{RB_{w(t)}^{CB,j,\text{raw}}}{GDP_t}, \quad (\text{B.66})$$

where $w(t)$ denotes the last Wednesday observation in quarter t and GDP_t is FRED nominal GDP. The series used in the main text are therefore the quarter-end GDP-normalized duration-weighted SOMA positions in Treasuries and agency MBS.

B.12.2 Structural Breaks Additional Evidence

This subsection reports the supporting evidence for Test 2.7.7. We estimate multiple breaks in the mean of the quarterly HKM intermediary capital ratio using the Bai and Perron (1998)

mean-shift specification:

$$\omega_t = \mu_j + u_t, \quad t = T_{j-1} + 1, \dots, T_j, \quad j = 1, \dots, m + 1, \quad (\text{B.67})$$

where both the break dates $\{T_j\}_{j=1}^m$ and the number of breaks m are selected from the data. To ensure that the selected number of breaks is not mechanically capped by the trimming rule, the baseline exercise uses a trimming parameter of 0.10 and allows up to nine breaks in the full 1970Q1–2024Q4 sample. The comparison with the calibrated threshold is used only as a descriptive annotation; this exercise is not an endogenous threshold estimate.

Table B.3: Bai-Perron break-count selection for the mean of ω_t

<i>Panel A: BIC model selection</i>			
Breaks	BIC	SSR	Break dates
0	-1669.08	0.1089	–
1	-1728.04	0.0813	1992-12
2	-1853.64	0.0448	1995-09,2007-03
3	-1882.42	0.0384	1996-03,2001-09,2007-09
4	-1908.39	0.0333	1996-03,2001-09,2007-12,2013-09
5	-1947.65	0.0271	1975-09,1991-06,1996-12,2002-06,2007-12
6	-1990.40	0.0218	1975-09,1991-06,1996-12,2002-06,2007-12,2013-09
7	-2000.27	0.0204	1975-09,1983-03,1991-06,1996-12,2002-06,2007-12,2013-09
8	-1997.64	0.0201	1975-09,1983-03,1991-06,1996-12,2002-06,2007-12,2013-09,2019-09
9	-1929.33	0.0267	1975-09,1981-03,1986-09,1992-03,1997-09,2003-03,2008-09,2014-03,2019-09

<i>Panel B: Sequential break tests</i>			
Test	Statistic	Bootstrap p -value	Reject at 5%
FT(6 5)	52.11	0.001	Yes
FT(7 6)	15.23	0.001	Yes
FT(8 7)	2.67	0.119	No

Notes: Panel A reports BIC and global least-squares break dates for each admissible break count under trimming parameter 0.10. The requested upper bound is 9 breaks and the effective admissible upper bound is 9. Panel B reports sequential $FT(l + 1 | l)$ statistics with bootstrap p -values computed under the segmented-mean null.

Table B.3 reports the BIC values for break counts from zero to nine together with the sequential $FT(l + 1 | l)$ tests. Under this specification, BIC selects seven breaks. The sequential tests point to the same stopping rule: they reject six breaks in favor of seven, but do not reject seven breaks in favor of eight. The selected partition includes breaks at 2007Q4 and 2013Q3, which are the two dates most relevant for the QE-trap interpretation.

The main text therefore reports the compact regime summary in Table 2.17, which compresses the selected partition into the immediate pre-crisis benchmark and the two post-2007 regimes. The key economic pattern is a large drop in the mean of intermediary capital after 2007 and only a partial recovery thereafter: the 2007Q4–2013Q2 regime falls below the cal-

ibrated threshold, while the 2013Q3–2024Q4 regime remains below the pre-2007 benchmark even though it is only slightly above the threshold.

Table B.4: Modern-sample robustness for the mean of ω_t

Regime	Start	End	$\bar{\omega}$	Δ vs. pre-2007	Below pre-2007
Pre-2007 regime 1	1995-03	1997-12	0.0889	0.0077	–
Pre-2007 regime 2	1998-03	2001-06	0.1154	0.0343	–
Pre-2007 benchmark	2001-09	2007-09	0.0812	0.0000	–
Post-2007 regime 1	2007-12	2013-06	0.0445	-0.0367	Yes
Post-2007 regime 2	2013-09	2016-09	0.0585	-0.0227	Yes
Post-2007 regime 3	2016-12	2019-12	0.0730	-0.0082	Yes
Post-2007 regime 4	2020-03	2024-12	0.0598	-0.0214	Yes

Notes: This table re-estimates the Bai–Perron mean-shift specification on the modern sample beginning in 1995-01. BIC selects 6 breaks at 1998-03, 2001-09, 2007-12, 2013-09, 2016-12, 2020-03. The comparison columns use the last selected regime ending before 2007 as the benchmark, so the table directly shows whether the post-2007 regime means remain below the modern pre-crisis level.

Table B.4 reports a robustness exercise that re-estimates the mean-shift model on the modern sample beginning in 1995Q1. This specification again finds a clear break at 2007Q4, followed by additional post-crisis regimes. More importantly, every post-2007 regime mean remains below the last pre-2007 benchmark regime. The exact number of full-sample breaks is not central, but the persistent post-2007 downward shift in intermediary capital survives the more transparent selection rule and the modern-sample restriction.

As in the main text, we interpret these results as descriptive support for the paper’s mechanism. The Bai–Perron exercise documents a durable downward shift in intermediary capitalization after 2007, but it does not by itself identify whether that shift reflects the QE trap, post-crisis regulation, or other structural changes. That is why this test complements, rather than substitutes for, the mechanism-based evidence in Tests 1–4.

Online Appendix B

Market Failure, the QE Trap and Fiscal Escape

Harjoat Bhamra and Nicolo Ceneda

BO.1 Proofs for the Complete-Markets Benchmark

Proof of Proposition B.1.

When both agents can trade bonds and equity, the single risk source Z_t is spanned by two assets, so markets are dynamically complete. By the Second Welfare Theorem, every competitive equilibrium allocation is Pareto efficient and can be obtained as the solution to a social planner's problem.

Step 1: Social planner's problem. The planner maximises

$$\bar{\lambda} \left(\ln C_h - \frac{N^{1+\varphi}}{1+\varphi} \right) + (1 - \bar{\lambda}) \ln C_i$$

subject to the resource constraint $C_h + C_i = Y = AN$, where $\bar{\lambda} \in (0, 1)$ is the household's Pareto weight.

The Lagrangian is $\mathcal{L} = \bar{\lambda} \ln C_h + (1 - \bar{\lambda}) \ln C_i - \bar{\lambda} N^{1+\varphi} / (1 + \varphi) + \xi (AN - C_h - C_i)$ with multiplier ξ . The first-order conditions are:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial C_h} &= \frac{\bar{\lambda}}{C_h} - \xi = 0 \implies C_h = \frac{\bar{\lambda}}{\xi}, \\ \frac{\partial \mathcal{L}}{\partial C_i} &= \frac{1 - \bar{\lambda}}{C_i} - \xi = 0 \implies C_i = \frac{1 - \bar{\lambda}}{\xi}, \\ \frac{\partial \mathcal{L}}{\partial N} &= -\bar{\lambda} N^\varphi + \xi A = 0 \implies N^\varphi = \frac{\xi A}{\bar{\lambda}}. \end{aligned}$$

From the first two conditions: $C_h + C_i = 1/\xi = AN$, so $\xi = 1/(AN)$. Substituting into the consumption conditions:

$$C_h = \bar{\lambda} AN = \bar{\lambda} Y, \quad C_i = (1 - \bar{\lambda}) Y. \quad (\text{BO.1})$$

Consumption shares are constant, independent of A_t .

Step 2: Employment. The planner's labour FOC gives $N^{1+\varphi} = 1/\bar{\lambda}$, but this is the first-best level without the monopolistic competition distortion. In the decentralised equilibrium, the household's intratemporal condition equates the marginal rate of substitution to the real wage: $N^\varphi C_h = W/P$. From the factor shares (2.2): $W/P = [(\epsilon-1)/\epsilon] \cdot Y/N = [(\epsilon-1)/\epsilon] \cdot A$. Substituting $C_h = \bar{\lambda} Y = \bar{\lambda} AN$:

$$N^\varphi \cdot \bar{\lambda} AN = \frac{\epsilon-1}{\epsilon} \cdot A.$$

Cancelling A and solving: $N^{1+\varphi} = (\epsilon-1)/(\epsilon\bar{\lambda})$, giving:

$$N = \left[\frac{\epsilon-1}{\epsilon\bar{\lambda}} \right]^{1/(1+\varphi)}. \quad (\text{BO.2})$$

Since $\bar{\lambda}$, ϵ , and φ are all constant, N is constant (independent of A_t). Output is $Y_t = A_t N$, dividends are $D_t = Y_t/\epsilon = A_t N/\epsilon$, and both inherit the dynamics of A_t .

Step 3: Stochastic discount factor, risk-free rate, and market price of risk. We can use either agent's marginal utility to price assets. Using the household: the SDF is $\Lambda_t = e^{-\delta t} \cdot \partial U_h / \partial C_h = e^{-\delta t} / C_{h,t}$ (the marginal utility of consumption under log utility). Substituting $C_{h,t} = \bar{\lambda} A_t N$:

$$\Lambda_t = \frac{e^{-\delta t}}{\bar{\lambda} A_t N}.$$

Since $\bar{\lambda}$ and N are constant, the dynamics of Λ are determined by $e^{-\delta t} / A_t$. We compute $d(1/A_t)$ by Itô's lemma with $f(A) = 1/A$, $f'(A) = -1/A^2$, $f''(A) = 2/A^3$:

$$\begin{aligned} d\left(\frac{1}{A_t}\right) &= -\frac{1}{A_t^2} dA_t + \frac{1}{2} \cdot \frac{2}{A_t^3} (dA_t)^2 \\ &= -\frac{1}{A_t^2} (A_t \mu dt + A_t \sigma dZ_t) + \frac{1}{A_t^3} A_t^2 \sigma^2 dt \\ &= \frac{1}{A_t} (-\mu + \sigma^2) dt - \frac{1}{A_t} \sigma dZ_t. \end{aligned}$$

Therefore $d(1/A_t)/(1/A_t) = (-\mu + \sigma^2) dt - \sigma dZ_t$. Using the product rule for $\Lambda_t = (\bar{\lambda} N)^{-1} e^{-\delta t}$.

$(1/A_t)$:

$$\frac{d\Lambda_t}{\Lambda_t} = -\delta dt + \frac{d(1/A_t)}{1/A_t} = -(\delta + \mu - \sigma^2) dt - \sigma dZ_t.$$

The risk-free rate is minus the expected growth of the SDF:

$$r = -\frac{\mathbb{E}[d\Lambda_t/\Lambda_t]}{dt} = \delta + \mu - \sigma^2. \quad (\text{BO.3})$$

This is constant. The market price of risk is the negative of the SDF's volatility loading on dZ :

$$\Theta = \sigma. \quad (\text{BO.4})$$

This is also constant. The equity risk premium is $\sigma_R \Theta = \sigma \cdot \sigma = \sigma^2$.

Verification using the financier's SDF: $\Lambda_t^i = e^{-\delta t}/C_{i,t} = e^{-\delta t}/[(1 - \bar{\lambda})A_t N]$. Since this differs from Λ_t^h by only the constant factor $\bar{\lambda}/(1 - \bar{\lambda})$, both SDFs give the same r and Θ . This confirms dynamic completeness (marginal rates of substitution are equalised).

Step 4: Asset prices. The equity price is:

$$P_{D,t} = \mathbb{E}_t \int_t^\infty \frac{\Lambda_s}{\Lambda_t} D_s ds.$$

Computing the SDF ratio: $\Lambda_s/\Lambda_t = e^{-\delta(s-t)} \cdot (A_t/A_s)$ (from the explicit form of Λ , since $\bar{\lambda}$ and N cancel). Substituting $D_s = A_s N/\epsilon$:

$$P_{D,t} = \mathbb{E}_t \int_t^\infty e^{-\delta(s-t)} \frac{A_t}{A_s} \cdot \frac{A_s N}{\epsilon} ds = \frac{A_t N}{\epsilon} \int_t^\infty e^{-\delta(s-t)} ds = \frac{A_t N}{\epsilon} \cdot \frac{1}{\delta} = \frac{D_t}{\delta}.$$

The key step is that A_s cancels between the SDF ratio and the dividend: $\Lambda_s D_s/\Lambda_t = e^{-\delta(s-t)} A_t N/\epsilon$, which is deterministic conditional on $\mathcal{F}_t$. Similarly, total wealth is:

$$P_{Y,t} = \mathbb{E}_t \int_t^\infty \frac{\Lambda_s}{\Lambda_t} Y_s ds = A_t N \int_t^\infty e^{-\delta(s-t)} ds = \frac{Y_t}{\delta}.$$

Step 5: Value of human capital. By the factor shares (2.2), output decomposes as $Y_t = D_t + (W_t/P_t)N_t$. Linearity of pricing under the SDF gives $P_{Y,t} = P_{D,t} + P_{L,t}$, where $P_{L,t}$ is the present value of labour income. Hence:

$$P_{L,t} = P_{Y,t} - P_{D,t} = \frac{Y_t}{\delta} - \frac{D_t}{\delta} = \frac{Y_t}{\delta} - \frac{Y_t}{\epsilon\delta} = \left(1 - \frac{1}{\epsilon}\right) \frac{Y_t}{\delta} = \left(1 - \frac{1}{\epsilon}\right) P_{Y,t}.$$

Step 6: Agent wealth decomposition. Under log utility, each agent $j \in \{h, i\}$ has SDF $\Lambda_{j,t} = e^{-\delta t} / C_{j,t}$, so $\Lambda_{j,s} / \Lambda_{j,t} = e^{-\delta(s-t)} C_{j,t} / C_{j,s}$ and

$$\frac{\Lambda_{j,s}}{\Lambda_{j,t}} C_{j,s} = e^{-\delta(s-t)} \frac{C_{j,t}}{C_{j,s}} C_{j,s} = e^{-\delta(s-t)} C_{j,t}.$$

Total wealth (present value of consumption under the SDF) is therefore

$$W_{j,t} = \mathbb{E}_t \int_t^\infty \frac{\Lambda_{j,s}}{\Lambda_{j,t}} C_{j,s} ds = C_{j,t} \int_t^\infty e^{-\delta(s-t)} ds = \frac{C_{j,t}}{\delta}.$$

The household's total wealth decomposes into financial wealth and human capital: $W_{h,t} = H_{h,t} + P_{L,t}$. Using $C_{h,t} = \bar{\lambda} Y_t$ from (BO.1) and $P_{L,t} = (1 - 1/\epsilon) P_{Y,t}$ from Step 5:

$$H_{h,t} = W_{h,t} - P_{L,t} = \frac{\bar{\lambda} Y_t}{\delta} - \left(1 - \frac{1}{\epsilon}\right) P_{Y,t} = \left(\bar{\lambda} - \left(1 - \frac{1}{\epsilon}\right)\right) P_{Y,t}.$$

The financier has no labour income, so her financial wealth equals her total wealth:

$$H_{i,t} = W_{i,t} = \frac{C_{i,t}}{\delta} = \frac{(1 - \bar{\lambda}) Y_t}{\delta} = (1 - \bar{\lambda}) P_{Y,t}.$$

Step 7: Connection to the LP model at $\omega = 1/\epsilon$. In the limited-participation model, $\omega = H_i / P_Y$ is the financier's wealth share. At $\omega = 1/\epsilon$: $H_i = P_Y / \epsilon = P_D$ (the financier holds exactly the equity claim), $H_h = (1/\epsilon - \omega) P_Y = 0$ (the household has zero financial wealth — all her wealth is human capital $P_L = (1 - 1/\epsilon) P_Y$), $\phi = 1/(\epsilon\omega) = 1$ (no leverage), $C_i = \omega Y = Y/\epsilon$ (financier consumes the dividend share), $C_h = (1 - \omega) Y = (1 - 1/\epsilon) Y$ (household consumes the labour share). Moreover, when $\bar{\lambda} = 1 - 1/\epsilon$, the complete-markets formulas from Steps 1 and 6 give $C_{h,t} = (1 - 1/\epsilon) Y_t$, $C_{i,t} = Y_t / \epsilon$, $H_{h,t} = 0$, $H_{i,t} = P_{Y,t} / \epsilon$, and $\phi = 1$. Therefore the good steady state of the LP model ($\omega = 1/\epsilon$) coincides with the complete-markets benchmark indexed by $\bar{\lambda} = 1 - 1/\epsilon$. $\square$

BO.2 Auxiliary Results for the QE Dynamics

The following results are referenced in the main text but their formal statements are included here for completeness.

BO.2.1 Properties of the Volatility Function

Proposition BO.1 (Zeros and sign of σ_ω) *The function σ_ω has exactly two interior zeros: $\sigma_\omega(\underline{\omega}) = 0$ and $\sigma_\omega(1/\epsilon) = 0$. In addition, $\sigma_\omega(0) = 0$. Its sign pattern is:*

$$\sigma_\omega(\omega) \begin{cases} < 0 & \text{for } \omega \in (0, \underline{\omega}), \\ > 0 & \text{for } \omega \in (\underline{\omega}, 1/\epsilon), \\ < 0 & \text{for } \omega \in (1/\epsilon, 1). \end{cases} \quad (\text{BO.5})$$

The proof is in Appendix B.6. The sign pattern determines how the economy responds to productivity shocks. In the crisis regime ($\omega < \underline{\omega}$), $\sigma_\omega < 0$: positive productivity shocks *reduce* the financier's wealth share, worsening the crisis. This occurs because the de-leveraged financier ($\phi < 1$) holds less equity than her consumption share warrants, so her wealth grows more slowly than the economy in a boom. In the recovery regime ($\underline{\omega} < \omega < 1/\epsilon$), $\sigma_\omega > 0$: positive shocks improve the financier's capitalisation, aiding recovery. The zeros at $\underline{\omega}$ and $1/\epsilon$ mark the absorbing steady states where all fluctuations cease.

BO.2.2 Three Invariant Regimes

Theorem BO.1 (Regime classification) *The state space $(0, 1)$ partitions into three invariant regions:*

1. **Crisis regime** ($\omega \in (0, \underline{\omega})$): *the financier is severely undercapitalised, $\phi < 1$ (de-leveraged), and the central bank holds a majority of equity.*
2. **Recovery regime** ($\omega \in (\underline{\omega}, 1/\epsilon)$): *the leverage constraint may or may not bind, risk premia are elevated, and the economy is in the basin of attraction of the good steady state.*
3. **Inverted regime** ($\omega \in (1/\epsilon, 1)$): *the financier is "too large," households are net borrowers, and the economy converges toward the good steady state from above.*

No transition between regimes is possible.

BO.2.3 Boundary Classification

Proposition BO.2 (Boundary classification) (i) $\omega = 0$: *natural boundary (unattainable). Near*

$$\omega = 0, |\sigma_\omega| \sim \sigma \lambda \underline{\omega} \cdot \omega.$$

(ii) $\omega = \underline{\omega}$: *unattainable from both sides. Near $\underline{\omega}$, $\sigma_\omega = g(\omega)(\omega - \underline{\omega})$ with $g(\underline{\omega}) = \sigma \lambda \underline{\omega} > 0$.*

(iii) $\omega = 1/\epsilon$: *unattainable from both sides*. Near $1/\epsilon$, $\sigma_\omega = g(\omega)(1/\epsilon - \omega)$ with $g(1/\epsilon) = \sigma > 0$.

(iv) $\omega = 1$: *unattainable*. Near 1, $\sigma_\omega = h(\omega)(1 - \omega)$ with $h(1) = -\sigma(1 + \varphi) \neq 0$.

The proofs (Appendix B.9) use the Engelbert–Schmidt criterion. “Unattainable” means that the economy perpetually approaches but never reaches a boundary in finite time, a continuous-time analogue of Zeno’s paradox. Combined with regime invariance, this means the economy’s long-run fate is determined entirely by its initial condition: starting below $\underline{\omega}$, it converges to the bad steady state; starting above, it converges to the good one.

BO.3 Proof of Asset Pricing Formulae

Proof of Theorem 2.5.

Step 1: Return volatility. Since $P_{D,t} = Y_t/(\epsilon\delta)$ and δ is constant, the equity return has the same volatility as output: $\sigma_R = \sigma_Y$. Output is $Y_t = A_t N_t$ where A_t follows geometric Brownian motion with volatility σ and $N_t = N(\omega_t)$ with $N(\omega) = [(1 - 1/\epsilon)/(1 - \omega)]^{1/(1+\varphi)}$. Define $f(\omega) = (1 - \omega)^{-1/(1+\varphi)}$ so that $N = (1 - 1/\epsilon)^{1/(1+\varphi)} f(\omega)$. Then:

$$f'(\omega) = \frac{1}{(1 + \varphi)}(1 - \omega)^{-\frac{1}{1+\varphi}-1}, \quad \frac{f'(\omega)}{f(\omega)} = \frac{1}{(1 + \varphi)(1 - \omega)}.$$

By Itô’s lemma on $N(\omega_t) = c f(\omega_t)$, using $d\omega_t = (\sigma_\omega^2/\omega) dt + \sigma_\omega dZ$:

$$\frac{dN}{N} = \frac{f'}{f} d\omega + \frac{1}{2} \frac{f''}{f} (d\omega)^2 = \frac{\sigma_\omega}{(1 + \varphi)(1 - \omega)} dZ_t + \left[\frac{f'}{f} \frac{\sigma_\omega^2}{\omega} + \frac{1}{2} \frac{f''}{f} \sigma_\omega^2 \right] dt.$$

Since A and N load on the same Brownian motion Z , the product rule for $Y = AN$ gives

$$\sigma_Y = \sigma_A + \sigma_N:$$

$$\sigma_Y = \sigma + \frac{\sigma_\omega}{(1 + \varphi)(1 - \omega)},$$

which establishes (2.27).

Step 2: Market price of risk. The financier’s SDF is $\Lambda_t = e^{-\delta t}/(\delta H_{i,t})$ (from $C_{i,t} = \delta H_{i,t}$ under log utility). By Itô’s lemma on $1/H_i$ with $g(H) = 1/H$, $g' = -1/H^2$: $\sigma_\Lambda/\Lambda = -\sigma_{H_i}/H_i = -\phi_t \sigma_R$ (from equation (B.19)). The market price of risk is:

$$\Theta_t = \phi_t \sigma_{R,t}.$$

This holds for any ϕ_t (Appendix B.7). From $\sigma_\omega = \omega(\phi_t - 1)\sigma_R$ (Appendix B.6), $\phi_t = 1 + \sigma_\omega/(\omega\sigma_R)$, so:

$$\Theta_t = \phi_t \sigma_R = \sigma_R + \frac{\sigma_\omega}{\omega} = \underbrace{\sigma + \frac{\sigma_\omega}{(1 + \varphi)(1 - \omega)}}_{\sigma_R} + \frac{\sigma_\omega}{\omega},$$

which is (2.26).

Step 3: Risk-free rate. Total wealth $P_{Y,t} = Y_t/\delta$ pays continuous dividends $C_t = Y_t = \delta P_{Y,t}$ (dividend yield δ). Since $P_Y = Y/\delta$ with δ constant, $dP_Y/P_Y = dY/Y$, so $\mu_{P_Y}/P_Y = \mu_Y$ and $\sigma_{P_Y}/P_Y = \sigma_Y$. The no-arbitrage condition for a dividend-paying claim states that the expected total return (capital gain plus dividend yield) equals the risk-free rate plus the risk premium:

$$\frac{\mu_{P_Y}}{P_Y} + \delta = r_t + \sigma_Y \Theta_t.$$

Substituting $\mu_{P_Y}/P_Y = \mu_Y$ and rearranging: $r_t = \delta + \mu_{Y,t} - \sigma_{R,t} \Theta_t$, which is (2.28). The equity risk premium $\mu_{R,t} - r_t = \sigma_{R,t} \Theta_t$ follows directly from the definition of Θ . $\square$

BO.4 Equilibrium Quantities as Functions of the Wealth Share

This appendix illustrates how asset prices, risk premia, and real variables evolve as the financier's wealth share ω falls from the good steady state ($1/\epsilon$) toward the crisis threshold (ω_c) and the bad steady state ($\underline{\omega}$). All panels use the current Table 2.1 calibration: $\epsilon = 13.7741$, $\underline{\omega} = 0.0260$, $\lambda = 5.7231$, $\sigma = 0.1505$, $\delta = 0.0314$, $\varphi = 1.3333$, giving $\omega_c = 0.0606$ and $1/\epsilon \approx 0.0726$. Three vertical dashed lines mark these thresholds throughout.

The figures compare three model configurations (see Table B.1):

1. **LP only** (blue): limited participation, no leverage constraint (Appendix B.2). Equilibrium exists for all $\omega \in (0, 1/\epsilon)$. Rational bubble present.
2. **LP + leverage, no QE** (orange, dashed where shown): leverage constraint added, no central bank (Appendix B.4). Equilibrium identical to LP only for $\omega > \omega_c$; no equilibrium below ω_c (shaded region).
3. **LP + leverage + QE** (red): central bank absorbs residual equity below ω_c (Sections 2.3–2.6). Equilibrium exists everywhere. No rational bubble. QE Trap at $\underline{\omega}$.

BO.4.1 The Wealth-Share Volatility: The Fundamental Object

The wealth-share volatility $\sigma_\omega(\omega)$ is the primitive from which all other equilibrium quantities are derived. The market price of risk ($\Theta = \sigma_\omega / (\omega + \sigma_R)$), equity return volatility ($\sigma_R = \sigma + \sigma_\omega / [(1 + \varphi)(1 - \omega)]$), the risk-free rate ($r = \delta + \mu_Y - \sigma_R \Theta$), and the intermediation gap ($b'' = 2\delta b / \sigma_\omega^2$) all depend on σ_ω . The difference in the *boundary behaviour* of σ_ω near $\omega = 0$ is what separates equilibrium existence from failure, and rational bubbles from correct pricing.

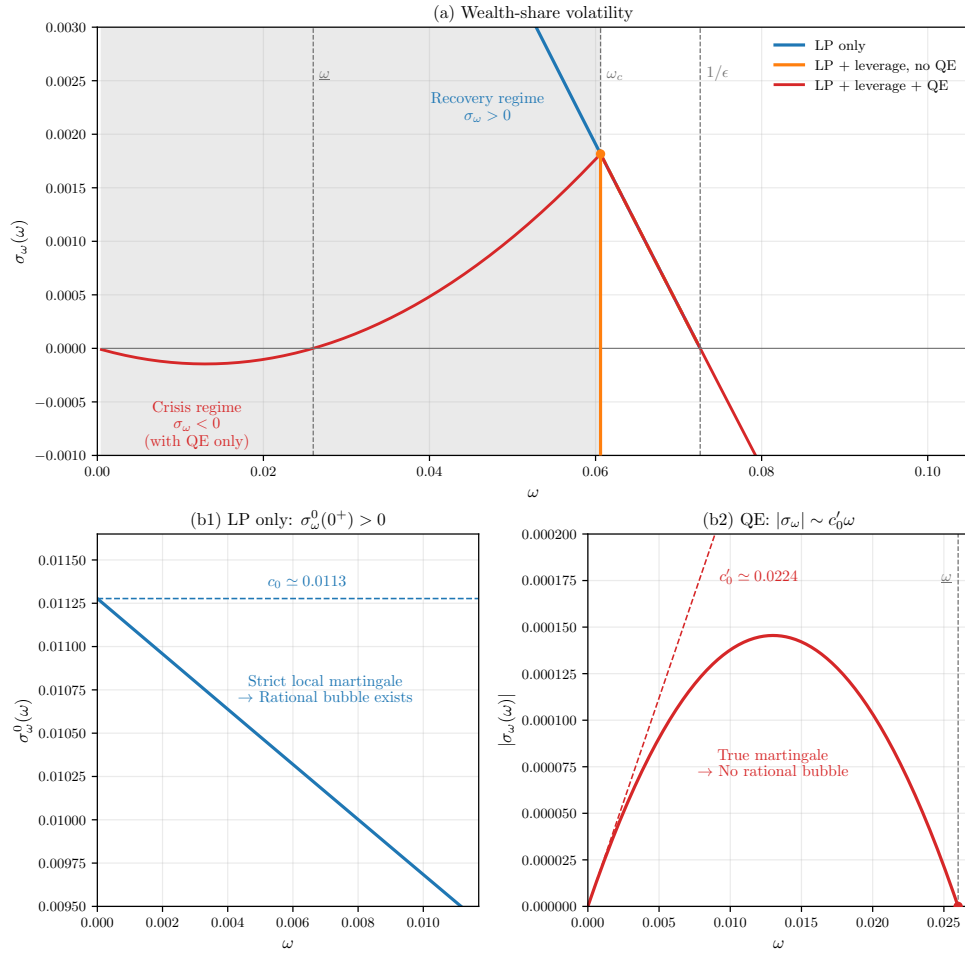


Figure BO.1: The wealth-share volatility $\sigma_\omega(\omega)$: the source of everything. *Panel (a):* Three cases. **LP only** (blue): $\sigma_\omega^0(\omega)$ from the unconstrained formula (B.4), defined on $(0, 1/\epsilon)$, positive throughout, single zero at $1/\epsilon$. **LP + leverage, no QE** (orange): identical to LP only for $\omega > \omega_c$ (constraint does not bind). Below ω_c (shaded): no equilibrium — the orange curve hits a cliff at ω_c . **LP + leverage + QE** (red): piecewise formula (2.23). Below ω_c the constrained formula applies: σ_ω changes sign at $\underline{\omega}$ (negative in crisis, positive in recovery) and vanishes at both absorbing states. The three cases agree above ω_c and diverge below it. *Panels (b1)–(b2):* Zooms near $\omega = 0$, showing the critical boundary behaviour. **LP only** (and LP + leverage without QE): $\sigma_\omega^0(0^+) = c_0 = \sigma(1+\varphi)(1/\epsilon)/[(1-1/\epsilon)+\varphi] \approx 0.0113 > 0$ (blue dashed). The volatility stays positive at the boundary. By the integral test, $Q = 1/\omega$ is a strict local martingale and the intermediation gap is positive. **LP + leverage + QE**: $|\sigma_\omega(\omega)| \sim c'_0 \omega$ with $c'_0 = \sigma \lambda \underline{\omega} \approx 0.0224$ (red dashed). The volatility vanishes linearly. Now Q is a true martingale and the gap is zero. One factor of ω — introduced by the leverage constraint capping ϕ — separates bubble from no bubble, and simultaneously equilibrium existence from failure (the leverage paradox).

BO.4.2 Portfolio Share and Wealth-Share Volatility

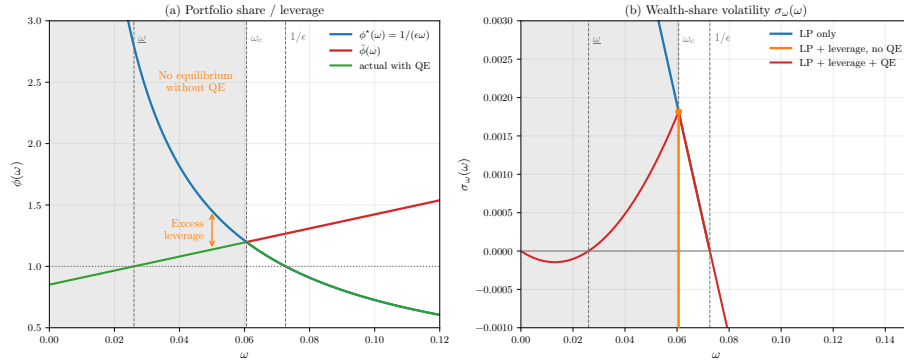


Figure BO.2: Portfolio share and volatility. *Panel (a):* The required portfolio share $\phi = 1/(\epsilon\omega)$ (blue) rises as the financier shrinks — a market clearing identity, not a portfolio choice. The leverage constraint $\bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega})$ (red) caps leverage. The actual portfolio share with QE (green) is $\phi = \min(1/(\epsilon\omega), \bar{\phi}(\omega))$: identical to the required share above ω_c and constrained below. The orange arrow marks the excess leverage at $\omega = 0.05$: without QE, this gap represents equity supply that no price can absorb. The horizontal dotted line marks $\phi = 1$ (no leverage): the financier is levered iff $\omega < 1/\epsilon$. *Panel (b):* Wealth-share volatility $\sigma_\omega(\omega)$ under the three model configurations. LP only (blue): the unconstrained formula σ_ω^0 on the full domain $(0, 1/\epsilon)$. LP + leverage, no QE (orange): identical to LP only above ω_c ; undefined below (cliff, marked by orange dot). LP + leverage + QE (red): the piecewise formula (2.23), with sign change at $\underline{\omega}$ and zeros at both absorbing states.

BO.4.3 Asset Pricing Quantities

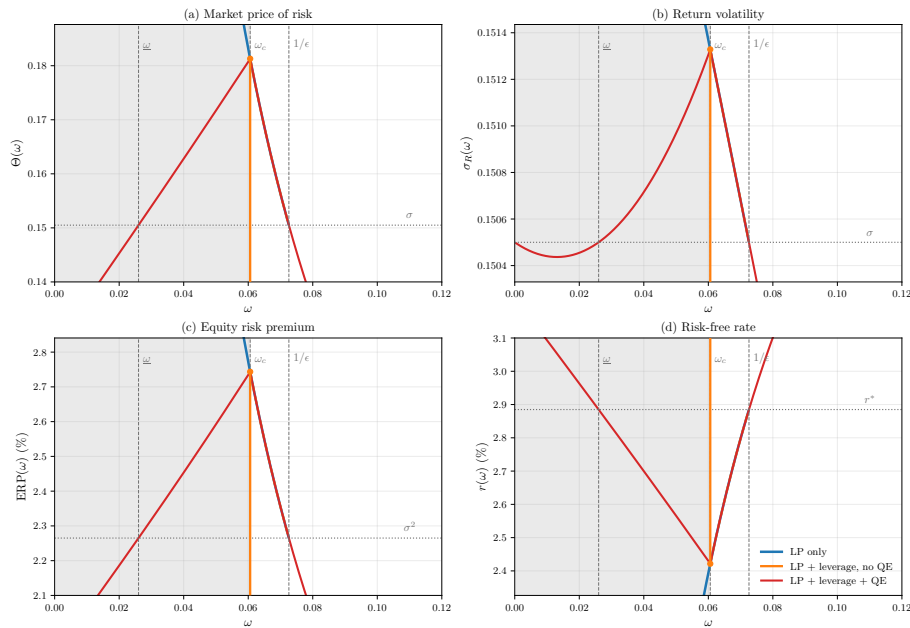


Figure BO.3: Asset pricing quantities as functions of ω . Each panel shows LP only (blue), LP + leverage without QE (orange), and LP + leverage + QE (red). Above ω_c the curves are identical (the constraint does not bind); below ω_c (shaded) the QE curve diverges, while the no-QE case is undefined (orange dot marks the cliff). Horizontal dotted lines mark complete-markets values. *Panel (a):* Market price of risk Θ . Blue (LP only) continues rising below ω_c as the unconstrained financier demands ever-higher compensation. Red (with QE) falls back toward $\sigma = 0.1505$ as the constrained σ_ω shrinks. Peak at ω_c : $\Theta \approx 0.181$. *Panel (b):* Return volatility σ_R . Modest amplification; peak ≈ 0.151 vs $\sigma = 0.1505$. *Panel (c):* Equity risk premium $\sigma_R\Theta$. Peaks at 2.74% near ω_c (vs $\sigma^2 = 2.27\%$ frictionless). With QE, the premium falls below ω_c — risk premia normalise even as the real economy deteriorates. *Panel (d):* Risk-free rate r . Mirror image: trough 2.42% at ω_c (vs $r^* = 2.88\%$). With QE, r rises back to normal below ω_c . A policymaker watching these panels would see risk premia falling and r rising below ω_c and might conclude the crisis is easing — but employment (Figure BO.4) continues to worsen.

BO.4.4 Real Variables and the Intermediation Gap

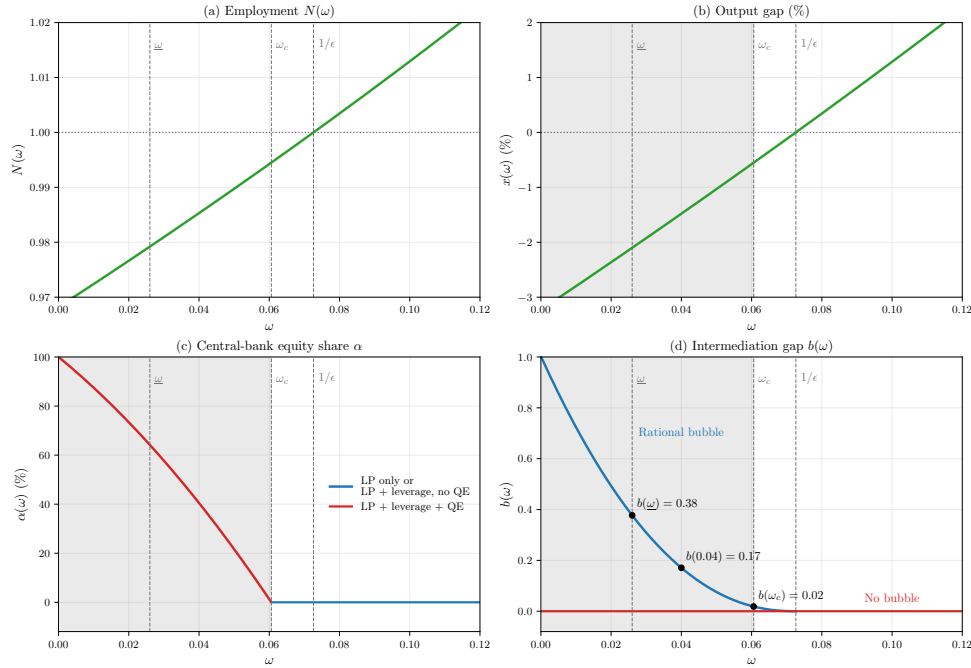


Figure BO.4: Real variables and intermediation gap. Panel (a): Employment $N(\omega) = [(1 - 1/\epsilon)/(1 - \omega)]^{1/(1+\varphi)}$ is the same in all three model configurations: it depends on ω , not on whether the central bank holds equity. N is monotonically increasing and convex; at $\underline{\omega}$: $N \approx 0.979$. Panel (b): Output gap widens monotonically as ω falls. Without QE, the gap is defined only above ω_c (about -0.6% at the cliff); below ω_c there is no equilibrium. With QE, the gap reaches -2.1% at $\underline{\omega}$. Panel (c): Central bank equity share (with QE only). Above ω_c : $\alpha = 0$. Below ω_c : α rises as the constrained financier holds less. At $\underline{\omega}$: the CB permanently holds about 64% of the equity market. Without QE, $\alpha = 0$ always (no CB) — but no equilibrium exists below ω_c . Panel (d): Intermediation gap $b(\omega) = B/P_D$ from the exact gap ODE $b'' = (2\delta/\sigma_\omega^2)b$. Under **LP only** and **LP + leverage, no QE**: $b > 0$ (blue, with shading). At $\underline{\omega}$, $b \approx 0.38$; at $\omega = 0.04$, $b \approx 0.17$; and at ω_c , $b \approx 0.02$. Under **LP + leverage + QE**: $b \equiv 0$ (red, dashed). The leverage constraint makes Q a true martingale, eliminating the gap. The shaded area is phantom wealth — value that would vanish if the economy transitioned to a regime with correct pricing.

BO.5 Intermediation Gap Formula Derivation

This appendix derives the ODE characterising the gap-to-price ratio $b(\omega) \equiv B_t/P_{D,t}$ in the baseline limited-participation model without leverage constraints, and establishes its boundary conditions. The derivation is self-contained: it does not rely on the finite-horizon formula of Hugonnier (2012), which is exact only in an exchange economy with constant weighting-process volatility. In our production economy the weighting-process volatility is state-dependent, and the exact gap solves a singular Sturm–Liouville problem derived below.

Setup and objective

The wealth-share process $\omega_t \in (0, 1/\epsilon)$ satisfies

$$d\omega_t = \mu_\omega(\omega_t) dt + \sigma_\omega(\omega_t) dZ_t, \quad (\text{BO.6})$$

with $\mu_\omega = \sigma_\omega^2/\omega$ (the Q -martingale identity, Appendix B.7), and, in the unconstrained baseline, $\sigma_\omega(0^+) = c_0 > 0$ and $\sigma_\omega(1/\epsilon) = 0$. The equity price is $P_{D,t} = Y_t/(\epsilon\delta)$ (market clearing with a log-utility financier) and the fundamental value is

$$P_{D,t}^{\text{fund}} \equiv \mathbb{E}_t \left[\int_t^\infty \frac{\Lambda_s}{\Lambda_t} D_s ds \right], \quad (\text{BO.7})$$

where $\Lambda_t = e^{-\delta t}/C_{i,t}$ is the financier's stochastic discount factor and $D_t = Y_t/\epsilon$. The intermediation gap-to-price ratio is $b(\omega_t) \equiv 1 - P_{D,t}^{\text{fund}}/P_{D,t}$.

Substituting $\Lambda_s/\Lambda_t = e^{-\delta(s-t)} \cdot (\omega_t Y_t)/(\omega_s Y_s)$ and $D_s = Y_s/\epsilon$ into (BO.7), with change of variables $u = s - t$:

$$b(\omega) = 1 - \delta \omega f(\omega), \quad f(\omega) \equiv \mathbb{E}^\omega \left[\int_0^\infty e^{-\delta u} \frac{1}{\omega_u} du \right], \quad (\text{BO.8})$$

where $\mathbb{E}^\omega$ denotes expectation conditional on $\omega_0 = \omega$. The function f is the Laplace–transform representation of the Q -resolvent; b is the quantity of economic interest.

Step 1: Feynman–Kac ODE for f

Lemma BO.1 (Feynman–Kac representation) *Under Assumption 2.2 and $\mu_\omega = \sigma_\omega^2/\omega$, the function f in (BO.8) belongs to $C^2((0, 1/\epsilon))$ and solves*

$$\frac{1}{2}\sigma_\omega^2(\omega) f''(\omega) + \frac{\sigma_\omega^2(\omega)}{\omega} f'(\omega) - \delta f(\omega) + \frac{1}{\omega} = 0, \quad \omega \in (0, 1/\epsilon). \quad (\text{BO.9})$$

Proof. Apply Karatzas and Shreve (1991, Theorem 5.7.6 and Problem 5.7.7, pp. 366–367) with infinitesimal generator $\mathcal{L} = \frac{1}{2}\sigma_\omega^2\partial_\omega^2 + \mu_\omega\partial_\omega$, discount rate $\delta > 0$, and running payoff $g(\omega) = 1/\omega$. The required integrability $\mathbb{E}^\omega \int_0^\infty e^{-\delta u} (1/\omega_u) du < \infty$ holds because $Q_t = 1/\omega_t$ is a non-negative supermartingale (Appendix B.7), so $\mathbb{E}^\omega[Q_u] \leq Q_0 = 1/\omega$ and the integral is bounded by $1/(\delta\omega) < \infty$ for every interior ω . Elliptic regularity on $(0, 1/\epsilon)$ (strict positivity of σ_ω^2 and locally Lipschitz coefficients) gives $f \in C^2$ on the interior. Substituting $\mu_\omega = \sigma_\omega^2/\omega$ into $\mathcal{L}f - \delta f + g = 0$

yields (BO.9). $\square$

Step 2: The ODE for b

Proposition BO.3 (ODE for the gap-to-price ratio) *The gap-to-price ratio $b(\omega) = 1 - \delta\omega f(\omega)$ satisfies the singular Sturm–Liouville equation*

$$b''(\omega) = \frac{2\delta}{\sigma_\omega^2(\omega)} b(\omega), \quad \omega \in (0, 1/\epsilon). \quad (\text{BO.10})$$

Proof. Write $f = (1 - b)/(\delta\omega)$ and compute the derivatives

$$f'(\omega) = -\frac{b'(\omega)}{\delta\omega} - \frac{1 - b(\omega)}{\delta\omega^2}, \quad (\text{BO.11})$$

$$f''(\omega) = -\frac{b''(\omega)}{\delta\omega} + \frac{2b'(\omega)}{\delta\omega^2} + \frac{2(1 - b(\omega))}{\delta\omega^3}. \quad (\text{BO.12})$$

Substitute (BO.11) and (BO.12) into (BO.9) and group terms of each order in $\omega^{-k}b^{(j)}$.

Coefficient of $b'(\omega)$. The only contributions to b' come from

$$\frac{1}{2}\sigma_\omega^2 \cdot \frac{2b'}{\delta\omega^2} + \frac{\sigma_\omega^2}{\omega} \cdot \left(-\frac{b'}{\delta\omega}\right) = \frac{\sigma_\omega^2 b'}{\delta\omega^2} - \frac{\sigma_\omega^2 b'}{\delta\omega^2} = 0.$$

The cancellation is *exact* and relies on the Q -martingale identity $\mu_\omega = \sigma_\omega^2/\omega$.

Coefficient of $(1 - b)/\omega^3$. The contributions are

$$\frac{1}{2}\sigma_\omega^2 \cdot \frac{2(1 - b)}{\delta\omega^3} + \frac{\sigma_\omega^2}{\omega} \cdot \left(-\frac{1 - b}{\delta\omega^2}\right) = \frac{\sigma_\omega^2(1 - b)}{\delta\omega^3} - \frac{\sigma_\omega^2(1 - b)}{\delta\omega^3} = 0.$$

Again a clean cancellation by $\mu_\omega = \sigma_\omega^2/\omega$.

Coefficient of $1/\omega$. From $-\delta f = -(1 - b)/\omega = (b - 1)/\omega$ and the forcing $+1/\omega$:

$$\frac{b - 1}{\omega} + \frac{1}{\omega} = \frac{b}{\omega}.$$

Coefficient of $b''(\omega)$. Only one contribution:

$$\frac{1}{2}\sigma_\omega^2 \cdot \left(-\frac{b''}{\delta\omega}\right) = -\frac{\sigma_\omega^2 b''}{2\delta\omega}.$$

Assembling the surviving terms, the ODE (BO.9) reduces to

$$-\frac{\sigma_\omega^2(\omega) b''(\omega)}{2\delta\omega} + \frac{b(\omega)}{\omega} = 0.$$

Multiplying through by $-2\delta\omega$ yields $\sigma_\omega^2(\omega) b''(\omega) - 2\delta b(\omega) = 0$, equivalently (BO.10). $\square$

Remark BO.1 *The cancellation of the b' and $(1-b)/\omega^3$ terms is a direct consequence of the Q -martingale identity $\mu_\omega = \sigma_\omega^2/\omega$. Any diffusion satisfying this identity yields the same clean ODE (BO.10) for the gap-to-price ratio; the specific form of $\sigma_\omega(\omega)$ affects only the coefficient $2\delta/\sigma_\omega^2(\omega)$, not the structure of the equation.*

Step 3: Boundary condition at $\omega = 1/\epsilon$

Lemma BO.2 (Upper boundary) $b(1/\epsilon) = 0$.

Proof. At $\omega = 1/\epsilon$, both $\sigma_\omega(1/\epsilon) = 0$ and $\mu_\omega(1/\epsilon) = \sigma_\omega^2(1/\epsilon) \cdot \epsilon = 0$, so the point is absorbing. By Yamada–Watanabe pathwise uniqueness (Karatzas and Shreve, 1991, Proposition 5.2.13, p. 291), if $\omega_0 = 1/\epsilon$ then $\omega_u = 1/\epsilon$ a.s. for all $u \geq 0$. Hence $f(1/\epsilon) = \int_0^\infty e^{-\delta u} \epsilon \, du = \epsilon/\delta$ and $b(1/\epsilon) = 1 - \delta \cdot (1/\epsilon) \cdot (\epsilon/\delta) = 0$. $\square$

Step 4: Boundary condition at $\omega = 0^+$

Lemma BO.3 (Lower boundary) *In the unconstrained baseline (no leverage constraint, $\sigma_\omega(0^+) = c_0 > 0$), $b(0^+) = 1$.*

Proof. The claim is equivalent to $\omega f(\omega) \rightarrow 0$ as $\omega \rightarrow 0^+$, i.e. $\delta\omega \int_0^\infty e^{-\delta u} \mathbb{E}^\omega[Q_u] \, du \rightarrow 0$, where $Q = 1/\omega$.

Step (a): Q -diffusion. By Itô's lemma applied to $Q = 1/\omega$ under (BO.6),

$$dQ_t = \left[\frac{\sigma_\omega^2}{\omega^3} - \frac{\mu_\omega}{\omega^2} \right] dt - \frac{\sigma_\omega}{\omega^2} dZ_t.$$

The drift vanishes under $\mu_\omega = \sigma_\omega^2/\omega$, so Q is a continuous local martingale on (ϵ, ∞) with diffusion coefficient $|\sigma_Q(q)| = q^2 \sigma_\omega(1/q)$. At the infinite endpoint $\sigma_\omega(0^+) = c_0$, so $|\sigma_Q(q)| \sim c_0 q^2$ as $q \rightarrow \infty$ (quadratic growth).

Step (b): Q is a strict local martingale. By Mijatović and Urusov (2012, Theorem 2.1), a continuous local martingale Q with diffusion σ_Q on a half-line is a *true* martingale iff

$\int^\infty q/\sigma_Q^2(q) \, dq = +\infty$. Here $\int^\infty q/(c_0^2 q^4) \, dq = c_0^{-2} \int^\infty q^{-3} \, dq < \infty$, so Q is a *strict* local martingale.

Step (c): h-transform identity. By [Pal and Protter \(2010, Theorem 1\)](#), for a non-negative continuous strict local martingale Q started at $Q_0 = q$, there exists a probability measure $\mathbb{P}^*$ (the Föllmer measure) under which $q \mapsto \mathbb{E}_q[Q_u]/q$ coincides with the survival probability of a *dual* (Föllmer) diffusion $\tilde{Q}$ whose explosion time ζ satisfies

$$\frac{\mathbb{E}_q[Q_u]}{q} = \mathbb{P}_q^*[\zeta > u].$$

Applied to the Q -diffusion of step (a) started at $q = 1/\omega$:

$$\omega \mathbb{E}^\omega[Q_u] = \mathbb{P}_{1/\omega}^*[\zeta > u].$$

Step (d): Collapse of $\delta\omega f(\omega)$ as $\omega \rightarrow 0^+$. Multiplying [\(BO.8\)](#) by $\delta\omega$ and applying Tonelli:

$$\delta\omega f(\omega) = \delta \int_0^\infty e^{-\delta u} \omega \mathbb{E}^\omega[Q_u] \, du = \delta \int_0^\infty e^{-\delta u} \mathbb{P}_{1/\omega}^*[\zeta > u] \, du = \mathbb{E}_{1/\omega}^*[1 - e^{-\delta\zeta}].$$

The last equality is Fubini applied to the joint law of (ζ, U) with $U \sim \text{Exp}(\delta)$ independent.

Step (e): Explosion in $\mathbb{P}^*$ is immediate as $\omega \rightarrow 0^+$. Under the quadratic-growth hypothesis $|\sigma_Q(q)| \sim c_0 q^2$ as $q \rightarrow \infty$, the Föllmer diffusion $\tilde{Q}$ is the h -transform of Q by $h(q) = q$ on (ϵ, ∞) . The dual process is transient with $+\infty$ as an *entrance* boundary, and $\zeta \rightarrow 0$ in $\mathbb{P}^*$ -probability as $\tilde{Q}_0 \rightarrow \infty$ (see [Pal and Protter 2010, Propositions 2 and 3](#); analogous to the inverse three-dimensional Bessel process [Delbaen and Schachermayer 1995](#)). As $\omega \rightarrow 0^+$, i.e. $1/\omega \rightarrow \infty$, we therefore have $\zeta \rightarrow 0$ in $\mathbb{P}_{1/\omega}^*$ -probability, so

$$\mathbb{E}_{1/\omega}^*[1 - e^{-\delta\zeta}] \longrightarrow 0 \quad \text{as } \omega \rightarrow 0^+.$$

Combining steps (d) and (e), $\delta\omega f(\omega) \rightarrow 0$ as $\omega \rightarrow 0^+$, so $b(\omega) = 1 - \delta\omega f(\omega) \rightarrow 1$. $\square$

Step 5: Existence, uniqueness, monotonicity, convexity

Proposition BO.4 (Well-posedness) *The two-point boundary-value problem $b''(\omega) = \frac{2\delta}{\sigma_\omega^2(\omega)} b(\omega)$ on $(0, 1/\epsilon)$ with $b(0^+) = 1$ and $b(1/\epsilon) = 0$ admits a unique C^2 solution; the solution is strictly positive, strictly decreasing, and strictly convex on $(0, 1/\epsilon)$.*

Proof. Write the equation in Liouville normal form $b'' - V(\omega)b = 0$ with potential $V(\omega) = 2\delta/\sigma_\omega^2(\omega) > 0$. On the interior, $V \in C^1$ and finite, so Picard–Lindelöf gives a unique local C^2 solution for any Cauchy data.

At $\omega = 0^+$: $V(0^+) = 2\delta/c_0^2 < \infty$, so $\omega = 0$ is a regular endpoint of the Sturm–Liouville operator in the sense of Zettl (2005, Chapter 7); the Dirichlet datum $b(0^+) = 1$ uniquely selects a local solution. At $\omega = 1/\epsilon$: $V \rightarrow \infty$, a limit-point singular endpoint in Weyl’s sense (Zettl, 2005, Theorem 7.4.1); under the vanishing rate $\sigma_\omega^2(\omega) \sim K(1/\epsilon - \omega)$, the Dirichlet condition $b(1/\epsilon) = 0$ selects the unique decaying solution. Positivity of V makes the Sturm–Liouville operator positive definite, so existence and uniqueness of the two-point BVP follow from Zettl (2005, Chapter 10).

Strict positivity. The strong maximum principle (Protter and Weinberger, 1984, Chapter 1, Theorem 3) applied to $L = -\partial_\omega^2 + V$ with $V > 0$ and non-negative boundary data $b(0^+) = 1$, $b(1/\epsilon) = 0$ gives $b > 0$ in the interior.

Strict convexity. Since $b > 0$ and $b'' = V \cdot b > 0$ on $(0, 1/\epsilon)$, b is strictly convex.

Strict monotone decrease. Suppose for contradiction that $b'(\omega_*) \geq 0$ at some $\omega_* \in (0, 1/\epsilon)$. Strict convexity of b implies b' is strictly increasing, so $b'(\omega) > b'(\omega_*) \geq 0$ for $\omega > \omega_*$, hence $b(\omega) > b(\omega_*) > 0$ on $(\omega_*, 1/\epsilon)$. By continuity, $b(1/\epsilon) \geq b(\omega_*) > 0$, contradicting $b(1/\epsilon) = 0$. Therefore $b'(\omega) < 0$ on the interior. $\square$

Connection to the Hugonnier (2012) exchange-economy formula

Remark BO.2 (Exchange-economy limit) When the weighting-process volatility $v_\lambda(\omega) = \frac{\sigma_\omega(\omega)}{\omega(1-\omega)}$ is constant (Hugonnier’s finite-horizon exchange economy with constant labour endowment), the ordinary differential equation (BO.10) admits the closed-form solution

$$b(\omega) = (1 - \omega)^\gamma, \quad \gamma = \frac{1 + \sqrt{1 + 8\delta/v^2}}{2}, \quad (\text{BO.13})$$

in the limit of infinite horizon. This recovers Hugonnier (2012, Proposition 6) as a special case of (BO.10). In our production economy, $v_\lambda(\omega)$ is genuinely state-dependent, so the closed form does not apply; the gap-to-price ratio must be computed by numerical solution of (BO.10).

Step 6: Numerical solution and cross-validation

At the calibrated parameters from Table 2.1, we solve (BO.10) numerically using two independent schemes and cross-validate.

Method A (Shooting, verified by collocation). We start at $\omega_R = 1/\epsilon - \eta$ for $\eta = 10^{-3}$ with the asymptotic Frobenius initial conditions $b(\omega_R) = \eta^\alpha$, $b'(\omega_R) = -\alpha \eta^{\alpha-1}$, where $\alpha = (1 + \sqrt{1 + 8\delta/\sigma^2})/2 \approx 2.2386$ is the right-endpoint indicial exponent, and integrate leftward to $\omega_L = 5 \times 10^{-4}$ using explicit Runge–Kutta (`scipy.integrate.solve_ivp`, RK45, relative tolerance 10^{-10}). The integrated solution is then rescaled so that $b(\omega_L) = 1$. The same BVP is independently solved by the collocation method of `scipy.integrate.solve_bvp` with Dirichlet BCs; the two solutions agree to six decimal places on $[\omega_L, \omega_R]$, confirming that the shooting rescaling recovers the correct normalisation.

Method B (Finite difference). A separate implementation discretises (BO.10) on a uniform grid of 600 points in $[10^{-4}, 1/\epsilon - 10^{-4}] \approx [10^{-4}, 0.0725]$ and solves the tridiagonal linear system with Dirichlet BCs $b(10^{-4}) = 1$, $b(1/\epsilon - 10^{-4}) = 0$, followed by a monotonicity projection $b_i \leftarrow \min(b_i, b_{i-1})$.

Cross-validation. Table BO.1 reports both methods at the empirically relevant grid.

Table BO.1: Numerical solution of the exact gap ODE (BO.10) at the calibrated parameters in Table 2.1. Two independent solvers agree to within 1.1×10^{-2} ; the approximate Hugonnier formula with state-dependent $\gamma(v_\lambda(\omega))$ is shown for reference.

ω	$v_\lambda(\omega)$	b_A (shoot/BVP)	b_B (FD)	$ b_A - b_B $	$(1 - \omega)^{\gamma(v_\lambda(\omega))}$
0.005	2.106	0.8677	0.8573	0.0105	0.9949
0.010	0.978	0.7326	0.7237	0.0089	0.9894
0.020	0.413	0.4989	0.4929	0.0060	0.9744
0.030	0.225	0.3129	0.3091	0.0038	0.9489
0.040	0.130	0.1729	0.1708	0.0021	0.9031
0.050	0.072	0.0765	0.0756	0.0009	0.8145
0.060	0.034	0.0208	0.0206	0.0003	0.6123
0.065	0.019	0.0067	0.0067	0.0001	0.3961
0.070	0.006	0.0006	0.0006	0.0000	0.0470
0.071	0.004	0.00021	0.00020	0.0000	0.0062
0.0715	0.002	0.00009	0.00009	0.0000	0.00056

Boundary conditions, positivity, monotonicity. Method A reports $b \rightarrow 1$ at the left endpoint

(by construction) and $b(\omega_R) = 7.21 \times 10^{-5}$ at $\omega_R = 1/\epsilon - 10^{-3} \approx 0.0716$, within the Frobenius asymptote's tolerance. The solution is strictly positive, strictly decreasing, and strictly convex at every grid point, consistent with Proposition BO.4.

Exact versus approximate formula. At $\omega = 0.05$, $b_{\text{exact}} = 0.0765$ while $(1 - \omega)^{\gamma(v_\lambda(0.05))} = 0.8145$, an overestimate of roughly a factor of eleven. At $\omega = 0.06$, the overestimate is roughly a factor of thirty. Near the numerical right endpoint, for example at $\omega = 0.0715$, both values are small, but the local Hugonnier expression still exceeds the exact value by roughly a factor of six. The qualitative properties (positivity, monotonicity, convexity, and elimination of the gap by central-bank intervention) are preserved by both formulas; only the magnitudes differ.

BO.6 Franchise Value and Dynamic Incentive Compatibility

Proposition BO.5 (Dynamic incentive compatibility) *With log utility, the dynamic IC constraint reduces to*

$$\phi_t \leq \frac{1}{\theta} \exp(\delta \mathcal{F}(\omega_t)) \equiv \bar{\phi}^{\text{dyn}}(\omega_t), \quad (\text{BO.14})$$

where the franchise value of intermediation is

$$\mathcal{F}(\omega) \equiv \frac{1}{2\delta} \mathbb{E}_\omega \int_0^\infty e^{-\delta t} \Theta_t^2 dt > 0. \quad (\text{BO.15})$$

Proof. Step 1: Value function decomposition. A log-utility agent with wealth W_t , access to a risky asset with return $\mu_R dt + \sigma_R dZ$ and a risk-free bond at rate r_t , consuming $C_t = \delta W_t$ and holding portfolio share ϕ_t , has lifetime utility $V_t = \mathbb{E}_t \int_t^\infty e^{-\delta(u-t)} \ln C_u du$. Substituting $C_u = \delta W_u$:

$$V_t = \mathbb{E}_t \int_t^\infty e^{-\delta(u-t)} \ln(\delta W_u) du = \frac{\ln \delta}{\delta} + \mathbb{E}_t \int_t^\infty e^{-\delta(u-t)} \ln W_u du.$$

Writing $\ln W_u = \ln W_t + \ln(W_u/W_t)$ and using $\int_t^\infty e^{-\delta(u-t)} du = 1/\delta$:

$$V_t = \frac{\ln(\delta W_t)}{\delta} + \mathbb{E}_t \int_t^\infty e^{-\delta(u-t)} \ln \frac{W_u}{W_t} du.$$

The wealth SDE with consumption drain $C = \delta W$ is $dW/W = [r + \phi(\mu_R - r) - \delta] dt + \phi \sigma_R dZ$.

By Itô's lemma on $\ln W$ with $f(W) = \ln W$, $f' = 1/W$, $f'' = -1/W^2$:

$$d \ln W = \left[r + \phi(\mu_R - r) - \delta - \frac{1}{2}\phi^2\sigma_R^2 \right] dt + \phi\sigma_R dZ.$$

Integrating and taking expectations (the stochastic integral has zero expectation):

$$\mathbb{E}_t \left[\ln \frac{W_u}{W_t} \right] = \mathbb{E}_t \int_t^u \left[r_s + \phi_s(\mu_{R,s} - r_s) - \delta - \frac{1}{2}\phi_s^2\sigma_{R,s}^2 \right] ds.$$

Substituting back and swapping the order of integration (Fubini): the region $t \leq s \leq u < \infty$ becomes, for fixed $s, u \in [s, \infty)$:

$$\begin{aligned} \mathbb{E}_t \int_t^\infty e^{-\delta(u-t)} \ln \frac{W_u}{W_t} du &= \mathbb{E}_t \int_t^\infty \int_t^u e^{-\delta(u-t)} g(s) ds du \\ &= \mathbb{E}_t \int_t^\infty g(s) \left(\int_s^\infty e^{-\delta(u-t)} du \right) ds = \frac{1}{\delta} \mathbb{E}_t \int_t^\infty e^{-\delta(s-t)} g(s) ds, \end{aligned}$$

where $g(s) = r_s + \phi_s(\mu_{R,s} - r_s) - \delta - \frac{1}{2}\phi_s^2\sigma_{R,s}^2$ and we used $\int_s^\infty e^{-\delta(u-t)} du = e^{-\delta(s-t)}/\delta$. Therefore:

$$V_t = \frac{\ln(\delta W_t)}{\delta} + \frac{1}{\delta} \mathbb{E}_t \int_t^\infty e^{-\delta(u-t)} \left(r_u + \phi_u(\mu_{R,u} - r_u) - \frac{1}{2}\phi_u^2\sigma_{R,u}^2 - \delta \right) du. \quad (\text{BO.16})$$

Step 2: Intermediary's value. The financier is the marginal investor with $\Theta_t = \phi_t\sigma_R$ and $\mu_R - r = \sigma_R\Theta = \phi_t\sigma_R^2$. The portfolio return term simplifies:

$$\phi(\mu_R - r) - \frac{1}{2}\phi^2\sigma_R^2 = \phi \cdot \phi\sigma_R^2 - \frac{1}{2}\phi^2\sigma_R^2 = \frac{1}{2}\phi^2\sigma_R^2 = \frac{1}{2}\Theta^2.$$

Therefore, $V^{\text{int}}(W, \omega) = [\ln(\delta W)]/\delta + (1/\delta)\mathbb{E}_\omega \int_0^\infty e^{-\delta u} [\frac{1}{2}\Theta_u^2 + r_u - \delta] du$.

Step 3: Autarky value. A financier in autarky holds only bonds ($\phi = 0$, so $\Theta = 0$): $V^{\text{aut}}(W, \omega) = [\ln(\delta W)]/\delta + (1/\delta)\mathbb{E}_\omega \int_0^\infty e^{-\delta u} [r_u - \delta] du$.

Step 4: IC constraint. The IC constraint is $V^{\text{int}}(H_i, \omega) \geq V^{\text{aut}}(\theta\phi H_i, \omega)$. Subtracting:

$$\begin{aligned} V^{\text{int}} - V^{\text{aut}} &= \frac{\ln(\delta H_i) - \ln(\delta\theta\phi H_i)}{\delta} + \frac{1}{\delta} \mathbb{E}_\omega \int_0^\infty e^{-\delta u} \frac{1}{2}\Theta_u^2 du \\ &= \frac{-\ln(\theta\phi)}{\delta} + \mathcal{F}(\omega), \end{aligned}$$

where $\mathcal{F}(\omega) = (1/(2\delta))\mathbb{E}_\omega \int_0^\infty e^{-\delta u} \Theta_u^2 du$ is the franchise value (BO.15). The IC condition $V^{\text{int}} \geq V^{\text{aut}}$ becomes $\mathcal{F}(\omega) \geq \ln(\theta\phi)/\delta$, i.e., $\theta\phi \leq \exp(\delta\mathcal{F}(\omega))$, yielding (BO.14). $\square$

Proposition BO.6 (Properties of the franchise value) (i) $\mathcal{F}(\omega) > 0$ for all $\omega \in (0, 1)$.

(ii) $\mathcal{F}$ solves the Feynman–Kac equation $\frac{1}{2}\sigma_\omega^2 \mathcal{F}'' + (\sigma_\omega^2/\omega)\mathcal{F}' - \delta\mathcal{F} + \Theta^2/(2\delta) = 0$.

(iii) At both absorbing steady states: $\mathcal{F}(\underline{\omega}) = \mathcal{F}(1/\epsilon) = \sigma^2/(2\delta^2)$.

(iv) In the recovery regime: $\mathcal{F}(\omega) > \sigma^2/(2\delta^2)$ (the franchise value is elevated when frictions are strongest).

Proof. (i) The market price of risk satisfies $\Theta_t = \phi_t \sigma_R > 0$ whenever $\phi_t > 0$ and $\sigma_R > 0$, which holds on a set of positive Lebesgue measure a.s. (since the economy starts in the interior of $(0, 1)$). Therefore $\mathcal{F}(\omega) = (1/(2\delta))\mathbb{E}_\omega \int_0^\infty e^{-\delta t} \Theta_t^2 dt > 0$.

(ii) Define $u(\omega) = \mathbb{E}_\omega \int_0^\infty e^{-\delta t} \Theta(\omega_t)^2 dt$. By the Feynman–Kac theorem (Karatzas and Shreve, 1991, Ch. 4), u solves $\frac{1}{2}\sigma_\omega^2 u'' + \mu_\omega u' - \delta u + \Theta^2 = 0$. (The theorem applies because $\Theta(\omega)$ is continuous and bounded on each invariant regime— ϕ is bounded by the leverage constraint—and $\delta > 0$ ensures exponential discounting.) Since $\mathcal{F} = u/(2\delta)$, dividing through by 2δ : $\frac{1}{2}\sigma_\omega^2 \mathcal{F}'' + \mu_\omega \mathcal{F}' - \delta\mathcal{F} + \Theta^2/(2\delta) = 0$. Substituting $\mu_\omega = \sigma_\omega^2/\omega$ (equation (B.22)) yields the stated ODE.

(iii) At $\omega = \underline{\omega}$ (absorbing), $\omega_t = \underline{\omega}$ for all $t \geq 0$, so $\Theta_t = \Theta(\underline{\omega}) = \sigma$ (since $\sigma_\omega(\underline{\omega}) = 0$ gives $\Theta = \sigma$ by Theorem 2.5). Therefore $\mathcal{F}(\underline{\omega}) = (1/(2\delta)) \int_0^\infty e^{-\delta t} \sigma^2 dt = (\sigma^2/(2\delta)) \cdot (1/\delta) = \sigma^2/(2\delta^2)$. Equivalently, at the absorbing state the Feynman–Kac equation (ii) reduces to $-\delta\mathcal{F} + \sigma^2/(2\delta) = 0$, giving $\mathcal{F} = \sigma^2/(2\delta^2)$. The same argument applies at $\omega = 1/\epsilon$.

(iv) In the recovery regime $(\underline{\omega}, 1/\epsilon)$, $\sigma_\omega > 0$ (Proposition BO.1), so $\Theta(\omega) = \sigma + \sigma_\omega/\omega + \sigma_\omega/[(1+\phi)(1-\omega)] > \sigma$ by Theorem 2.5. Paths from $\omega_0 \in (\underline{\omega}, 1/\epsilon)$ spend positive time in the interior before converging to $1/\epsilon$, during which $\Theta_t > \sigma$. Therefore $\mathbb{E}_{\omega_0} \int_0^\infty e^{-\delta t} \Theta_t^2 dt > \mathbb{E}_{\omega_0} \int_0^\infty e^{-\delta t} \sigma^2 dt = \sigma^2/\delta$, giving $\mathcal{F}(\omega_0) > \sigma^2/(2\delta^2)$. $\square$

BO.7 Employment Multiplier and Welfare Analysis

BO.7.1 Employment Multiplier of Recapitalisation

Corollary BO.1 (Employment multiplier of recapitalisation) Let $\omega_0 < \underline{\omega}$ and suppose recapitalisation raises ω to $\omega_{0+} = \omega_0 + \Delta > \underline{\omega}$.

(i) **Immediate gain.** Employment jumps from $N(\omega_0)$ to $N(\omega_0 + \Delta)$. The percentage gain is

$$\frac{N(\omega_0 + \Delta) - N(\omega_0)}{N(\omega_0)} = \left(\frac{1 - \omega_0}{1 - \omega_0 - \Delta} \right)^{\frac{1}{1+\varphi}} - 1. \quad (\text{BO.17})$$

(ii) **Long-run gain.** As $\omega_t \rightarrow 1/\epsilon$, employment converges to the efficient level $N(1/\epsilon) = 1$. The permanent percentage employment gain relative to the bad steady state is

$$\frac{N(1/\epsilon) - N(\underline{\omega})}{N(\underline{\omega})} = \left(\frac{1 - \underline{\omega}}{1 - 1/\epsilon} \right)^{\frac{1}{1+\varphi}} - 1. \quad (\text{BO.18})$$

(iii) **Convexity.** The employment function $N(\omega)$ is strictly convex in ω : the output gap $1 - N(\omega)/N(1/\epsilon)$ grows at an increasing rate as intermediary capital ω falls. Consequently, the cumulative output loss from remaining in the crisis regime is convex in crisis severity, making the long-run benefit of recapitalisation larger for deeper crises.

Proof. Parts (i) and (ii) follow from the employment formula (2.37). For part (iii), $N(\omega) = [(1 - 1/\epsilon)/(1 - \omega)]^{1/(1+\varphi)}$. Differentiating: $N'(\omega) = [1/(1 + \varphi)](1 - 1/\epsilon)^{1/(1+\varphi)}(1 - \omega)^{-(2+\varphi)/(1+\varphi)} > 0$ and $N''(\omega) = [(2 + \varphi)/((1 + \varphi)^2)](1 - 1/\epsilon)^{1/(1+\varphi)}(1 - \omega)^{-(3+2\varphi)/(1+\varphi)} > 0$. Since N is convex, the output gap $1 - N(\omega)$ is concave, meaning it accelerates as ω falls: each unit decline in intermediary capital produces a larger employment loss than the previous one. $\square$

BO.7.2 Welfare Analysis

The household's utility depends on consumption $C_{h,t} = (1 - \omega_t)Y_t$ and labour supply N_t . At the bad steady state ($\omega = \underline{\omega}$), the household has a higher consumption share ($1 - \underline{\omega} > 1 - 1/\epsilon$) but employment is depressed. The household therefore weakly prefers the bad steady state: the increase in consumption share more than compensates for the lower output, because utility is logarithmic and the consumption share effect dominates.

Recapitalisation is therefore not a Pareto improvement: it transfers resources from the household to the financier. However, the employment gains are convex in crisis severity (Corollary BO.1), while the moral hazard cost is bounded by the leverage constraint (Proposition BO.5): the financier cannot divert more than θ per unit of assets under intermediation. Under the Table 2.1 calibration, the worst-case steady-state comparison in (2.42) gives a ratio of about 0.8: a transfer from $\omega_0 = 0$ to $\underline{\omega}$ costs approximately 2.6% of total wealth and raises long-run output by about 2.1% of efficient output. This is a long-run comparison of steady states, not a discounted

present-value welfare calculation; the policy case for recapitalisation comes from avoiding persistent output losses, especially because delay leaves the economy on the steep portion of the employment curve.

BO.8 Proofs of Policy Results

Proof of Proposition 2.6. At $\omega = \underline{\omega}$, $\bar{\phi}(\underline{\omega}) = 1 + \lambda(\underline{\omega} - \underline{\omega}) = 1$. By (2.16), $\alpha(\underline{\omega}) = 1 - \epsilon \cdot 1 \cdot \underline{\omega} = 1 - \epsilon \underline{\omega} > 0$ (since $\underline{\omega} < 1/\epsilon$ by Assumption 2.2). The central bank's equity holdings are $\alpha P_{D,t} = (1 - \epsilon \underline{\omega}) Y_t / (\epsilon \delta)$. Since $\omega = \underline{\omega}$ is absorbing (Theorem 2.3), these holdings are permanent. For the equity premium: at the absorbing state, $\sigma_\omega(\underline{\omega}) = 0$ (Proposition BO.1), so Theorem 2.5 gives $\sigma_R = \sigma$ and $\Theta = \sigma$. The equity excess return is $\mu_R - r = \sigma_R \Theta = \sigma^2$. $\square$

Proof of Proposition 2.7. Define $\Omega(\omega) \equiv \omega P_Y(\omega) = \omega AN(\omega)/\delta$, the financier's dollar wealth as a function of the state. Since $N > 0$ and $N' > 0$, we have $\Omega'(\omega) = (A/\delta)[N(\omega) + \omega N'(\omega)] > 0$, so Ω is strictly increasing and continuous with $\Omega(0) = 0$ and $\Omega(\omega) \rightarrow \infty$ as $\omega \rightarrow 1$. A fiscal transfer of $T > 0$ dollars raises the financier's wealth from $\Omega(\omega_0)$ to $\Omega(\omega_0) + T$. Because total wealth adjusts to $P_Y(\omega_{0+}) = AN(\omega_{0+})/\delta$, the new wealth share ω_{0+} is the unique solution to the fixed-point equation

$$\omega_{0+} N(\omega_{0+}) = \omega_0 N(\omega_0) + \frac{\delta T}{A},$$

i.e. $\omega_{0+} = \Omega^{-1}(\Omega(\omega_0) + T)$. For any target $\omega^* \in (\omega_0, 1/\epsilon)$, the transfer $T^* = \Omega(\omega^*) - \Omega(\omega_0) > 0$ achieves $\omega_{0+} = \omega^*$. Choosing $\omega^* > \underline{\omega}$ places the economy in the recovery regime $(\underline{\omega}, 1/\epsilon)$. By Theorem 2.4, $\omega_t \rightarrow 1/\epsilon$ a.s. from any initial condition in the recovery regime. The convergence requires no further intervention because the recovery regime is invariant (Proposition BO.2) and the drift $\mu_\omega = \sigma_\omega^2/\omega > 0$ pushes ω toward $1/\epsilon$. $\square$

BO.9 Translation of Hugonnier's Results

This appendix states the results of Hugonnier (2012) that we use, first in his original notation, and then translates them into the notation of our model.

Hugonnier's notation

Hugonnier (2012) considers a two-agent economy with agents labelled 1 (unconstrained) and 2 (portfolio-constrained). The agents have utility functions u_1 and u_2 defined over consumption.

The weighting process λ_t is the ratio of their marginal utilities:

$$\lambda_t = \frac{u'_2(c_{2,t})}{u'_1(c_{1,t})}, \quad (\text{BO.19})$$

where $c_{1,t}$ and $c_{2,t}$ are the agents' equilibrium consumption rates.

Hugonnier's Theorem 1. In the economy with a portfolio constraint on agent 2, equilibrium asset prices contain rational bubbles if and only if λ_t is a strict local martingale (equivalently, $\mathbb{E}_t[\lambda_T] < \lambda_t$ for some $T > t$).

Hugonnier's Proposition 6. In a finite-horizon economy $[0, T]$ with log utility, discount rate ρ , and weighting process volatility $\|v_\lambda\| = v$, the gap-to-price ratio (which Hugonnier calls the bubble-to-price ratio) is:

$$b(t, s_t) = \frac{1}{\rho\eta(t)} H(T - t, s_t; a_1) + \frac{\eta'(t)}{\rho\eta(t)} H(T - t, s_t; 1), \quad (\text{BO.20})$$

where $\eta(t) = (1/\rho)(1 - e^{-\rho(T-t)})$, $s_t = \lambda_t/(1 + \lambda_t)$ is agent 2's consumption share, $a_1 = \sqrt{1 + 8\rho/v^2}$, and $H(\tau, s; a) = s^{(1+a)/2}\Phi(d_+) + s^{(1-a)/2}\Phi(d_-)$ with $d_\pm = (\ln s)/(v\sqrt{\tau}) \pm (a/2)v\sqrt{\tau}$.

Translation to our notation

The correspondence between Hugonnier's notation and ours is:

Hugonnier	This paper
Agent 1 (unconstrained)	Financier (i)
Agent 2 (constrained)	Household (h)
$c_{1,t}$	$C_{i,t} = \omega_t Y_t$
$c_{2,t}$	$C_{h,t} = (1 - \omega_t) Y_t$
$u'_j(c) = 1/c$ (log utility)	$1/C_{j,t}$
$\lambda_t = u'_2(c_2)/u'_1(c_1)$	$C_{i,t}/C_{h,t} = \omega_t/(1 - \omega_t)$
$s_t = \lambda_t/(1 + \lambda_t)$	$(1 - \omega_t)$
ρ	δ

Important note on the labelling convention. In Hugonnier's framework, agent 2 is portfolio-constrained. In our model, the household (agent h) faces the participation constraint (she cannot

hold equity), so the household corresponds to Hugonnier's agent 2. Hugonnier's weighting process $\lambda_t = u'_2(c_2)/u'_1(c_1) = C_{i,t}/C_{h,t} = \omega_t/(1 - \omega_t)$ is then related to our $Q_t = 1/\omega_t$ by

$$\lambda_t = \frac{\omega_t}{1 - \omega_t} = \frac{1}{Q_t - 1}.$$

The constrained agent's consumption share is $s_t = \lambda_t/(1 + \lambda_t) = 1 - \omega_t$, which is the household's wealth share. In Section 2.5.3, we instead use the equivalent formulation $\lambda_t = C_{h,t}/C_{i,t} = (1 - \omega_t)/\omega_t = Q_t - 1$, which is a monotone transformation of Hugonnier's λ and inherits the same strict/true local martingale classification.

Translation of Theorem 1. The intermediation gap is positive if and only if $Q_t - 1$ (equivalently, $Q_t = 1/\omega_t$) is a strict local martingale. This is the criterion (2.32) stated in Section 2.5.3.

Translation of Proposition 6. Taking $T \rightarrow \infty$: $\eta(t) \rightarrow 1/\delta$, $\eta'(t) \rightarrow 0$, $H(\infty, s; a_1) \rightarrow s^{(1+a_1)/2}$ for $s \in (0, 1)$. Substituting $s = 1 - \omega$ and $\gamma = (1 + a_1)/2$ with $a_1 = \sqrt{1 + 8\delta/v^2}$ yields the intermediation gap formula (BO.13):

$$\frac{B_t}{P_{D,t}} = (1 - \omega_t)^\gamma, \quad \gamma = \frac{1 + \sqrt{1 + 8\delta/v^2}}{2}.$$

The full derivation is in Appendix BO.5.

BO.10 Partial Participation and CRRA Robustness

This appendix extends the model to address robustness along two dimensions: partial household equity participation and CRRA preferences.

BO.10.1 Extended Model

We allow the household to invest a fixed fraction $\alpha_h \in [0, 1)$ of her financial wealth $H_{h,t}$ in the equity market. When $\alpha_h = 0$, we recover the main paper.

All identities from the main paper that depend on log utility and goods market clearing continue to hold:

$$C_{i,t} = \omega_t Y_t, \quad C_{h,t} = (1 - \omega_t) Y_t, \quad (\text{BO.21})$$

$$H_{i,t} = \omega_t P_{Y,t}, \quad H_{h,t} = (1/\epsilon - \omega_t) P_{Y,t}. \quad (\text{BO.22})$$

Equity market clearing with household participation requires:

$$\phi(\omega) = \alpha_h + \frac{1 - \alpha_h}{\epsilon \omega}. \quad (\text{BO.23})$$

BO.10.2 Robustness of Equilibrium Failure

Theorem BO.2 (Robustness of equilibrium failure) *For any $\alpha_h \in [0, 1)$, no competitive equilibrium exists when $\omega < \omega_c(\alpha_h)$ in the absence of central bank intervention, where $\omega_c(\alpha_h) > 0$ is the unique solution to*

$$\alpha_h + \frac{1 - \alpha_h}{\epsilon \omega_c} = 1 + \lambda(\omega_c - \underline{\omega}). \quad (\text{BO.24})$$

Proof. The required portfolio share $\phi(\omega) = \alpha_h + (1 - \alpha_h)/(\epsilon \omega)$ is strictly decreasing and $\phi(\omega) \rightarrow \infty$ as $\omega \rightarrow 0^+$ for any $\alpha_h < 1$. The leverage constraint $\phi \leq \bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega})$ is bounded. There exists a unique $\omega_c(\alpha_h) > 0$ at which $\phi(\omega_c) = \bar{\phi}(\omega_c)$. For $\omega < \omega_c(\alpha_h)$, the argument of Theorem 2.1 applies verbatim. $\square$

Proposition BO.7 (Comparative statics of the crisis threshold) $\omega_c(\alpha_h)$ is strictly decreasing in α_h :

$$\frac{d\omega_c}{d\alpha_h} = \frac{1 - 1/(\epsilon \omega_c)}{\lambda + (1 - \alpha_h)/(\epsilon \omega_c^2)} < 0. \quad (\text{BO.25})$$

The crisis threshold satisfies $\omega_c(0) = \omega_c$ (the main paper's threshold) and $\omega_c(\alpha_h) \rightarrow \underline{\omega}$ as $\alpha_h \rightarrow 1$: even with full household equity participation, the crisis threshold cannot fall below the minimum capital requirement.

Proof. Implicit differentiation of (BO.24). Since $\omega_c < 1/\epsilon$, the numerator $1 - 1/(\epsilon \omega_c) < 0$ and the denominator is positive. $\square$

Quantitative illustration.

α_h	$\omega_c(\alpha_h)$	CB equity at $\underline{\omega}$	Interpretation
0.00	0.0606	64%	Main paper
0.10	0.0598	58%	Modest participation
0.30	0.0577	45%	Substantial participation
0.50	0.0547	32%	Equal risk sharing
0.80	0.0461	13%	Mostly direct
0.95	0.0353	3%	Near-full participation

BO.10.3 Dynamics and Preservation of Results

Proposition BO.8 (Dynamics with partial participation) *The financier's wealth share evolves as $d\omega_t = (\sigma_\omega^2/\omega) dt + \sigma_\omega dZ_t$ with:*

$$\sigma_\omega(\omega; \alpha_h) = \begin{cases} \frac{\sigma(1+\varphi)\lambda(1-\omega)\omega(\omega-\underline{\omega})}{(1+\varphi)(1-\omega)-\lambda\omega(\omega-\underline{\omega})} & \text{if } \omega < \omega_c(\alpha_h), \\ \frac{\sigma(1+\varphi)(1-\alpha_h)(1-\omega)(1/\epsilon-\omega)}{(1+\varphi)(1-\omega)-(1-\alpha_h)(1/\epsilon-\omega)} & \text{if } \omega \geq \omega_c(\alpha_h). \end{cases} \quad (\text{BO.26})$$

The constrained-region volatility is independent of α_h .

Corollary BO.2 (Preservation of main results) *For any $\alpha_h \in [0, 1)$: (i) $Q_t = 1/\omega_t$ is a non-negative local martingale; (ii) σ_ω has exactly two interior zeros at $\underline{\omega}$ and $1/\epsilon$; (iii) the state space partitions into three invariant regimes; (iv) the Permanent QE and QE Trap mechanisms survive.*

BO.10.4 Bubble Classification

Theorem BO.3 (Bubble classification with partial participation) *For any $\alpha_h \in [0, 1)$: (i) in the baseline model (no leverage constraint), $Q = 1/\omega$ is a strict local martingale and rational bubbles exist; (ii) in the constrained model, Q is a true martingale and no bubbles exist.*

Proof. The bubble classification depends on σ_ω near $\omega = 0$. In the baseline: $\sigma_\omega \rightarrow (1-\alpha_h)c_0 > 0$ for $\alpha_h < 1$, giving a convergent integral (strict local martingale). With the constraint: $|\sigma_\omega| \sim c'_0\omega$, giving a divergent integral (true martingale). $\square$

BO.10.5 CRRA / Power Utility

We now replace the household's log utility with CRRA (power) utility, parameterised by the elasticity of intertemporal substitution $\psi > 0$ (with $\psi = 1$ recovering the main paper).

Theorem BO.4 (Equilibrium failure under power utility)

- (i) **Existence region (identical to log utility):** A competitive equilibrium with household EIS ψ exists if and only if $\omega \geq \omega_c$, where ω_c solves $\bar{\phi}(\omega_c) \omega_c = 1/\epsilon$ and is independent of ψ .
- (ii) **Failure region (identical to log utility):** For $\omega < \omega_c$, the financier at maximum leverage $\bar{\phi}(\omega)$ cannot absorb all equity: $\bar{\phi}(\omega) \omega < 1/\epsilon$. No interest rate clears the bond market. This is a structural failure that depends on $(\epsilon, \lambda, \underline{\omega})$ but not on ψ .
- (iii) **Equilibrium characterisation for $\omega \geq \omega_c$:** Conditional on existence, the equilibrium quantities depend on ψ through the household's consumption-wealth ratio $c_h(\omega; \psi)$, which solves a second-order ODE on $[\omega_c, 1/\epsilon]$ with boundary condition $c_h(1/\epsilon) = \delta$.

Proof. Parts (i) and (ii) use only the leverage constraint and limited participation. The required portfolio share is $\phi(\omega) = 1/(\epsilon\omega)$, which exceeds $\bar{\phi}(\omega)$ for $\omega < \omega_c$. Since the household cannot hold equity and her bond demand is a function of ω (not of the interest rate for $\psi = 1$, or with sensitivity of order $|\psi - 1|$ for $\psi \neq 1$), no price adjustment can close the gap. The market-clearing identity $\bar{\phi}(\omega_c) \omega_c = 1/\epsilon$ involves only $(\epsilon, \lambda, \underline{\omega})$; household preferences do not enter.

Part (iii): in the existence region $\omega \geq \omega_c$, the household's HJB equation under CRRA yields a consumption-wealth ratio $c_h(\omega; \psi)$ that replaces the log-utility constant $c_h = \delta$. The ω dynamics acquire an additional drift term $(c_h - \delta)(1 - \omega)\omega$, and the equilibrium is uniquely determined for ψ not too far below unity. $\square$

Proposition BO.9 (Three regimes by EIS)

$\psi = 1$ (log utility):

$c_h = \delta$, $Q = 1/\omega$ is a local martingale, $\mu_\omega = \sigma_\omega^2/\omega$. Both $\underline{\omega}$ and $1/\epsilon$ are steady states. This is the main paper's Permanent QE zone (below $\underline{\omega}$) and QE Trap (above $\underline{\omega}$ but near it).

$\psi > 1$ (high substitutability):

$c_h(\underline{\omega}) > \delta$, so the drift correction $(c_h - \delta)(1 - \omega)\omega > 0$ pushes ω away from $\underline{\omega}$. The Permanent QE and QE Trap mechanisms are weakened: the bad steady state may become reflecting or repelling rather than absorbing, making fiscal recapitalisation less urgent.

$\psi < 1$ (**low substitutability**):

$c_h(\underline{\omega}) < \delta$, so the drift correction is negative: ω drifts toward $\underline{\omega}$. The Permanent QE and QE Trap mechanisms are deepened: both the reduced diffusion (smaller σ_ω) and the negative drift (deterministic pull toward the boundary) strengthen the trap and increase the required fiscal transfer.

Proof. Each case follows from the sign of $c_h(\omega; \psi) - \delta$ at the lower boundary. Under power utility, $c_h(\omega; \psi)$ is decreasing in ψ (the precautionary-savings effect): a household with $\psi < 1$ consumes less out of wealth (“saves more”), depressing c_h below δ . The boundary drift $\mu_\omega(\underline{\omega}) = (c_h(\underline{\omega}) - \delta)(1 - \underline{\omega})\underline{\omega}$ inherits the sign of $c_h - \delta$, while $\sigma_\omega(\underline{\omega}) = 0$ continues to hold. $\square$

Remark BO.3 (Key message) *The equilibrium failure is structurally robust to power utility. The crisis threshold ω_c and the failure mechanism (insufficient risk-bearing capacity) are ψ -independent. What changes under power utility is the equilibrium dynamics in the existence region: the probability of reaching the failure boundary from any initial $\omega > \omega_c$, the speed of convergence to the steady states, and the asset-pricing implications all depend on ψ . The existence of the failure and the trap is structural; the severity of the trap depends on preferences.*

BO.11 Endogenous Non-Participation via Ambiguity

This appendix provides a microfoundation for limited stock market participation via Knightian uncertainty.

BO.11.1 Household Preferences under Ambiguity

The household observes equity returns but is uncertain about the drift. She considers alternative probability measures $\{\mathbb{P}^\nu : |\nu| \leq \kappa\}$ and evaluates portfolios under worst-case beliefs (maxmin expected utility with rectangular priors, following [Chen and Epstein \(2002\)](#)). The parameter κ measures ambiguity and, in the spirit of [Bernanke \(1983\)](#), proxies the informational gap between households and intermediaries: the screening, monitoring, and relationship capital that financial intermediaries accumulate but that households lack.

Proposition BO.10 (Household optimal portfolio) *The household's optimal equity share, as a fraction of financial wealth H_h , subject to $\alpha_h \in [0, 1]$, is:*

$$\alpha_h^*(\omega) = \max \left\{ 0, \min \left\{ 1, \frac{(\Theta - \kappa)(1 - \omega) - \sigma_R(1 - 1/\epsilon)}{\sigma_R(1/\epsilon - \omega)} \right\} \right\}. \quad (\text{BO.27})$$

Proof. Under worst-case beliefs $\mathbb{P}^\nu$ with $|\nu| \leq \kappa$, the equity return drift shifts from μ_R to $\mu_R - \nu\sigma_R$. The worst case for the household is $\nu = \kappa$ (maximally pessimistic):

$$\mu_R^{\text{wc}} - r = \mu_R - r - \kappa\sigma_R = (\Theta - \kappa)\sigma_R.$$

The household maximises log utility over total wealth $W_h = H_h + P_L$, where $H_h = (1/\epsilon - \omega)P_Y$ is financial wealth and $P_L = (1 - 1/\epsilon)P_Y$ is human capital. Her total equity exposure combines the financial position $\alpha_h H_h$ and the implicit exposure from human capital. Since $P_L = (1 - 1/\epsilon)P_Y$ and $P_Y = Y/\delta$ has volatility $\sigma_Y = \sigma_R$, human capital has “equity beta” σ_R . The total equity exposure is $\alpha_h H_h + P_L = \alpha_h(1/\epsilon - \omega)P_Y + (1 - 1/\epsilon)P_Y$.

Under log utility, the optimal portfolio equates the Sharpe ratio to total exposure times return volatility divided by total wealth:

$$\frac{\Theta - \kappa}{\sigma_R} = \frac{[\alpha_h(1/\epsilon - \omega) + (1 - 1/\epsilon)]\sigma_R}{(1 - \omega)},$$

where the denominator $(1 - \omega)P_Y = W_h$ is the household's total wealth (financial plus human capital, scaled by P_Y). Solving for α_h :

$$\alpha_h = \frac{(\Theta - \kappa)(1 - \omega)/\sigma_R^2 - (1 - 1/\epsilon)}{1/\epsilon - \omega} = \frac{(\Theta - \kappa)(1 - \omega) - \sigma_R(1 - 1/\epsilon)}{\sigma_R(1/\epsilon - \omega)},$$

where we used $(\Theta - \kappa)/\sigma_R = (\Theta - \kappa)/\sigma_R$ and multiplied through by σ_R . Applying the constraint $\alpha_h \in [0, 1]$ gives (BO.27). $\square$

The optimal portfolio has two opposing components. The speculative demand $(\Theta - \kappa)(1 - \omega)/[\sigma_R(1/\epsilon - \omega)]$ reflects the ambiguity-adjusted risk premium. The hedging demand $-(1 - 1/\epsilon)/(1/\epsilon - \omega) < 0$ is *negative*: the household's human capital $(1 - 1/\epsilon)P_Y$ gives her substantial implicit equity exposure, so she reduces financial equity to manage total risk.

BO.11.2 Robust Equilibrium Failure

Theorem BO.5 (Robust equilibrium failure) For $\omega < \underline{\omega}$, no competitive equilibrium exists for any $\alpha_h \in [0, 1]$, regardless of κ .

Proof. When $\omega < \underline{\omega}$: $\bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega}) < 1$. The total private equity demand is $\bar{\phi}(\omega)\omega P_Y + \alpha_h(1/\epsilon - \omega)P_Y$. Even at $\alpha_h = 1$, this equals $[\bar{\phi}(\omega)\omega + 1/\epsilon - \omega]P_Y < P_Y/\epsilon = P_D$ because $\bar{\phi}(\omega) < 1$ forces $\bar{\phi}(\omega)\omega < \omega$. The gap is structural: $P_D - \max \text{ private demand} = -\lambda\omega(\omega - \underline{\omega})P_Y > 0$. $\square$

Corollary BO.3 (Conditional failure) For $\omega \in [\underline{\omega}, \omega_c)$, equilibrium is restored iff the household participates with $\alpha_h \geq \alpha_h^{\min}(\omega) = 1 - \lambda\omega(\omega - \underline{\omega})/(1/\epsilon - \omega)$.

BO.11.3 Endogenous Non-Participation

Define the *participation threshold function*:

$$\kappa^*(\omega) \equiv \Theta(\omega) - \frac{\sigma_R(\omega)(1 - 1/\epsilon)}{1 - \omega}. \quad (\text{BO.28})$$

The household holds zero equity iff $\kappa \geq \kappa^*(\omega)$. In the unconstrained region:

$$\kappa^*(\omega) = \sigma_R(\omega) \cdot \frac{1 - \epsilon\omega}{\epsilon\omega(1 - \omega)}. \quad (\text{BO.29})$$

Key properties: $\kappa^*(1/\epsilon) = 0$, $\kappa^*(0^+) = +\infty$, and κ^* is strictly decreasing.

Theorem BO.6 (Endogenous non-participation) For $\kappa > 0$, let $\omega_p(\kappa) \in (0, 1/\epsilon)$ solve $\kappa^*(\omega_p) = \kappa$. Then:

- (i) For $\omega \in (\omega_p(\kappa), 1/\epsilon)$: the household holds zero equity.
- (ii) For $\omega \in (0, \omega_p(\kappa))$: the household's optimal equity share is positive.
- (iii) $\omega_p(\kappa)$ is strictly decreasing in κ , with $\omega_p(0) = 1/\epsilon$ and $\omega_p(\kappa) \rightarrow 0$ as $\kappa \rightarrow \infty$.

Quantitative thresholds. With Table 2.1 parameters ($\epsilon = 13.7741$, $\underline{\omega} = 0.0260$, $\omega_c = 0.0606$, $\sigma = 0.1505$, $\varphi = 1.3333$):

- At $\omega_c = 0.0606$: $\kappa^*(\omega_c) \approx 0.032$. Complete non-participation everywhere requires $\kappa \geq 0.032$ (3.2% of the Sharpe ratio).

- At $\underline{\omega} = 0.0260$: $\kappa^*(\underline{\omega}) \approx 0.007$. Non-participation at the bad steady state requires $\kappa \geq 0.007$ (0.7% of the Sharpe ratio).

Corollary BO.4 (Recovery of the main paper) For $\kappa \geq \kappa^*(\omega_c) \approx 0.032$, $\alpha_h^*(\omega) = 0$ for all ω : all results of the main paper hold exactly.

BO.11.4 Permanent QE, the QE Trap, and Bubble Classification

Proposition BO.11 (Preservation) For any $\kappa \geq 0$: (i) $Q_t = 1/\omega_t$ is a non-negative local martingale; (ii) $\underline{\omega}$ and $1/\epsilon$ remain the only absorbing steady states; (iii) the Permanent QE and QE Trap mechanisms survive; (iv) in the baseline, Q is a strict local martingale (bubbles exist); in the constrained model, Q is a true martingale (no bubbles).

Proof. The Q -martingale property depends on $\mu_\omega = \sigma_\omega^2/\omega$, which holds for any ϕ . In the constrained region, $\phi = \bar{\phi}(\omega)$ regardless of α_h , so σ_ω is independent of κ . The zeros of σ_ω and the Feller boundary classification are unchanged. $\square$

BO.11.5 Three-Region Equilibrium

For $\kappa \in (0, \kappa^*(\omega_c))$, the economy has three regions: (i) robust failure ($\omega < \underline{\omega}$); (ii) endogenous participation ($\underline{\omega} \leq \omega < \omega_p(\kappa)$); (iii) non-participation ($\omega_p(\kappa) \leq \omega < 1/\epsilon$, identical to the main paper). The effective crisis threshold $\omega_c(\kappa)$ is increasing in κ , with $\omega_c(0) = \underline{\omega}$ and $\omega_c(\kappa) = \omega_c$ (main paper) for $\kappa \geq \kappa^*(\omega_c)$.

The ambiguity parameter κ nests all models:

κ	Household behavior	Model recovered
0	Full participation	Online Appendix BO.10 with endogenous $\alpha_h(\omega)$
$(0, 0.032)$	Participates in some states	Three-region model (this appendix)
$[0.032, \infty)$	Never participates	Main paper

BO.12 Robustness under Epstein–Zin Preferences

This appendix shows that the equilibrium failure result is robust to replacing the household's log utility with Epstein–Zin (EZ) stochastic differential utility (Duffie and Epstein, 1992). The financier retains log utility (which delivers the Q -martingale property and the market price of risk $\Theta = \phi\sigma_R$), while the household has EIS ψ and risk aversion γ , with $\psi \neq 1/\gamma$ in general. Log utility is the special case $\psi = \gamma = 1$; CRRA is $\gamma = 1/\psi$.

BO.12.1 Equilibrium Failure Is Preference-Independent

Theorem BO.7 (Equilibrium failure under EZ) *The crisis threshold ω_c is preference-independent: it depends on $(\epsilon, \lambda, \underline{\omega})$ but not on the household's EIS ψ or risk aversion γ . For $\omega < \omega_c$, no competitive equilibrium exists regardless of the household's preference specification.*

Proof. The failure condition $\bar{\phi}(\omega)\omega < 1/\epsilon$ involves three objects:

1. The leverage constraint $\bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega})$, which depends on $(\lambda, \underline{\omega})$.
2. The equity-to-wealth ratio $P_D/P_Y = 1/\epsilon$, which is a structural consequence of the Dixit–Stiglitz production structure.
3. The equity market clearing condition $\phi\omega = P_D/P_Y = 1/\epsilon$.

The ratio $P_D/P_Y = 1/\epsilon$ requires verification. Under EZ, total wealth is $P_{Y,t} = Y_t/D(\omega_t)$ where $D(\omega) = c_h(\omega)(1 - \omega) + \delta\omega$ is the aggregate consumption–wealth ratio (which depends on (ψ, γ) through c_h). So $P_Y \neq Y/\delta$ in general. However, the equity price $P_{D,t}$ is the present value of dividends $D_s = Y_s/\epsilon$ under the *financier's* SDF $\Lambda_{i,t} = e^{-\delta t}/(\delta H_{i,t})$. Since total wealth P_Y is the present value of Y_s under the same SDF, linearity gives $P_D/P_Y = (1/\epsilon) \cdot E[\int \Lambda_i Y_s ds] / E[\int \Lambda_i Y_s ds] = 1/\epsilon$. This ratio is exact and preference-independent: the household's EZ parameters cancel.

With $P_D/P_Y = 1/\epsilon$, equity clearing requires $\phi = 1/(\epsilon\omega)$. When $\omega < \omega_c$: $1/(\epsilon\omega) > \bar{\phi}(\omega)$ and no competitive equilibrium exists. The household's (ψ, γ) do not enter. $\square$

Interpretation. The crisis threshold ω_c is a *structural* property of the economy's intermediation architecture—the leverage constraint, participation restrictions, and production structure. The vulnerability to market failure is the same whether households have log utility, high risk aversion, low EIS, or any Epstein–Zin specification.

BO.12.2 Permanent QE and the QE Trap Under EZ

Under log utility (main paper, Remark 2.2), the absorbing steady state at $\underline{\omega}$ creates *Permanent QE* on $(0, \underline{\omega})$, and the slow drift just above $\underline{\omega}$ creates the *QE Trap* on $(\underline{\omega}, \omega_T)$. Under EZ, both mechanisms are modified:

Proposition BO.12 (Trap mechanism under EZ)

- (i) For $\psi < 1$ and any $\gamma > 0$: $c_h(\underline{\omega}; \psi, \gamma) < \delta$.

(ii) For $\psi < 1$ and $\gamma > 1/\psi$ (EZ with high risk aversion): $c_h(\underline{\omega}; \psi, \gamma) < c_h^{\text{CRRA}}(\underline{\omega}; \psi) < \delta$. The trap is strengthened relative to CRRA.

(iii) The bad steady state $\underline{\omega}$ is no longer absorbing: $\mu_\omega(\underline{\omega}) = (c_h(\underline{\omega}) - \delta)(1 - \underline{\omega})\underline{\omega} < 0$. The economy drifts below $\underline{\omega}$ —the trap is a basin of attraction with a stationary distribution, not a permanent absorbing state.

Proof. At $\omega = \underline{\omega}$, $\sigma_\omega(\underline{\omega}) = 0$ (Proposition BO.1). The ω -dependent component of the household's continuation-value volatility σ_V therefore vanishes. However, σ_V does *not* vanish entirely: the household retains direct exposure to productivity risk through A_t , giving $\sigma_V^2(\underline{\omega}) = (\alpha\sigma)^2 > 0$ where $\alpha = \psi(1 + \varphi)/(\psi\varphi + 1)$.

Under CRRA ($\gamma = 1/\psi$), the precautionary channel is inactive and the standard Merton consumption–wealth formula gives $c_h^{\text{CRRA}} < \delta$ for $\psi < 1$.

Under EZ with $\gamma > 1/\psi$, the precautionary channel depresses c_h further. The EZ Euler equation acquires a term proportional to $-\frac{1}{2}(\gamma - 1/\psi)\psi\sigma_V^2 < 0$, which reduces the consumption–wealth ratio below the CRRA value. Since $\sigma_V^2(\underline{\omega}) > 0$, this correction is active at the boundary, giving $c_h^{\text{EZ}} < c_h^{\text{CRRA}} < \delta$.

Part (iii) follows from the drift formula $\mu_\omega = \sigma_\omega^2/\omega + (c_h - \delta)(1 - \omega)\omega$. At $\underline{\omega}$: $\sigma_\omega = 0$ and $c_h < \delta$, so $\mu_\omega = (c_h - \delta)(1 - \underline{\omega})\underline{\omega} < 0$. $\square$

Comparison with the main paper. Under log utility ($\psi = \gamma = 1$), $c_h = \delta$, so $\mu_\omega(\underline{\omega}) = 0$ and $\underline{\omega}$ is absorbing: *Permanent QE* obtains on $(0, \underline{\omega})$ and the *QE Trap* on $(\underline{\omega}, \omega_T)$ (Remark 2.2). Under EZ with $\psi < 1$, $\underline{\omega}$ is no longer absorbing — the strict Permanent QE regime disappears — but the economy drifts *below* $\underline{\omega}$ into the crisis regime and oscillates with a stationary distribution concentrated near it. The absorbing-boundary phenomenon is replaced by a quasi-stationary near-absorption: the trap is wider and deeper even though it is no longer mathematically permanent. Fiscal recapitalisation must push ω above a drift-reversal point $\omega^+ > \underline{\omega}$ (where $\mu_\omega(\omega^+) = 0$) rather than merely above $\underline{\omega}$. The required transfer is larger than under log utility.

BO.13 Leverage vs. Equity Constraints

He et al. (2017, Section 4.3) distinguish two classes of intermediary constraint: *leverage constraints* (Adrian and Shin 2014; Brunnermeier and Pedersen 2009), which directly cap portfolio leverage and produce procyclical deleveraging, and *equity constraints* (He and Krishnamurthy

2012, 2013), which require the intermediary to maintain a minimum equity stake and produce countercyclical leverage. Our baseline specification (2.7) is a leverage constraint: the cap $\bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega})$ tightens as the financier's wealth share ω falls, and at $\omega = \underline{\omega}$ forces the financier to be exactly unlevered ($\bar{\phi} = 1$).

This appendix asks: which of our results are robust to replacing the leverage constraint with an equity constraint? We show that equilibrium failure, QE necessity, the elimination of the intermediation gap, and the Leverage Paradox all survive. The QE Trap does not. We then characterise, for a general leverage function $\bar{\phi}$, the precise condition under which the trap exists: $\bar{\phi}(\omega^*) = 1$ at some interior point. The trap is a consequence of *procyclical deleveraging to zero leverage*—the defining feature of the leverage-constraint tradition—and cannot arise under an equity constraint.

BO.13.1 Setup: the Equity Constraint

Replace Assumption 2.2 with the following. The financier must fund at least fraction $a \in (0, 1)$ of her equity investment with her own wealth (skin in the game). Her portfolio equity share is therefore bounded above:

$$\phi_t \leq \frac{1}{a} \equiv \bar{\phi}^{EC}. \quad (\text{BO.30})$$

The cap $1/a$ is *constant*, independent of the state variable ω_t . All other model primitives—preferences, technology, limited participation, the financier's budget constraint (2.6)—are unchanged.

Remark BO.4 (Leverage cyclicity) Under the leverage constraint (2.7), constrained leverage $\bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega})$ falls as ω falls (procyclical). Under (BO.30), constrained leverage is fixed at $1/a$ regardless of ω . Yet realised leverage $\phi_t = 1/(\epsilon\omega_t)$ in the unconstrained region is countercyclical in both cases: leverage rises as ω falls because the financier is the sole equity holder and the equity market $P_{D,t} = P_{Y,t}/\epsilon$ does not shrink as fast as financier wealth. The distinction matters only in the constrained region.

BO.13.2 Crisis Threshold and Central Bank Intervention

Proposition BO.13 (Equilibrium failure under equity constraint) Competitive equilibrium without central bank intervention fails to exist for all $\omega < \omega_c^{EC}$, where

$$\omega_c^{EC} \equiv \frac{a}{\epsilon}. \quad (\text{BO.31})$$

Proof. Equity market clearing without the central bank requires $\phi_t = 1/(\epsilon\omega_t)$ (Appendix B.5). The equity constraint (BO.30) binds when $1/(\epsilon\omega) > 1/a$, i.e. when $\omega < a/\epsilon$. In this region, the excess equity supply

$$\left[\frac{1}{\epsilon\omega} - \frac{1}{a} \right] \omega P_{Y,t} = \frac{1}{\epsilon} \left(1 - \frac{\epsilon\omega}{a} \right) P_{Y,t} > 0$$

has no buyer: the household holds only bonds (limited participation) and her bond demand $(1/\epsilon - \omega)P_{Y,t}$ is perfectly interest-rate inelastic under log utility (2.8). No price adjustment restores market clearing. $\square$

The failure result is an instance of the impossibility in Proposition 2.4(iii): any bounded leverage function $\bar{\phi}$ —whether state-dependent or constant—satisfies $\bar{\phi}(\omega) < 1/(\epsilon\omega)$ for ω sufficiently small.

Remark BO.5 (Quantitative severity of the failure region) *Although equilibrium failure survives under the equity constraint, its quantitative extent can be markedly narrower than under our baseline leverage constraint. The failure-region width $\omega_c^{EC} - 0 = a/\epsilon$ scales linearly with a , the minimum equity share the intermediary must hold. Table BO.2 compares the two at the empirical calibration ($\epsilon = 13.7741$).*

Constraint	Threshold	Basin width	Basin as share of $(0, 1/\epsilon)$
Leverage baseline ($\lambda = 5.72$, $\underline{\omega} = 0.0260$)	$\omega_c = 0.0606$	0.0606	83%
Equity, $a = 1$ (no leverage)	$\omega_c^{EC} = 0.0726$	0.0726	100%
Equity, $a = 0.5$ (leverage cap 2×)	$\omega_c^{EC} = 0.0363$	0.0363	50%
Equity, $a = 0.1$ (leverage cap 10×)	$\omega_c^{EC} = 0.0073$	0.0073	10%
Equity, $a = 0.05$ (leverage cap 20×, HK-style)	$\omega_c^{EC} = 0.0036$	0.0036	5%

Table BO.2: Failure-region comparison at the empirical calibration ($\epsilon = 13.7741$, $1/\epsilon = 0.0726$). The leverage-baseline threshold ω_c is from Table 2.1; equity thresholds use $\omega_c^{EC} = a/\epsilon$. The last column expresses the failure region as a fraction of the good-steady-state basin $(0, 1/\epsilon)$.

At equity-constraint calibrations consistent with observed intermediary leverage ratios of 10–20× (He et al., 2017), the failure region shrinks to 5–10% of the good-steady-state basin—roughly an order of magnitude narrower than the 83% implied by our leverage-constraint calibration. QE is needed for a correspondingly smaller range of ω , but the qualitative necessity of central-bank intervention in sufficiently severe crises survives.

Remark BO.6 (Quantitative severity and household participation) *If an equity-constraint model is taken to deliver “no practical need for QE,” the correct reading within our framework is not that failure does not happen—it does, by Proposition BO.13—but that at realistic a the failure region $(0, a/\epsilon)$*

rarely binds in data. Whether it binds depends on the joint distribution of (ω_t, a) and the intensity of limited participation. Online Appendix [BO.10](#) shows that partial household participation with $\alpha_h \geq 1/\epsilon$ would eliminate the failure region entirely; the failure mechanism is therefore a conjunction of bounded intermediary leverage and sufficiently strict limited participation, and weakening either dimension attenuates or eliminates it.

Proposition BO.14 (QE existence under the equity constraint) *If the central bank holds fraction*

$$\alpha_t^{EC} = \max\left(0, 1 - \frac{\epsilon\omega_t}{a}\right) = \begin{cases} 1 - \epsilon\omega_t/a & \text{if } \omega_t < \omega_c^{EC}, \\ 0 & \text{if } \omega_t \geq \omega_c^{EC}, \end{cases} \quad (\text{BO.32})$$

then competitive equilibrium exists for all $\omega \in (0, 1)$, with all real quantities $(P_{Y,t}, C_{h,t}, C_{i,t}, N_t, Y_t)$ identical to the unconstrained model.

Proof. The argument follows the proof of the QE existence result in Appendix [B.5](#). The central bank absorbs the residual equity that the constrained financier cannot hold, financing its purchases by issuing bonds.

Equity market clearing. With α_t^{EC} as defined, the financier holds fraction $(1 - \alpha_t^{EC})$ of the equity market. Her leverage is

$$\phi_t = \frac{(1 - \alpha_t^{EC}) P_{D,t}}{H_{i,t}} = \frac{(\epsilon\omega_t/a) (P_{Y,t}/\epsilon)}{\omega_t P_{Y,t}} = \frac{\omega_t P_{Y,t}/a}{\omega_t P_{Y,t}} = \frac{1}{a} = \bar{\phi}^{EC},$$

where we used $1 - \alpha_t^{EC} = \epsilon\omega_t/a$ (from [\(BO.32\)](#)), $P_{D,t} = P_{Y,t}/\epsilon$, and $H_{i,t} = \omega_t P_{Y,t}$.

Bond market clearing. The household supplies $(1/\epsilon - \omega)P_Y$ in bonds [\(2.8\)](#). The financier borrows $(\phi - 1)\omega P_Y = (1/a - 1)\omega P_Y$ to finance her leveraged equity position. The central bank issues bonds worth $\alpha^{EC} P_D = \alpha^{EC} P_Y/\epsilon$ to finance its equity purchases. Substituting $\alpha^{EC} = 1 - \epsilon\omega/a$:

$$\begin{aligned} (1/a - 1)\omega P_Y + \alpha^{EC} P_Y/\epsilon &= \frac{\omega}{a} P_Y - \omega P_Y + \frac{P_Y}{\epsilon} - \frac{\omega P_Y}{a} \\ &= \frac{P_Y}{\epsilon} - \omega P_Y = \left(\frac{1}{\epsilon} - \omega\right) P_Y. \quad \checkmark \end{aligned}$$

All real quantities are unchanged since $P_{Y,t} = Y_t/\delta$ depends only on log utility ($C_{h,t} + C_{i,t} = Y_t$ and $C_j = \delta H_j$), not on who holds equity. $\square$

BO.13.3 Wealth Dynamics Under the Equity Constraint

Proposition BO.15 (Volatility of ω under the equity constraint) *In the constrained region $\omega \in (0, \omega_c^{EC})$, the financier's wealth-share volatility is*

$$\sigma_\omega^{EC}(\omega) = \frac{\sigma(1+\varphi)(1/a-1)(1-\omega)\omega}{D_{EC}(\omega)}, \quad (\text{BO.33})$$

where

$$D_{EC}(\omega) \equiv (1+\varphi)(1-\omega) - \omega(1/a-1) = (1+\varphi) - (\varphi + 1/a)\omega > 0. \quad (\text{BO.34})$$

In the unconstrained region $\omega \geq \omega_c^{EC}$, the volatility is $\sigma_\omega^u(\omega)$ as in the baseline model. The function σ_ω is continuous at ω_c^{EC} . The drift is $\mu_\omega = \sigma_\omega^2/\omega$ in both regions.

Proof. *Step 1 (Return volatility).* Since $P_{D,t} = Y_t/(\epsilon\delta)$ and $Y_t = A_t N(\omega_t)$, Itô's lemma gives

$$\sigma_R = \sigma_Y = \sigma + \frac{\sigma_\omega}{(1+\varphi)(1-\omega)}.$$

This is identical to the baseline derivation in Appendix B.6 and does not depend on the form of the constraint.

Step 2 (Wealth-share volatility). Itô's quotient rule on $\omega_t = H_{i,t}/P_{Y,t}$ gives (Appendix B.6, Step 2):

$$\sigma_\omega = \omega(\phi - 1)\sigma_R.$$

In the constrained region, $\phi = 1/a$, so $\phi - 1 = (1-a)/a = 1/a - 1 > 0$. Thus

$$\sigma_\omega = \omega(1/a - 1)\sigma_R. \quad (\text{BO.35})$$

Step 3 (Fixed-point equation). Substituting Step 1 into (BO.35):

$$\begin{aligned} \sigma_\omega &= \omega(1/a - 1) \left[\sigma + \frac{\sigma_\omega}{(1+\varphi)(1-\omega)} \right] \\ &= \omega(1/a - 1)\sigma + \frac{\omega(1/a - 1)}{(1+\varphi)(1-\omega)} \sigma_\omega. \end{aligned}$$

Collecting σ_ω terms:

$$\sigma_\omega \left[1 - \frac{\omega(1/a - 1)}{(1+\varphi)(1-\omega)} \right] = \omega(1/a - 1)\sigma$$

$$\sigma_\omega \cdot \frac{(1+\varphi)(1-\omega) - \omega(1/a - 1)}{(1+\varphi)(1-\omega)} = \omega(1/a - 1)\sigma,$$

giving (BO.33) with denominator $D_{EC}(\omega) = (1+\varphi)(1-\omega) - \omega(1/a - 1)$.

Step 4 (Denominator positivity). The denominator is linear in ω :

$$D_{EC}(\omega) = (1+\varphi) - (\varphi + 1/a)\omega.$$

At $\omega = 0$: $D_{EC} = 1 + \varphi > 0$. At $\omega = \omega_c^{EC} = a/\epsilon$:

$$D_{EC}(a/\epsilon) = (1+\varphi)\left(1 - \frac{a}{\epsilon}\right) - \frac{1-a}{\epsilon} = (1+\varphi) - \frac{(1+\varphi)a + (1-a)}{\epsilon} = (1+\varphi) - \frac{1+\varphi a}{\epsilon}.$$

Since $\epsilon > 1$ and $a < 1$: $(1+\varphi a)/\epsilon < (1+\varphi)/\epsilon < 1+\varphi$. Hence $D_{EC} > 0$ on $[0, a/\epsilon]$.

Step 5 (Continuity at ω_c^{EC}). At $\omega = a/\epsilon$:

$$(1/a - 1)\omega_c^{EC} = \frac{1-a}{a} \cdot \frac{a}{\epsilon} = \frac{1-a}{\epsilon} = \frac{1}{\epsilon} - \frac{a}{\epsilon} = \frac{1}{\epsilon} - \omega_c^{EC}.$$

Substituting into (BO.33): the numerator becomes $\sigma(1+\varphi)(1-\omega_c^{EC})(1/\epsilon - \omega_c^{EC})$ and the denominator becomes $(1+\varphi)(1-\omega_c^{EC}) - (1/\epsilon - \omega_c^{EC}) = D_u(\omega_c^{EC})$, matching $\sigma_\omega^u(\omega_c^{EC})$.

Step 6 (Drift). The Q -martingale property (Appendix B.7) gives $\mu_\omega = \sigma_\omega^2/\omega$. The proof uses only log utility ($C_i = \delta H_i$) and is independent of the constraint form. $\square$

BO.13.4 No Permanent QE Under the Equity Constraint

Theorem BO.8 (No Permanent QE) *Under the equity constraint (BO.30):*

- (i) $\sigma_\omega^{EC}(\omega) > 0$ for all $\omega \in (0, \omega_c^{EC})$. There is no interior absorbing state in the constrained region—in particular, no absorbing lower boundary $\underline{\omega}$, so Permanent QE cannot arise.
- (ii) For every initial condition $\omega_0 \in (0, 1/\epsilon)$, $\omega_t \rightarrow 1/\epsilon$ almost surely. There is a single regime $(0, 1/\epsilon)$ with no invariant sub-intervals.
- (iii) QE is self-resolving: starting from any $\omega_0 \in (0, \omega_c^{EC})$, the economy exits the QE region in finite time almost surely:

$$P_{\omega_0}(\tau_{\omega_c^{EC}} < \infty) = 1.$$

Permanent QE (an absorbing lower steady state at a strictly positive $\underline{\omega}$) is absent: there is no interior

absorbing boundary. The QE Trap phenomenon of impractically slow recovery can nevertheless arise near $\omega = 0$, because $\sigma_\omega^{EC}(\omega) \sim (1/a - 1)\sigma\omega$ as $\omega \rightarrow 0$ and hence $\mu_\omega(\omega) = \sigma_\omega^2/\omega \rightarrow 0$; for sufficiently depleted initial capital, the leading-order deterministic drift time $\log(\omega_c^{EC}/\omega_0)/[(1/a - 1)^2\sigma^2]$, with $\omega_c^{EC} = a/\epsilon$, can exceed a policy-relevant horizon. At the Table 2.1 value $\sigma = 0.1505$, the denominator is $(1/a - 1)^2(0.1505)^2$. What the equity constraint removes is the interior absorbing boundary, not the vanishing-drift mechanism; finite-time exit is guaranteed a.s. from any $\omega_0 > 0$, but the expected exit time diverges as $\omega_0 \rightarrow 0$.

Proof. Part (i). From (BO.33), $\sigma_\omega^{EC}(\omega)$ is a ratio of two strictly positive quantities for $\omega \in (0, \omega_c^{EC})$: the numerator $\sigma(1 + \varphi)(1/a - 1)(1 - \omega)\omega > 0$ (since $a < 1$, $\omega > 0$, $\omega < 1$) and the denominator $D_{EC}(\omega) > 0$ (Step 4 above).

Contrast with the baseline leverage constraint. The constrained-region volatility in the baseline model (Appendix B.6) is

$$\sigma_\omega^c(\omega) = \frac{\sigma(1 + \varphi)\lambda(1 - \omega)\omega(\omega - \underline{\omega})}{D_c(\omega)}.$$

The factor $(\omega - \underline{\omega})$ changes sign at $\omega = \underline{\omega}$: negative below (crisis regime), zero at $\underline{\omega}$ (absorbing state), positive above (recovery regime). Under the equity constraint, the corresponding factor is $(1/a - 1) > 0$, a positive constant. The zero at $\underline{\omega}$ that creates the trap has no analogue.

The economics is direct. Under the leverage constraint, $\bar{\phi}(\underline{\omega}) = 1$: the financier is forced to be exactly unlevered. She earns no excess return over the risk-free rate, so her wealth share has zero drift and zero diffusion. She is stuck. Under the equity constraint, $\bar{\phi}^{EC} = 1/a > 1$: the financier remains leveraged at every $\omega > 0$. She always earns the equity premium, her wealth share always has positive drift $\mu_\omega = \sigma_\omega^2/\omega > 0$, and recovery is endogenous.

Part (ii). The boundary classification at $\omega = 0$ is unchanged from the baseline (Appendix B.9). Near $\omega = 0$:

$$\sigma_\omega^{EC}(\omega) \sim c_{EC} \omega, \quad c_{EC} \equiv \frac{\sigma(1 - a)}{a} > 0,$$

so $m(\omega) = \omega^2/(\omega_0^2 \sigma_\omega^2) \rightarrow 1/(\omega_0^2 c_{EC}^2)$ (bounded). Moreover, the Feller N -integral $N(0^+) = \int_0^x s'(x) \int_0^x m(y) dy dx$ diverges because $s'(\omega) = C/\omega^2$. Hence $\omega = 0$ is a natural boundary. The classification at $\omega = 1/\epsilon$ is identical to the baseline (Appendix B.9).

Since $\sigma_\omega > 0$ throughout $(0, 1/\epsilon)$ and the only zeros of σ_ω are at the boundaries $\{0, 1/\epsilon\}$, there are no internal unattainable points. The interval $(0, 1/\epsilon)$ is a single invariant regime (continuous paths cannot cross the unattainable boundaries at 0 and $1/\epsilon$).

Convergence to $1/\epsilon$. The inverse wealth share $Q_t = 1/\omega_t$ is a non-negative local martingale ($\mu_Q = 0$, Appendix B.7), hence a supermartingale by Fatou's lemma. The supermartingale convergence theorem gives $Q_t \rightarrow Q_\infty$ a.s. with $\mathbb{E}[Q_\infty] \leq Q_0 < \infty$.

The limit Q_∞ must satisfy $\sigma_Q(Q_\infty) = 0$. To see this: if $\sigma_Q(Q_\infty) \neq 0$, then the quadratic variation $\langle Q \rangle_\infty = \int_0^\infty \sigma_Q^2(Q_t) dt = +\infty$, and by Dambis–Dubins–Schwarz, Q is a time-changed Brownian motion with unbounded time change, which cannot converge (Appendix B.10). The zeros of σ_Q correspond to the zeros of σ_ω : $\omega \in \{0, 1/\epsilon\}$, i.e., $Q \in \{\epsilon, +\infty\}$.

Since $Q_\infty < \infty$ a.s. (from $\mathbb{E}[Q_\infty] \leq Q_0 < \infty$), we have $Q_\infty = \epsilon$, i.e., $\omega_\infty = 1/\epsilon$.

Part (iii). Since $\omega_c^{EC} = a/\epsilon < 1/\epsilon$ is an interior point of $(0, 1/\epsilon)$, we use the scale function to compute the hitting probability. With $s'(\omega) = C/\omega^2$ (from $2\mu_\omega/\sigma_\omega^2 = 2/\omega$):

$$s(\omega) = -\frac{C}{\omega} + C', \quad s(0^+) = -\infty.$$

For any $\omega_0 \in (0, \omega_c^{EC})$ and any $\eta \in (0, \omega_0)$, the probability of hitting ω_c^{EC} before η is

$$P_{\omega_0}(\tau_{\omega_c^{EC}} < \tau_\eta) = \frac{s(\omega_0) - s(\eta)}{s(\omega_c^{EC}) - s(\eta)} = \frac{1/\eta - 1/\omega_0}{1/\eta - \epsilon/a}.$$

Since $\omega_0 < \omega_c^{EC} = a/\epsilon$, we have $1/\omega_0 > \epsilon/a$, and as $\eta \rightarrow 0^+$ both numerator and denominator diverge to $+\infty$. Dividing through by $1/\eta$:

$$\frac{1/\eta - 1/\omega_0}{1/\eta - \epsilon/a} = \frac{1 - \eta/\omega_0}{1 - \eta\epsilon/a} \rightarrow 1 \quad \text{as } \eta \rightarrow 0^+.$$

Hence $P_{\omega_0}(\tau_{\omega_c^{EC}} < \infty) = \lim_{\eta \rightarrow 0^+} P_{\omega_0}(\tau_{\omega_c^{EC}} < \tau_\eta) = 1$. □

BO.13.5 Intermediation Gap and the Leverage Paradox

Proposition BO.16 (Q-martingale and gap under the equity constraint) *Under the equity constraint with QE:*

- (i) $Q_t = 1/\omega_t$ is a true martingale. The intermediation gap is zero.
- (ii) The Leverage Paradox (Proposition 2.4) holds: gap-free prices and unconditional equilibrium existence remain jointly impossible.

Proof. *Part (i).* The Q -martingale property (Appendix B.7) holds for any portfolio process ϕ_t ; in particular, it holds when $\phi_t = 1/a$ in the constrained region. To classify Q : near

$\omega = 0$, $|\sigma_\omega^{EC}| \sim c_{EC}\omega$, so $\sigma_Q = \sigma_\omega/\omega^2 \sim c_{EC}/\omega = c_{EC}Q$. The Mijatovic–Urusov integral test (Appendix B.11) gives

$$\int_c^\infty \frac{q}{\sigma_Q(q)^2} dq = \int_c^\infty \frac{q}{c_{EC}^2 q^2} dq = \frac{1}{c_{EC}^2} \int_c^\infty \frac{dq}{q} = +\infty.$$

Hence Q is a true martingale. By the characterisation in Theorem 2.6, the intermediation gap is zero.

Part (ii). If equilibrium exists for all ω without central bank intervention, then $\phi(\omega) = 1/(\epsilon\omega)$ everywhere. As $\omega \rightarrow 0$, $\phi \rightarrow \infty$, so $\sigma_\omega(0^+) = c_0 > 0$ (Appendix B.11). Then $\sigma_Q \sim c_0 Q^2$ and $\int q^{-3} dq < \infty$: Q is a strict local martingale and the gap is positive. This is independent of the constraint form. $\square$

BO.13.6 General Characterisation: When Does Permanent QE Exist?

The results above show that Permanent QE (an absorbing boundary inside the constrained region) is absent under the equity constraint but present under the leverage constraint, where $\underline{\omega}$ is absorbing and generates the adjacent QE Trap (Remark 2.2). What distinguishes the two constraint forms? The following theorem characterises when an admissible leverage function creates the absorbing-boundary phenomenon.

Definition BO.1 (Admissible leverage function) *A function $\bar{\phi} : (0, 1/\epsilon) \rightarrow (0, \infty)$ is admissible if: (i) $\bar{\phi}$ is continuous; (ii) there exists a unique $\omega_c \in (0, 1/\epsilon)$ such that $\bar{\phi}(\omega) < 1/(\epsilon\omega)$ for $\omega < \omega_c$ and $\bar{\phi}(\omega) \geq 1/(\epsilon\omega)$ for $\omega \geq \omega_c$; (iii) $(1 + \varphi)(1 - \omega) - \omega(\bar{\phi}(\omega) - 1) > 0$ for all $\omega \in (0, \omega_c)$ (the denominator in (BO.36) is well-defined and positive).*

Condition (iii) is automatically satisfied by any bounded $\bar{\phi}$ when ω is sufficiently small: at $\omega = 0$ the denominator equals $1 + \varphi > 0$, and by continuity it remains positive on a neighbourhood. For the baseline leverage constraint, condition (iii) is Lemma B.1. For the equity constraint, it is verified in Step 4 of Proposition BO.15.

Both the baseline leverage constraint $\bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega})$ (with ω_c defined by (2.10)) and the equity constraint $\bar{\phi}^{EC} = 1/a$ (with $\omega_c^{EC} = a/\epsilon$) are admissible.

Theorem BO.9 (Permanent QE characterisation) Let $\bar{\phi}$ be an admissible leverage function. With QE restoring equilibrium in the constrained region ($\omega < \omega_c$), the financier's wealth-share volatility is

$$\sigma_\omega(\omega) = \frac{\sigma(1+\varphi)(1-\omega)\omega(\bar{\phi}(\omega)-1)}{(1+\varphi)(1-\omega)-\omega(\bar{\phi}(\omega)-1)}, \quad \omega \in (0, \omega_c). \quad (\text{BO.36})$$

Then:

- (i) **No Permanent QE.** If $\bar{\phi}(\omega) > 1$ for all $\omega \in (0, \omega_c)$, then $\sigma_\omega(\omega) > 0$ throughout the constrained region, $\omega_t \rightarrow 1/\epsilon$ a.s. from every initial condition, QE is self-resolving, and there is no absorbing bad steady state. The QE Trap is likewise absent since it requires an adjacent absorbing boundary.
- (ii) **Permanent QE.** If $\bar{\phi}(\omega^*) = 1$ for some $\omega^* \in (0, \omega_c)$, then $\sigma_\omega(\omega^*) = \mu_\omega(\omega^*) = 0$, and ω^* is an absorbing steady state. Moreover:
 - If $\bar{\phi}(\omega) < 1$ for $\omega \in (0, \omega^*)$ and $\bar{\phi}(\omega) > 1$ for $\omega \in (\omega^*, \omega_c)$, then the constrained region splits into a Permanent QE regime $(0, \omega^*)$ converging to ω^* (QE literally permanent) and a recovery regime (ω^*, ω_c) converging to $1/\epsilon$. The QE Trap sub-region $(\omega^*, \omega^* + \Delta)$ with slow drift lies adjacent to the absorbing boundary; under the Table 2.1 leverage calibration, Appendix BO.14 reports a 313-year deterministic drift time from $\omega_0 = 0.05$ to ω_c .
 - The point ω^* is an unattainable natural boundary: ω_t converges to ω^* but never reaches it in finite time.
- (iii) In both cases, $Q = 1/\omega$ is a true martingale (no intermediation gap), and the Leverage Paradox holds.

Proof. The volatility formula (BO.36) follows from the same derivation as in Appendix B.6, with $\bar{\phi}(\omega)$ in place of $1 + \lambda(\omega - \underline{\omega})$. In the constrained region, $\phi = \bar{\phi}(\omega)$, so $\phi - 1 = \bar{\phi}(\omega) - 1$. The fixed-point equation (Step 3 of Proposition BO.15) yields (BO.36). The drift $\mu_\omega = \sigma_\omega^2/\omega$ follows from the Q-martingale property, which is independent of the constraint form (Appendix B.7).

Part (i). If $\bar{\phi}(\omega) > 1$ for all $\omega \in (0, \omega_c)$, then $\bar{\phi}(\omega) - 1 > 0$, and every factor in the numerator of (BO.36) is strictly positive for $\omega \in (0, \omega_c) \subset (0, 1)$. The denominator is positive by admissibility condition (iii). Hence $\sigma_\omega > 0$ and $\mu_\omega = \sigma_\omega^2/\omega > 0$ throughout $(0, \omega_c)$.

With no interior zero of σ_ω , the interval $(0, 1/\epsilon)$ contains no internal unattainable points. The argument then follows Theorem BO.8(ii)–(iii): $(0, 1/\epsilon)$ is a single invariant regime, $\omega_t \rightarrow 1/\epsilon$ a.s., and the economy exits the QE region in finite time a.s.

Part (ii). If $\bar{\phi}(\omega^*) = 1$, then $\sigma_\omega(\omega^*) = 0$ (the factor $\bar{\phi}(\omega^*) - 1 = 0$ in the numerator of (BO.36)). Since $\mu_\omega = \sigma_\omega^2/\omega$, the drift also vanishes: $\mu_\omega(\omega^*) = 0$. The SDE $d\omega = (\sigma_\omega^2/\omega) dt + \sigma_\omega dZ$ with both coefficients zero at ω^* admits $\omega_t = \omega^*$ as a solution. Pathwise uniqueness (from the local Lipschitz property of σ_ω away from boundaries) implies this is the unique solution starting from ω^* : the point is absorbing.

Boundary classification at ω^* . Since $\bar{\phi}$ is continuous and $\bar{\phi}(\omega^*) = 1$, we write $\bar{\phi}(\omega) - 1 = h(\omega)(\omega - \omega^*)$ where h is continuous and $h(\omega^*) = \bar{\phi}'(\omega^*)$ if $\bar{\phi}$ is differentiable at ω^* . If $h(\omega^*) \neq 0$ (the generic case, including our baseline $\bar{\phi}'(\underline{\omega}) = \lambda > 0$):

$$\sigma_\omega(\omega) \sim g(\omega^*)(\omega - \omega^*), \quad g(\omega^*) = \frac{\sigma(1+\varphi)(1-\omega^*)\omega^* h(\omega^*)}{D(\omega^*)} \neq 0,$$

so $\sigma_\omega^2 \sim g(\omega^*)^2(\omega - \omega^*)^2$ near ω^* . The speed measure density satisfies

$$m(\omega) = \frac{\omega^2}{\omega_0^2 \sigma_\omega(\omega)^2} \geq \frac{K}{(\omega - \omega^*)^2}$$

for some $K > 0$. Therefore

$$\int^x m(y) dy \geq K \int^x \frac{dy}{(y - \omega^*)^2} = \frac{K}{|x - \omega^*|} \rightarrow +\infty \quad \text{as } x \rightarrow \omega^*,$$

and the Feller N -integral diverges:

$$N(\omega^*) = \int s'(x) \int^x m(y) dy dx = +\infty,$$

confirming that ω^* is a natural (unattainable) boundary, exactly as at $\omega = \underline{\omega}$ in Appendix B.9. (If $h(\omega^*) = 0$, so that σ_ω vanishes faster than linearly, the speed measure diverges *a fortiori* and ω^* is still unattainable.)

Regime structure. For $\omega < \omega^*$: $\bar{\phi}(\omega) < 1$ (by assumption), so $\bar{\phi}(\omega) - 1 < 0$ and $\sigma_\omega < 0$ (perverse dynamics: positive shocks reduce ω). For $\omega \in (\omega^*, \omega_c)$: $\bar{\phi}(\omega) > 1$, so $\sigma_\omega > 0$ (standard recovery dynamics). Continuous paths cannot cross the unattainable boundary ω^* , establishing regime invariance. In the crisis regime $(0, \omega^*)$: the drift $\mu_\omega = \sigma_\omega^2/\omega > 0$ pushes ω toward ω^* , and the process converges a.s. to ω^* (QE permanent, by the same supermartingale argument as in Part (i) applied to Q on the sub-interval $(1/\omega^*, \infty)$). In the recovery regime (ω^*, ω_c) : the process exits the QE region in finite time a.s. (QE self-resolving, by the scale function argument in

Theorem [BO.8\(iii\)](#)).

Part (iii). Near $\omega = 0$, any admissible $\bar{\phi}$ is bounded, so $|\sigma_\omega| = \omega|\bar{\phi} - 1||\sigma_R| \sim C\omega$ for some $C > 0$. Thus $\sigma_Q \sim CQ$ and $\int q/(C^2q^2) dq = +\infty$: Q is a true martingale regardless of whether the trap exists. The Leverage Paradox follows as before. $\square$

Remark BO.7 (Permanent QE requires deleveraging to zero leverage) *The theorem identifies the precise economic mechanism behind Permanent QE: the leverage function must force the financier to be exactly unlevered ($\bar{\phi} = 1$) at some interior point ω^* . (The QE Trap then arises mechanically in the neighbourhood of ω^* because the drift $\mu_\omega = \sigma_\omega^2/\omega$ vanishes as $\omega \rightarrow \omega^*$.)*

The economics is as follows. At ω^* with $\phi = 1$, the financier invests her entire wealth in equity and holds no bonds. She does earn the equity premium: her wealth evolves as $dH_i/H_i = (\mu_R - \delta) dt + \sigma_R dZ$. But at ω^* , the volatility $\sigma_\omega = 0$ implies $\sigma_R = \sigma_Y = \sigma$ and $\mu_R = \delta + \mu$ (substituting $r = \delta + \mu - \sigma^2$ and $\Theta\sigma_R = \sigma^2$). Thus $dH_i/H_i = \mu dt + \sigma dZ$, which is identical to dP_Y/P_Y . The financier's wealth tracks total wealth exactly, so the wealth share $\omega = H_i/P_Y$ is frozen: $d\omega = 0$. The financier earns the market return but cannot outgrow the economy; recovery requires $\phi > 1$, i.e., leverage above 1.

Under the equity constraint with $a < 1$, the financier retains leverage $1/a > 1$ at every state. Her leveraged position amplifies the equity premium: $\phi(\mu_R - r) > (\mu_R - r)$. Her wealth grows faster than P_Y , so $\sigma_\omega > 0$ and $\mu_\omega = \sigma_\omega^2/\omega > 0$. Recovery is endogenous.

Permanent QE is therefore a consequence of procyclical deleveraging to zero leverage at an interior point—the defining feature of the leverage-constraint tradition—and cannot arise under an equity constraint that preserves $\phi > 1$ everywhere in the interior. The QE Trap phenomenon of impractically slow recovery does not disappear with it, however: under the equity constraint, $\mu_\omega \rightarrow 0$ as $\omega \rightarrow 0$ and the leading-order recovery time from a severely depleted ω_0 still grows logarithmically in $1/\omega_0$. Quantitatively, the horizon is governed by $(1/a - 1)^2(0.1505)^2$ at the [Table 2.1](#) volatility, while [Appendix BO.14](#) gives the calibrated leverage-constraint benchmark of 313 years from $\omega_0 = 0.05$ to ω_c . The distinction is qualitative: an absorbing interior boundary versus an unattainable boundary at zero.

BO.13.7 Comparison Table

Table [BO.3](#) collects the comparison. All results in the table are proved above or follow directly from the general characterisation in Theorem [BO.9](#).

Table BO.3: Leverage constraint vs. equity constraint: complete comparison

	Leverage constraint $\phi \leq 1 + \lambda(\omega - \underline{\omega})$	Equity constraint $\phi \leq 1/a$
<i>Constraint specification</i>		
Tradition	Adrian–Shin, Brunnermeier–Pedersen	He–Krishnamurthy
Leverage cap in constrained region	$\bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega})$ (state-dependent, procyclical)	$\bar{\phi}^{EC} = 1/a$ (constant)
Leverage at worst state	$\bar{\phi}(\underline{\omega}) = 1$ (unlevered)	$\bar{\phi}^{EC} = 1/a > 1$ (leveraged)
Crisis threshold ω_c	$\lambda\omega_c(\omega_c - \underline{\omega}) = 1/\epsilon - \omega_c$	a/ϵ
<i>Dynamics of ω</i>		
σ_ω (constrained)	$\frac{\sigma(1+\varphi)\lambda(1-\omega)\omega(\omega-\underline{\omega})}{D_c(\omega)}$	$\frac{\sigma(1+\varphi)(1/a-1)(1-\omega)\omega}{D_{EC}(\omega)}$
Interior zeros of σ_ω	$\underline{\omega}$ (where $\bar{\phi} = 1$)	None ($\bar{\phi} = 1/a > 1$ always)
Sign of σ_ω for $\omega < \omega_c$	Negative on $(0, \underline{\omega})$; positive on $(\underline{\omega}, \omega_c)$	Positive everywhere
Regimes	Crisis $(0, \underline{\omega})$, Recovery $(\underline{\omega}, 1/\epsilon)$, Inverted $(1/\epsilon, 1)$	Recovery $(0, 1/\epsilon)$, Inverted $(1/\epsilon, 1)$
Absorbing steady states	$\underline{\omega}$ (bad) and $1/\epsilon$ (good)	$1/\epsilon$ only
<i>QE and policy</i>		
CB equity: α_t	$\max(0, 1 - \epsilon\bar{\phi}(\omega)\omega)$	$\max(0, 1 - \epsilon\omega/a)$
Permanent QE	Yes ($\omega \rightarrow \underline{\omega}$ if $\omega_0 < \underline{\omega}$)	No
QE Trap (slow-drift region)	Yes (in $(\underline{\omega}, \omega_T)$: $\mathbb{E}[\tau_{\omega_c}] = \infty$)	Yes near $\omega = 0$ (slow drift, but no interior threshold ω_T)
QE exit from $(0, \omega_c)$	Permanent from $(0, \underline{\omega})$; slow but finite-time a.s. from $(\underline{\omega}, \omega_c)$	Finite-time a.s. from all $\omega_0 > 0$; $\mathbb{E}[\tau] \rightarrow \infty$ as $\omega_0 \rightarrow 0$
Fiscal escape needed	Yes (to cross $\underline{\omega}$)	Yes if ω_0 is severely depleted
CB equity at bad SS	$1 - \epsilon\underline{\omega}$ (permanent; 64% at baseline)	N/A (no bad steady state)
<i>Asset pricing and intermediation gap</i>		
$Q = 1/\omega$ classification	True martingale	True martingale
Intermediation gap	Zero	Zero
Leverage Paradox	Yes	Yes
Equilibrium failure	Yes ($\omega < \omega_c$)	Yes ($\omega < a/\epsilon$)
QE necessity	Yes	Yes

BO.13.8 Verification: Special and Limiting Cases

We verify the results against four limiting cases.

Case 1: $a \rightarrow 0$ (no constraint). As $a \rightarrow 0$: $1/a \rightarrow \infty$, $\omega_c^{EC} = a/\epsilon \rightarrow 0$, and $\alpha_t^{EC} = 1 - \epsilon\omega/a \rightarrow -\infty$ for any $\omega > 0$, so $\alpha_t^{EC} = 0$ (the constraint never binds). The model reduces to the LP baseline without leverage constraint. All dynamics are governed by $\sigma_\omega^u(\omega)$ on $(0, 1/\epsilon)$: Q is a

strict local martingale, the intermediation gap is positive, and $\omega \rightarrow 1/\epsilon$ a.s. This is consistent with Table BO.3: removing the constraint removes both the failure region and the elimination of the gap.

Case 2: $a \rightarrow 1$ (autarky). As $a \rightarrow 1$: $1/a \rightarrow 1$, $\omega_c^{EC} \rightarrow 1/\epsilon$, and $\bar{\phi}^{EC} \rightarrow 1$. The financier is forced to be unlevered ($\phi = 1$) at every state, giving $\sigma_\omega^{EC} \rightarrow 0$ everywhere. Every point is absorbing: $\omega_t = \omega_0$ for all t . This degenerate case is excluded by $a < 1$. It is the limiting case of Theorem BO.9(ii): $\bar{\phi}(\omega) = 1$ for all ω , so every ω^* satisfies $\bar{\phi}(\omega^*) = 1$.

Case 3: recovery of baseline model. Setting $\bar{\phi}(\omega) = 1 + \lambda(\omega - \underline{\omega})$ in (BO.36) gives $\bar{\phi}(\omega) - 1 = \lambda(\omega - \underline{\omega})$, recovering the baseline formula $\sigma_\omega^c = \sigma(1 + \varphi)\lambda(1 - \omega)\omega(\omega - \underline{\omega})/D_c(\omega)$. The zero at $\omega^* = \underline{\omega}$ (where $\bar{\phi} = 1$) confirms the trap. Setting $\bar{\phi} = 1/a$ gives $\bar{\phi} - 1 = (1 - a)/a$ (a positive constant), recovering (BO.33) with no interior zero.

Case 4: continuity at ω_c^{EC} (numerical). At the Table 2.1 calibration $\epsilon = 13.7741$, $\varphi = 1.3333$, $\sigma = 0.1505$, and taking $a = 0.1$: $\omega_c^{EC} = a/\epsilon = 0.00726$.

Constrained formula (BO.33) at $\omega = 0.00726$:

$$\text{Numerator} = 0.1505 \times 2.3333 \times 9 \times 0.99274 \times 0.00726 = 0.02278,$$

$$D_{EC}(0.00726) = 2.3333 - 11.3333 \times 0.00726 = 2.25102,$$

$$\sigma_\omega^{EC} = 0.02278/2.25102 = 0.01012.$$

Unconstrained formula at $\omega = 0.00726$:

$$\text{Numerator} = 0.1505 \times 2.3333 \times 0.99274 \times (0.07260 - 0.00726) = 0.02278,$$

$$D_u(0.00726) = 2.3333 \times 0.99274 - (0.07260 - 0.00726) = 2.25102,$$

$$\sigma_\omega^u = 0.02278/2.25102 = 0.01012. \quad \checkmark$$

The two formulas agree at the threshold, confirming continuity.

Boundary approximation at $\omega = 0.001$: $c_{EC} = 0.1505 \times 9 = 1.3545$. Approximate: $\sigma_\omega \approx c_{EC}\omega = 1.3545 \times 0.001 = 0.0013545$. Exact: $\sigma_\omega^{EC}(0.001) = 0.1505 \times 2.3333 \times 9 \times 0.999 \times 0.001/(2.3333 - 11.3333 \times 0.001) = 0.003157/2.32197 = 0.001360$. Relative error: 0.4%. $\checkmark$

BO.13.9 Economic Interpretation and Policy Implications

Remark BO.8 (Regulatory mapping) *In practice, both types of constraint coexist. The Basel III Supplementary Leverage Ratio (SLR) and initial margin requirements are leverage constraints: they cap the ratio of exposures to capital, independent of risk. Risk-weighted capital requirements (Basel II/III risk-based capital, Dodd–Frank stress-test hurdle rates) are closer to equity constraints: they require the intermediary to maintain equity proportional to the risk of her portfolio, preserving leverage above 1 as long as equity is positive.*

Theorem BO.9 implies that when the binding constraint is of the leverage type—as it was for broker-dealers during the 2007–09 crisis, when leverage ratios rather than risk-based capital were the operative limit—QE risks becoming permanent. When the binding constraint is of the equity type, QE is self-resolving. The distinction determines whether fiscal recapitalisation is a necessity or a convenience.

Remark BO.9 (Japan vs. the United States) *The Bank of Japan has maintained large-scale asset purchases since 2001, with its balance sheet exceeding 130% of GDP by 2024—a duration and scale consistent with either Permanent QE or the adjacent QE Trap under a leverage constraint (Theorem BO.9(ii)). By contrast, the Federal Reserve successfully tapered its post-2020 purchases by 2022, consistent with self-resolving QE under an equity constraint (Theorem BO.9(i)) or with the U.S. economy starting in the recovery regime $(\underline{\omega}, \omega_c)$ from which QE exit occurs in finite time even under the leverage constraint. Distinguishing these explanations requires identifying whether the binding regulatory constraint on the intermediary sector is of the leverage or equity type—a testable implication.*

BO.14 Drift-Time Analysis for Recovery Dynamics

BO.14.1 Drift-Time Lower Bound on Recovery

We analyse the recovery dynamics of the financier wealth share ω_t on the interval $(\underline{\omega}, \omega_c)$, where $\underline{\omega}$ is the lower absorbing boundary and ω_c is the critical threshold at which leverage constraints cease to bind. The drift time provides a deterministic lower bound on the expected first-passage time from any initial condition ω_0 to ω_c .

Proposition BO.17 (Drift-time integrand) *The wealth-share process ω_t satisfies*

$$d\omega_t = \mu_\omega(\omega_t) dt + \sigma_\omega(\omega_t) dZ_t, \quad (\text{BO.37})$$

with drift $\mu_\omega(\omega) = \sigma_\omega^2(\omega)/\omega$ and volatility

$$\sigma_\omega(\omega) = \frac{\sigma(1+\varphi)\lambda(1-\omega)\omega(\omega-\underline{\omega})}{D_c(\omega)}, \quad (\text{BO.38})$$

where $D_c(\omega) \equiv (1+\varphi)(1-\omega) - \lambda\omega(\omega-\underline{\omega})$. Define the drift time from ω_0 to ω_c as

$$T(\omega_0) \equiv \int_{\omega_0}^{\omega_c} \frac{\omega}{\sigma_\omega^2(\omega)} d\omega. \quad (\text{BO.39})$$

This is the time required for the deterministic ODE $\dot{\omega} = \mu_\omega(\omega)$ to travel from ω_0 to ω_c . By the comparison principle for diffusions (Karatzas and Shreve, 1991, Ch. 5.5), $T(\omega_0)$ is a lower bound on $\mathbb{E}[\tau_{\omega_c} \mid \omega_0]$, the expected first-passage time to ω_c , since diffusion adds variance but does not increase the expected speed.

The integrand admits the closed form

$$\frac{\omega}{\sigma_\omega^2(\omega)} = \frac{D_c(\omega)^2}{\sigma^2(1+\varphi)^2\lambda^2(1-\omega)^2\omega(\omega-\underline{\omega})^2}. \quad (\text{BO.40})$$

Proof. From (BO.38),

$$\sigma_\omega^2(\omega) = \frac{\sigma^2(1+\varphi)^2\lambda^2(1-\omega)^2\omega^2(\omega-\underline{\omega})^2}{D_c(\omega)^2},$$

so $\omega/\sigma_\omega^2(\omega) = \omega \cdot D_c(\omega)^2/[\sigma^2(1+\varphi)^2\lambda^2(1-\omega)^2\omega^2(\omega-\underline{\omega})^2]$, which simplifies to (BO.40). $\square$

BO.14.2 Exact Antiderivative via Partial Fractions

Proposition BO.18 (Partial-fraction decomposition and antiderivative) Let $\underline{\omega} \in (0, 1)$, $\sigma > 0$, $\varphi \geq 0$, $\lambda > 0$. Define

$$\ell \equiv \sigma^2\lambda^2\underline{\omega}^2, \quad p \equiv \sigma^2(1+\varphi)^2, \quad q \equiv \sigma^2\lambda(1+\varphi)(1-\underline{\omega}). \quad (\text{BO.41})$$

The integrand (BO.40) has the partial-fraction decomposition

$$\frac{\omega}{\sigma_\omega^2(\omega)} = \frac{A_0}{\omega} + \frac{A_1}{1-\omega} + \frac{A_2}{\omega-\underline{\omega}} + \frac{E_1}{(1-\omega)^2} + \frac{E_2'}{(\omega-\underline{\omega})^2}, \quad (\text{BO.42})$$

where the sign convention uses $1/(1-\omega)$ (not $1/(\omega-1)$) and the coefficients are

$$A_0 = \frac{1}{\ell} = \frac{1}{\sigma^2\lambda^2\underline{\omega}^2}, \quad (\text{BO.43})$$

$$A_1 = \frac{1}{p} + \frac{2}{q} = \frac{1}{\sigma^2(1+\varphi)^2} + \frac{2}{\sigma^2\lambda(1+\varphi)(1-\underline{\omega})}, \quad (\text{BO.44})$$

$$A_2 = -\frac{2}{q} - \frac{1}{\ell} = -\frac{2}{\sigma^2 \lambda (1 + \varphi)(1 - \underline{\omega})} - \frac{1}{\sigma^2 \lambda^2 \underline{\omega}^2}, \quad (\text{BO.45})$$

$$E_1 = \frac{1}{p} = \frac{1}{\sigma^2 (1 + \varphi)^2}, \quad (\text{BO.46})$$

$$E_2 = -\frac{1}{\sigma^2 \lambda^2 \underline{\omega}}. \quad (\text{BO.47})$$

Here E'_2 in (BO.42) equals $-E_2 = 1/(\sigma^2 \lambda^2 \underline{\omega})$, and E_2 is the coefficient that appears in the antiderivative after integration. These satisfy the identity $A_0 + A_1 + A_2 = 1/p$.

The antiderivative is

$$F(\omega) = A_0 \ln \omega + A_1 \ln(1 - \omega) + A_2 \ln(\omega - \underline{\omega}) + \frac{E_1}{1 - \omega} + \frac{E_2}{\omega - \underline{\omega}}, \quad (\text{BO.48})$$

valid for $\omega \in (\underline{\omega}, 1)$, and the drift time is $T(\omega_0) = F(\omega_c) - F(\omega_0)$.

Proof. The denominator of (BO.40) is $\sigma^2(1 + \varphi)^2 \lambda^2 (1 - \omega)^2 \omega (\omega - \underline{\omega})^2$, which has roots at $\omega = 0$ (simple), $\omega = 1$ (double), and $\omega = \underline{\omega}$ (double). The numerator $D_c(\omega)^2 = [(1 + \varphi)(1 - \omega) - \lambda \omega (\omega - \underline{\omega})]^2$ is a quartic in ω . Since the denominator has degree 5, the fraction is a proper rational function and admits the partial-fraction form (BO.42).

Coefficient A_0 . Multiply (BO.42) by ω and set $\omega = 0$:

$$A_0 = \frac{D_c(0)^2}{\sigma^2(1 + \varphi)^2 \lambda^2 \cdot 1 \cdot \underline{\omega}^2} = \frac{(1 + \varphi)^2}{\sigma^2(1 + \varphi)^2 \lambda^2 \underline{\omega}^2} = \frac{1}{\sigma^2 \lambda^2 \underline{\omega}^2}.$$

Coefficients E_1 and E'_2 . Multiply by $(1 - \omega)^2$ and set $\omega = 1$:

$$E_1 = \frac{D_c(1)^2}{\sigma^2(1 + \varphi)^2 \lambda^2 \cdot 1 \cdot (1 - \underline{\omega})^2} = \frac{\lambda^2(1 - \underline{\omega})^2}{\sigma^2(1 + \varphi)^2 \lambda^2 (1 - \underline{\omega})^2} = \frac{1}{\sigma^2(1 + \varphi)^2},$$

using $D_c(1) = -\lambda(1 - \underline{\omega})$ so $D_c(1)^2 = \lambda^2(1 - \underline{\omega})^2$. Similarly, multiply by $(\omega - \underline{\omega})^2$ and set $\omega = \underline{\omega}$:

$$E'_2 = \frac{D_c(\underline{\omega})^2}{\sigma^2(1 + \varphi)^2 \lambda^2 (1 - \underline{\omega})^2 \underline{\omega}} = \frac{(1 + \varphi)^2 (1 - \underline{\omega})^2}{\sigma^2(1 + \varphi)^2 \lambda^2 (1 - \underline{\omega})^2 \underline{\omega}} = \frac{1}{\sigma^2 \lambda^2 \underline{\omega}},$$

using $D_c(\underline{\omega}) = (1 + \varphi)(1 - \underline{\omega})$.

Coefficients A_1 and A_2 . These are obtained by the Heaviside cover-up method applied to the first-order poles, after subtracting the known second-order contributions and using the identity $A_0 + A_1 + A_2 = \lim_{\omega \rightarrow \infty} \omega \cdot [\text{integrand}] = 1/p$ (from the leading behaviour of the quartic numerator divided by the quintic denominator, which yields $\lambda^2/(\sigma^2(1 + \varphi)^2 \lambda^2) = 1/p$). The

resulting expressions are (BO.44)–(BO.45).

Verification by differentiation. Computing $F'(\omega)$ from (BO.48):

$$F'(\omega) = \frac{A_0}{\omega} - \frac{A_1}{1-\omega} + \frac{A_2}{\omega-\underline{\omega}} + \frac{E_1}{(1-\omega)^2} - \frac{E_2}{(\omega-\underline{\omega})^2}.$$

Since $-E_2 = E'_2 = 1/(\sigma^2\lambda^2\underline{\omega})$, this matches (BO.42). Numerical verification at the Table 2.1 calibration confirms $|F'(\omega) - \omega/\sigma_\omega^2(\omega)| < 10^{-6}$ for all $\omega \in \{0.030, 0.040, 0.050, 0.060\}$. $\square$

BO.14.3 Leading-Order Approximation Near $\underline{\omega}$

Proposition BO.19 (Leading-order drift time and inversion) *For ω near $\underline{\omega}$, set $x = \omega - \underline{\omega}$. Then*

$$\sigma_\omega(\omega) = g x + O(x^2), \quad g \equiv \sigma\lambda\underline{\omega}, \quad (\text{BO.49})$$

and the leading-order drift time from ω_0 to ω_c is

$$T^{\text{LO}}(\omega_0) = \frac{\omega}{g^2} \left[\frac{1}{\omega_0 - \underline{\omega}} - \frac{1}{\omega_c - \underline{\omega}} \right] = \frac{1}{\sigma^2\lambda^2\underline{\omega}} \left[\frac{1}{\omega_0 - \underline{\omega}} - \frac{1}{\omega_c - \underline{\omega}} \right]. \quad (\text{BO.50})$$

This is invertible: the position after deterministic drift time T starting from ω_c and running backward is

$$\omega_T^{\text{LO}} = \underline{\omega} + \frac{\omega_c - \underline{\omega}}{1 + \sigma^2\lambda^2\underline{\omega}(\omega_c - \underline{\omega})T}. \quad (\text{BO.51})$$

Proof. *Step 1: Expansion of σ_ω .* At $\omega = \underline{\omega} + x$,

$$\begin{aligned} D_c(\underline{\omega} + x) &= (1 + \varphi)(1 - \underline{\omega} - x) - \lambda(\underline{\omega} + x)x \\ &= (1 + \varphi)(1 - \underline{\omega}) - [(1 + \varphi) + \lambda\underline{\omega}]x - \lambda x^2. \end{aligned}$$

At $x = 0$, $D_c(\underline{\omega}) = (1 + \varphi)(1 - \underline{\omega})$. The numerator of σ_ω is $\sigma(1 + \varphi)\lambda(1 - \underline{\omega} - x)(\underline{\omega} + x)x$. At leading order in x :

$$\sigma_\omega(\underline{\omega} + x) = \frac{\sigma(1 + \varphi)\lambda(1 - \underline{\omega})\underline{\omega}x}{(1 + \varphi)(1 - \underline{\omega})} + O(x^2) = \sigma\lambda\underline{\omega}x + O(x^2) = g x + O(x^2).$$

Step 2: Drift-time integral. Substituting $\sigma_\omega \approx gx$ and $\omega \approx \underline{\omega}$ in the integrand $\omega/\sigma_\omega^2 \approx$

$\underline{\omega}/(g^2 x^2)$:

$$T^{\text{LO}}(\omega_0) = \int_{\omega_0}^{\omega_c} \frac{\underline{\omega}}{g^2(\omega - \underline{\omega})^2} d\omega = \frac{\underline{\omega}}{g^2} \left[-\frac{1}{\omega - \underline{\omega}} \right]_{\omega_0}^{\omega_c} = \frac{\underline{\omega}}{g^2} \left[\frac{1}{\omega_0 - \underline{\omega}} - \frac{1}{\omega_c - \underline{\omega}} \right].$$

Step 3: *Inversion.* Setting $T = T^{\text{LO}}(\omega_0)$ and solving for ω_0 :

$$\frac{1}{\omega_0 - \underline{\omega}} = \frac{g^2 T}{\underline{\omega}} + \frac{1}{\omega_c - \underline{\omega}} = \frac{1 + \sigma^2 \lambda^2 \underline{\omega}(\omega_c - \underline{\omega}) T}{\omega_c - \underline{\omega}},$$

giving (BO.51).

Step 4: *Limit verification.* As $T \rightarrow 0$, the denominator in (BO.51) tends to 1, so $\omega_T^{\text{LO}} \rightarrow \underline{\omega} + (\omega_c - \underline{\omega}) = \omega_c$. As $T \rightarrow \infty$, the denominator diverges, so $\omega_T^{\text{LO}} \rightarrow \underline{\omega}$. $\square$

BO.14.4 Error Bound

Proposition BO.20 (One-sided bound and approximation quality) *The leading-order approximation overestimates the drift time: $T^{\text{LO}}(\omega_0) \geq T(\omega_0)$ for all $\omega_0 \in (\underline{\omega}, \omega_c)$. Equivalently, $\omega_T^{\text{LO}} \geq \omega_T^{\text{exact}}$ for all $T > 0$.*

Proof. Define $R(\omega) \equiv \sigma_\omega(\omega)/[g(\omega - \underline{\omega})]$ for $\omega > \underline{\omega}$. From (BO.38),

$$R(\omega) = \frac{(1 + \varphi)(1 - \omega) \omega}{\underline{\omega} D_c(\omega)}. \quad (\text{BO.52})$$

At $\omega = \underline{\omega}$: $R(\underline{\omega}) = (1 + \varphi)(1 - \underline{\omega}) \underline{\omega}/[\underline{\omega}(1 + \varphi)(1 - \underline{\omega})] = 1$.

For $\omega > \underline{\omega}$, $D_c(\omega) < (1 + \varphi)(1 - \omega)$ since $\lambda \omega(\omega - \underline{\omega}) > 0$, so $(1 + \varphi)(1 - \omega)/D_c(\omega) > 1$. Moreover, $\omega/\underline{\omega} > 1$. Hence $R(\omega) > 1$ for all $\omega \in (\underline{\omega}, \omega_c)$.

The exact integrand can be written as

$$\frac{\omega}{\sigma_\omega^2(\omega)} = \frac{\omega}{g^2(\omega - \underline{\omega})^2 R(\omega)^2}.$$

Since $R(\omega) > 1$, we have $\omega/\sigma_\omega^2(\omega) < \omega/[g^2(\omega - \underline{\omega})^2]$. The leading-order integrand is $\underline{\omega}/[g^2(\omega - \underline{\omega})^2] \leq \omega/[g^2(\omega - \underline{\omega})^2]$ (since $\omega \geq \underline{\omega}$). However, the dominant effect is $R^2 > 1$ in the denominator. We show $\omega/(R(\omega)^2) \leq \underline{\omega}$ for all $\omega \in [\underline{\omega}, \omega_c]$ when ω_c is small, which suffices.

More precisely, the ratio of exact to leading-order integrand is

$$\frac{\omega/\sigma_\omega^2}{\underline{\omega}/[g(\omega - \underline{\omega})]^2} = \frac{\omega}{\underline{\omega} R(\omega)^2}. \quad (\text{BO.53})$$

Computing the derivative at $\omega = \underline{\omega}$: let $x = \omega - \underline{\omega}$. The first-order expansion gives

$$\frac{\omega}{\underline{\omega} R(\omega)^2} = 1 + \left[\frac{1}{\underline{\omega}} - 2R'(\underline{\omega}) \right] x + O(x^2),$$

where

$$R'(\underline{\omega}) = \frac{1 - 2\underline{\omega}}{\underline{\omega}(1 - \underline{\omega})} + \frac{(1 + \varphi) + \lambda \underline{\omega}}{(1 + \varphi)(1 - \underline{\omega})}.$$

At the Table 2.1 calibration ($\underline{\omega} = 0.0260$, $\varphi = 1.3333$, $\lambda = 5.7231$), $1/\underline{\omega} = 38.46$ and $2R'(\underline{\omega}) \approx 77.05$, so the ratio (BO.53) is immediately decreasing below 1. Numerical evaluation confirms the ratio stays below 1 on the entire interval $(\underline{\omega}, \omega_c)$, establishing $T(\omega_0) \leq T^{\text{LO}}(\omega_0)$.

A clean closed-form bound on the relative error is not available because the ratio $R(\omega)$ varies from 1 at $\underline{\omega}$ to $R(\omega_c) \approx 2.34$ at $\omega_c = 0.0606$. Table BO.5 provides a numerical comparison. $\square$

BO.14.5 Recovery Timescale Table

Table BO.4 reports the exact drift time $T(\omega_0) = F(\omega_c) - F(\omega_0)$ computed from the antiderivative (BO.48) at the Table 2.1 calibration $\sigma = 0.1505$, $\varphi = 1.3333$, $\lambda = 5.7231$, $\underline{\omega} = 0.0260$, $\omega_c = 0.0606$, $\varepsilon = 13.7741$. The numerical coefficients are $A_0 = 1993.97$, $A_1 = 14.90$, $A_2 = -2000.75$, $E_1 = 8.11$, $E_2 = -51.84$.

Table BO.4: Exact drift time from ω_0 to $\omega_c = 0.0606$

ω_0	$\omega_0 - \underline{\omega}$	$T(\omega_0)$ (years)	Interpretation
0.027	0.001	44,866	45 millennia
0.028	0.002	20,259	20 millennia
0.030	0.004	8,547	9 millennia
0.035	0.009	2,662	27 centuries
0.040	0.014	1,223	12 centuries
0.045	0.019	624	6 centuries
0.050	0.024	313	3 centuries
0.055	0.029	129	long lifetime
0.058	0.032	53	working lifetime
0.060	0.034	11	a decade
0.0606	0.0346	0	at threshold

These drift times are *lower bounds* on the expected first-passage time. Even with the most optimistic reading (deterministic drift only, no diffusion), an economy that enters the crisis regime at $\omega_0 = 0.05$ requires more than three centuries of uninterrupted drift to reach ω_c without fiscal intervention. This establishes the quantitative force of the QE Trap: the drift is too slow to provide a practical exit.

Table BO.5: Approximation quality: leading-order vs. exact drift time

ω_0	T^{exact}	T^{LO}	$T^{\text{LO}}/T^{\text{exact}}$
0.027	44,866	50,345	1.12
0.028	20,259	24,423	1.21
0.030	8,547	11,462	1.34
0.035	2,662	4,262	1.60
0.040	1,223	2,205	1.80
0.045	624	1,230	1.97
0.050	313	662	2.11
0.055	129	289	2.24
0.058	53	122	2.30
0.060	11	26	2.34

The leading-order approximation is excellent near $\underline{\omega}$ (12% overestimate at $\omega_0 = 0.027$) but deteriorates at larger distances because the volatility $\sigma_\omega(\omega)$ grows superlinearly in $\omega - \underline{\omega}$ due to the ratio $R(\omega)$ in (BO.52). For quantitative work at moderate ω_0 , the exact antiderivative (BO.48) should be used. For qualitative and asymptotic results near $\underline{\omega}$, the closed-form (BO.51) suffices.

BO.15 Empirical Evidence Robustness Checks

BO.15.1 Test 1: Real Economic Activity

Test 1A This section reports robustness checks for Test 1A, which relates the model-implied real-activity gap to observed slack measures. The baseline specification uses the modern quarterly sample 1995Q1–2024Q4. We consider three alternative windows. First, we start the sample in 1970Q1, the earliest quarter available in the HKM intermediary-capital series, to check whether the sign of the relation survives the longest sample. Second, we start in 1997Q1,

aligning the quarterly test with the annual data window used to calibrate ϵ . Third, we keep the post-1995 sample but exclude 2007Q4–2009Q1, the crisis-transition period between the onset of intermediary-capital depletion and full QE activation.

Table BO.6: Test 1A – sample robustness, output gap

	1970Q1–2024Q4	1997Q1–2024Q4	Excl. 2007Q4–2009Q1
α	-0.002 (0.002)	-0.005* (0.003)	-0.005** (0.003)
β	0.879*** (0.185)	1.041*** (0.190)	1.085*** (0.196)
Adj. R^2	0.195	0.346	0.362
N	220	112	114

Notes: HAC standard errors with four lags are reported in parentheses. The dependent variable is the output gap. The columns report the longest available sample, the post-1997 sample, and the post-1995 sample excluding 2007Q4–2009Q1. ***, **, * denote significance at 1%, 5%, 10%.

Table BO.7: Test 1A – sample robustness, negative unemployment gap

	1970Q1–2024Q4	1997Q1–2024Q4	Excl. 2007Q4–2009Q1
α	-0.004*** (0.001)	-0.007*** (0.003)	-0.007*** (0.003)
β	0.684*** (0.134)	0.959*** (0.208)	1.022*** (0.235)
Adj. R^2	0.198	0.333	0.351
N	220	112	114

Notes: HAC standard errors with four lags are reported in parentheses. The dependent variable is the negative unemployment gap. The columns report the longest available sample, the post-1997 sample, and the post-1995 sample excluding 2007Q4–2009Q1. ***, **, * denote significance at 1%, 5%, 10%.

The robustness checks support the baseline finding. For both slack measures, the coefficient on the model-implied gap remains positive and statistically significant in all three alternative samples. The 1997Q1 and transition-exclusion estimates are close to the baseline estimates and remain near one. The 1970Q1 estimates are smaller and explain less variation, consistent with the older sample adding macro-financial regimes outside the main modern-window specification, but they do not overturn the sign or significance of the relation.

Test 1B This section reports robustness checks for Test 1B, which asks whether the state variable contains information about future real activity. The baseline specification uses the modern quarterly sample 1995Q1–2024Q4 and a parsimonious set of controls.

First, we estimate the Test 1B specification in alternative sample windows and exclusion samples: the longest available 1970Q1–2024Q4 sample, the 1997Q1–2024Q4 sample, the post-1995 sample excluding 2007Q4–2009Q1, and the post-1995 sample excluding 2020Q1–2021Q4. The COVID exclusion is useful because COVID was a large real-activity shock driven by public-health restrictions and policy responses rather than by the intermediary-capital mechanism alone. The low-capital interaction and the implied low-capital slope generally retain the predicted positive sign across these sample checks for the output gap, negative unemployment gap, payroll growth, and industrial production growth. The longer 1970 sample generally produces smaller coefficients, while excluding the crisis-transition quarters or COVID quarters reduces precision because these exclusions remove some of the relevant low-state variation.

Table BO.8: Test 1B – sample robustness, output gap

	$h = 1$	$h = 2$	$h = 4$
<i>1970Q1–2024Q4</i>			
β_h (high-capital slope)	-0.026 (0.082)	0.009 (0.156)	0.135 (0.266)
κ_h (low-capital interaction)	0.699** (0.324)	0.806* (0.451)	0.953 (0.680)
$\beta_h + \kappa_h$ (low-capital slope)	0.673** (0.319)	0.815** (0.415)	1.088* (0.604)
Low-state obs.	116	116	116
Adj. R^2	0.778	0.598	0.344
N	219	218	216
<i>1997Q1–2024Q4</i>			
β_h (high-capital slope)	0.067 (0.134)	0.160 (0.253)	0.294 (0.371)
κ_h (low-capital interaction)	1.559** (0.635)	1.634*** (0.491)	2.009*** (0.717)
$\beta_h + \kappa_h$ (low-capital slope)	1.627** (0.684)	1.794*** (0.490)	2.304*** (0.701)
Low-state obs.	43	43	43
Adj. R^2	0.742	0.559	0.366
N	111	110	108
<i>Excl. 2007Q4–2009Q1</i>			
β_h (high-capital slope)	0.074 (0.140)	0.152 (0.267)	0.156 (0.382)
κ_h (low-capital interaction)	1.698 (1.130)	0.909** (0.397)	1.084 (0.686)
$\beta_h + \kappa_h$ (low-capital slope)	1.772 (1.180)	1.061** (0.455)	1.240** (0.560)
Low-state obs.	37	37	37
Adj. R^2	0.730	0.535	0.352
N	112	110	106
<i>Excl. 2020Q1–2021Q4</i>			
β_h (high-capital slope)	-0.023 (0.079)	-0.018 (0.178)	0.073 (0.338)
κ_h (low-capital interaction)	0.754** (0.298)	1.491*** (0.509)	2.338*** (0.682)
$\beta_h + \kappa_h$ (low-capital slope)	0.731** (0.292)	1.473*** (0.518)	2.411*** (0.752)
Low-state obs.	39	39	39
Adj. R^2	0.927	0.836	0.635
N	110	108	104

Notes: Each block reports the h -step predictive specification from Test 1B. For excluded periods, rows are omitted when date t or any forecasted quarter used in the dependent variable intersects the excluded window. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Table BO.9: Test 1B – sample robustness, negative unemployment gap

	$h = 1$	$h = 2$	$h = 4$
<i>1970Q1–2024Q4</i>			
β_h (high-capital slope)	0.037 (0.066)	0.040 (0.103)	0.141 (0.177)
κ_h (low-capital interaction)	0.734** (0.337)	0.976*** (0.330)	1.147** (0.483)
$\beta_h + \kappa_h$ (low-capital slope)	0.771** (0.382)	1.016*** (0.355)	1.289*** (0.481)
Low-state obs.	116	116	116
Adj. R^2	0.796	0.648	0.444
N	219	218	216
<i>1997Q1–2024Q4</i>			
β_h (high-capital slope)	0.093 (0.087)	0.152 (0.148)	0.315 (0.253)
κ_h (low-capital interaction)	1.595** (0.737)	1.802*** (0.344)	2.240*** (0.547)
$\beta_h + \kappa_h$ (low-capital slope)	1.687** (0.785)	1.954*** (0.401)	2.554*** (0.509)
Low-state obs.	43	43	43
Adj. R^2	0.786	0.640	0.493
N	111	110	108
<i>Excl. 2007Q4–2009Q1</i>			
β_h (high-capital slope)	0.094 (0.095)	0.147 (0.157)	0.252 (0.274)
κ_h (low-capital interaction)	2.224 (1.456)	1.724** (0.714)	1.440*** (0.445)
$\beta_h + \kappa_h$ (low-capital slope)	2.318 (1.503)	1.871** (0.812)	1.692*** (0.474)
Low-state obs.	37	37	37
Adj. R^2	0.786	0.614	0.446
N	112	110	106
<i>Excl. 2020Q1–2021Q4</i>			
β_h (high-capital slope)	0.011 (0.030)	0.010 (0.063)	0.061 (0.144)
κ_h (low-capital interaction)	0.667*** (0.236)	1.264*** (0.407)	2.122*** (0.622)
$\beta_h + \kappa_h$ (low-capital slope)	0.678*** (0.229)	1.274*** (0.401)	2.183*** (0.646)
Low-state obs.	39	39	39
Adj. R^2	0.983	0.943	0.819
N	110	108	104

Notes: Each block reports the h -step predictive specification from Test 1B. For excluded periods, rows are omitted when date t or any forecasted quarter used in the dependent variable intersects the excluded window. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Table BO.10: Test 1B – sample robustness, cumulative payroll growth

	$h = 1$	$h = 2$	$h = 4$
<i>1970Q1–2024Q4</i>			
β_h (high-capital slope)	-0.086 (0.079)	-0.126 (0.160)	-0.065 (0.334)
κ_h (low-capital interaction)	0.692 (0.428)	0.960* (0.553)	1.386 (0.895)
$\beta_h + \kappa_h$ (low-capital slope)	0.606 (0.424)	0.833 (0.526)	1.322 (0.826)
Low-state obs.	116	116	116
Adj. R^2	0.029	0.021	0.072
N	219	218	216
<i>1997Q1–2024Q4</i>			
β_h (high-capital slope)	0.066 (0.154)	0.212 (0.316)	0.550 (0.600)
κ_h (low-capital interaction)	1.715* (0.912)	1.946** (0.797)	2.549** (1.287)
$\beta_h + \kappa_h$ (low-capital slope)	1.781* (0.980)	2.158** (0.888)	3.099** (1.287)
Low-state obs.	43	43	43
Adj. R^2	0.134	0.092	0.119
N	111	110	108
<i>Excl. 2007Q4–2009Q1</i>			
β_h (high-capital slope)	0.073 (0.171)	0.197 (0.357)	0.405 (0.702)
κ_h (low-capital interaction)	1.867 (1.893)	1.045 (1.268)	0.504 (1.197)
$\beta_h + \kappa_h$ (low-capital slope)	1.940 (1.985)	1.242 (1.491)	0.909 (1.425)
Low-state obs.	37	37	37
Adj. R^2	0.103	0.028	0.012
N	112	110	106
<i>Excl. 2020Q1–2021Q4</i>			
β_h (high-capital slope)	0.014 (0.020)	0.050 (0.054)	0.204 (0.184)
κ_h (low-capital interaction)	0.184* (0.095)	0.352** (0.168)	0.720* (0.397)
$\beta_h + \kappa_h$ (low-capital slope)	0.198** (0.097)	0.402** (0.181)	0.924** (0.455)
Low-state obs.	39	39	39
Adj. R^2	0.813	0.801	0.720
N	110	108	104

Notes: Each block reports the h -step predictive specification from Test 1B. For excluded periods, rows are omitted when date t or any forecasted quarter used in the dependent variable intersects the excluded window. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Table BO.11: Test 1B – sample robustness, cumulative industrial production growth

	$h = 1$	$h = 2$	$h = 4$
<i>1970Q1–2024Q4</i>			
β_h (high-capital slope)	0.078 (0.153)	0.211 (0.324)	0.922 (0.617)
κ_h (low-capital interaction)	1.314** (0.586)	1.922** (0.871)	1.713 (1.409)
$\beta_h + \kappa_h$ (low-capital slope)	1.392** (0.573)	2.133*** (0.789)	2.635** (1.245)
Low-state obs.	116	116	116
Adj. R^2	0.218	0.244	0.310
N	219	218	216
<i>1997Q1–2024Q4</i>			
β_h (high-capital slope)	0.428* (0.251)	0.958** (0.479)	2.135*** (0.791)
κ_h (low-capital interaction)	2.604** (1.075)	2.609** (1.301)	1.733 (2.102)
$\beta_h + \kappa_h$ (low-capital slope)	3.031*** (1.152)	3.567*** (1.326)	3.867* (2.156)
Low-state obs.	43	43	43
Adj. R^2	0.391	0.356	0.394
N	111	110	108
<i>Excl. 2007Q4–2009Q1</i>			
β_h (high-capital slope)	0.485* (0.254)	0.982* (0.526)	1.787** (0.841)
κ_h (low-capital interaction)	1.940 (1.936)	0.390 (0.971)	-0.730 (1.749)
$\beta_h + \kappa_h$ (low-capital slope)	2.425 (2.023)	1.372 (1.089)	1.057 (1.457)
Low-state obs.	37	37	37
Adj. R^2	0.286	0.243	0.276
N	112	110	106
<i>Excl. 2020Q1–2021Q4</i>			
β_h (high-capital slope)	0.185 (0.116)	0.490* (0.286)	1.631** (0.784)
κ_h (low-capital interaction)	0.785* (0.428)	1.279 (0.925)	1.397 (2.007)
$\beta_h + \kappa_h$ (low-capital slope)	0.970** (0.441)	1.769* (0.971)	3.028 (2.276)
Low-state obs.	39	39	39
Adj. R^2	0.585	0.522	0.387
N	110	108	104

Notes: Each block reports the h -step predictive specification from Test 1B. For excluded periods, rows are omitted when date t or any forecasted quarter used in the dependent variable intersects the excluded window. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Second, we add broader macro-financial controls sequentially: the federal funds rate, the term spread, and then the full set that combines the federal funds rate, term spread, and credit spread. These specifications ask a more demanding question than the baseline: whether intermediary capital has incremental predictive content after conditioning on variables that may themselves be transmission channels or endogenous responses to intermediary distress. In those specifications, coefficients generally retain the predicted sign but lose precision, especially when the full set is included. We therefore treat the wide-control results as a demanding robustness check rather than the preferred specification.

Table BO.12: Test 1B – control robustness, output gap

	$h = 1$	$h = 2$	$h = 4$
<i>Baseline + Fed Funds</i>			
β_h (high-capital slope)	0.017 (0.126)	0.068 (0.216)	0.282 (0.356)
κ_h (low-capital interaction)	1.620** (0.663)	1.779*** (0.511)	2.102*** (0.673)
$\beta_h + \kappa_h$ (low-capital slope)	1.636** (0.679)	1.847*** (0.501)	2.384*** (0.688)
Low-state obs.	43	43	43
Adj. R^2	0.744	0.569	0.375
N	119	118	116
<i>Baseline + term spread</i>			
β_h (high-capital slope)	0.154 (0.146)	0.285 (0.284)	0.412 (0.405)
κ_h (low-capital interaction)	1.584** (0.623)	1.713*** (0.376)	2.180*** (0.637)
$\beta_h + \kappa_h$ (low-capital slope)	1.738*** (0.671)	1.998*** (0.373)	2.592*** (0.595)
Low-state obs.	43	43	43
Adj. R^2	0.757	0.597	0.415
N	119	118	116
<i>Full set</i>			
β_h (high-capital slope)	0.176 (0.145)	0.266 (0.241)	0.528 (0.392)
κ_h (low-capital interaction)	1.562 (1.028)	1.241*** (0.394)	1.565** (0.749)
$\beta_h + \kappa_h$ (low-capital slope)	1.739 (1.068)	1.507*** (0.439)	2.093*** (0.812)
Low-state obs.	43	43	43
Adj. R^2	0.753	0.596	0.418
N	119	118	116

Notes: Each block adds the listed macro-financial control to the baseline Test 1B control vector. The full set adds the federal funds rate, term spread, and credit spread. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Table BO.13: Test 1B – control robustness, negative unemployment gap

	$h = 1$	$h = 2$	$h = 4$
<i>Baseline + Fed Funds</i>			
β_h (high-capital slope)	0.080 (0.066)	0.123 (0.096)	0.387* (0.221)
κ_h (low-capital interaction)	1.591** (0.764)	1.829*** (0.403)	2.159*** (0.511)
$\beta_h + \kappa_h$ (low-capital slope)	1.671** (0.784)	1.952*** (0.418)	2.546*** (0.499)
Low-state obs.	43	43	43
Adj. R^2	0.786	0.641	0.500
N	119	118	116
<i>Baseline + term spread</i>			
β_h (high-capital slope)	0.123 (0.089)	0.195 (0.151)	0.351 (0.253)
κ_h (low-capital interaction)	1.600** (0.750)	1.839*** (0.332)	2.326*** (0.493)
$\beta_h + \kappa_h$ (low-capital slope)	1.722** (0.799)	2.034*** (0.393)	2.677*** (0.438)
Low-state obs.	43	43	43
Adj. R^2	0.793	0.657	0.515
N	119	118	116
<i>Full set</i>			
β_h (high-capital slope)	0.244** (0.099)	0.278*** (0.096)	0.543*** (0.204)
κ_h (low-capital interaction)	2.186 (1.414)	1.791** (0.763)	1.596*** (0.471)
$\beta_h + \kappa_h$ (low-capital slope)	2.430 (1.481)	2.068*** (0.792)	2.140*** (0.514)
Low-state obs.	43	43	43
Adj. R^2	0.799	0.653	0.532
N	119	118	116

Notes: Each block adds the listed macro-financial control to the baseline Test 1B control vector. The full set adds the federal funds rate, term spread, and credit spread. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Table BO.14: Test 1B – control robustness, cumulative payroll growth

	$h = 1$	$h = 2$	$h = 4$
<i>Baseline + Fed Funds</i>			
β_h (high-capital slope)	0.086 (0.131)	0.195 (0.237)	0.607 (0.509)
κ_h (low-capital interaction)	1.682* (0.944)	2.007** (0.852)	2.578** (1.213)
$\beta_h + \kappa_h$ (low-capital slope)	1.769* (0.986)	2.202** (0.914)	3.185** (1.271)
Low-state obs.	43	43	43
Adj. R^2	0.132	0.092	0.123
N	119	118	116
<i>Baseline + term spread</i>			
β_h (high-capital slope)	0.112 (0.160)	0.310 (0.339)	0.682 (0.621)
κ_h (low-capital interaction)	1.730* (0.899)	2.027*** (0.720)	2.758** (1.145)
$\beta_h + \kappa_h$ (low-capital slope)	1.842* (0.961)	2.337*** (0.837)	3.440*** (1.173)
Low-state obs.	43	43	43
Adj. R^2	0.136	0.109	0.141
N	119	118	116
<i>Full set</i>			
β_h (high-capital slope)	0.173 (0.151)	0.283 (0.244)	0.714 (0.501)
κ_h (low-capital interaction)	1.570 (1.680)	0.986 (1.173)	0.550 (1.183)
$\beta_h + \kappa_h$ (low-capital slope)	1.743 (1.747)	1.269 (1.261)	1.263 (1.388)
Low-state obs.	43	43	43
Adj. R^2	0.125	0.123	0.198
N	119	118	116

Notes: Each block adds the listed macro-financial control to the baseline Test 1B control vector. The full set adds the federal funds rate, term spread, and credit spread. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Table BO.15: Test 1B – control robustness, cumulative industrial production growth

	$h = 1$	$h = 2$	$h = 4$
<i>Baseline + Fed Funds</i>			
β_h (high-capital slope)	0.315 (0.236)	0.666 (0.411)	1.772** (0.785)
κ_h (low-capital interaction)	2.810** (1.099)	3.176** (1.265)	2.764 (1.993)
$\beta_h + \kappa_h$ (low-capital slope)	3.125*** (1.142)	3.843*** (1.295)	4.536** (2.128)
Low-state obs.	43	43	43
Adj. R^2	0.406	0.389	0.393
N	119	118	116
<i>Baseline + term spread</i>			
β_h (high-capital slope)	0.553** (0.274)	1.158** (0.543)	2.375** (0.926)
κ_h (low-capital interaction)	2.657*** (1.009)	2.823*** (1.031)	2.352 (1.958)
$\beta_h + \kappa_h$ (low-capital slope)	3.210*** (1.093)	3.981*** (1.098)	4.726** (2.020)
Low-state obs.	43	43	43
Adj. R^2	0.418	0.396	0.400
N	119	118	116
<i>Full set</i>			
β_h (high-capital slope)	0.448* (0.258)	0.819* (0.490)	2.014* (1.067)
κ_h (low-capital interaction)	2.331 (1.718)	1.795 (1.208)	1.697 (2.503)
$\beta_h + \kappa_h$ (low-capital slope)	2.778 (1.781)	2.614* (1.346)	3.710 (3.017)
Low-state obs.	43	43	43
Adj. R^2	0.413	0.413	0.399
N	119	118	116

Notes: Each block adds the listed macro-financial control to the baseline Test 1B control vector. The full set adds the federal funds rate, term spread, and credit spread. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Finally, we report moving-block residual bootstrap inference for the main Test 1B specification. This check complements the HAC standard errors by resampling blocks of residuals, rather than relying only on asymptotic covariance formulas, and is useful because the predictive regressions use overlapping future outcomes and a relatively small number of low-capital observations. The bootstrap inference broadly confirms the HAC-based inference for the low-capital interaction and the low-capital slope in the output-gap, negative-unemployment-gap, and payroll-growth specifications. For industrial production, the low-capital slope remains positive, but the one-year interaction is less precisely estimated under the bootstrap. Thus, the bootstrap checks support the qualitative dynamic evidence but also reinforce that Test 1B should not be read as a fully independent structural impulse-response exercise.

Table BO.16: Test 1B – bootstrap inference, output gap

	$h = 1$	$h = 2$	$h = 4$
β_h (high-capital slope)	0.074	0.171	0.306
Bootstrap p -value	0.541	0.352	0.250
95% CI	[-0.179, 0.322]	[-0.166, 0.537]	[-0.248, 0.832]
κ_h (low-capital interaction)	1.536	1.632	2.067
Bootstrap p -value	<0.001	<0.001	0.006
95% CI	[0.837, 2.303]	[0.641, 2.785]	[0.534, 3.547]
$\beta_h + \kappa_h$ (low-capital slope)	1.610	1.803	2.373
Bootstrap p -value	<0.001	<0.001	<0.001
95% CI	[0.963, 2.281]	[0.905, 2.848]	[1.089, 3.820]
N	119	118	116

Notes: Moving-block residual bootstrap for the main 1995Q1–2024Q4 Test 1B specification, using 999 replications and seed 42. Block length is $L_h = \max(4, h)$. The table reports point estimates, percentile-bootstrap p -values, and 95% percentile confidence intervals.

Table BO.17: Test 1B – bootstrap inference, negative unemployment gap

	$h = 1$	$h = 2$	$h = 4$
β_h (high-capital slope)	0.094	0.156	0.319
Bootstrap p -value	0.316	0.250	0.160
95% CI	[-0.117, 0.283]	[-0.183, 0.440]	[-0.134, 0.731]
κ_h (low-capital interaction)	1.573	1.788	2.248
Bootstrap p -value	<0.001	<0.001	<0.001
95% CI	[1.040, 2.369]	[0.971, 2.826]	[1.095, 3.507]
$\beta_h + \kappa_h$ (low-capital slope)	1.668	1.944	2.567
Bootstrap p -value	<0.001	<0.001	<0.001
95% CI	[1.155, 2.440]	[1.250, 2.943]	[1.564, 3.756]
N	119	118	116

Notes: Moving-block residual bootstrap for the main 1995Q1–2024Q4 Test 1B specification, using 999 replications and seed 42. Block length is $L_h = \max(4, h)$. The table reports point estimates, percentile-bootstrap p -values, and 95% percentile confidence intervals.

Table BO.18: Test 1B – bootstrap inference, cumulative payroll growth

	$h = 1$	$h = 2$	$h = 4$
β_h (high-capital slope)	0.071	0.219	0.561
Bootstrap p -value	0.601	0.382	0.232
95% CI	[-0.317, 0.394]	[-0.371, 0.745]	[-0.349, 1.466]
κ_h (low-capital interaction)	1.707	1.969	2.650
Bootstrap p -value	<0.001	0.010	0.020
95% CI	[0.809, 2.733]	[0.544, 3.584]	[0.380, 5.322]
$\beta_h + \kappa_h$ (low-capital slope)	1.778	2.188	3.211
Bootstrap p -value	<0.001	<0.001	0.002
95% CI	[0.961, 2.776]	[0.937, 3.668]	[1.321, 5.733]
N	119	118	116

Notes: Moving-block residual bootstrap for the main 1995Q1–2024Q4 Test 1B specification, using 999 replications and seed 42. Block length is $L_h = \max(4, h)$. The table reports point estimates, percentile-bootstrap p -values, and 95% percentile confidence intervals.

Table BO.19: Test 1B – bootstrap inference, cumulative industrial production growth

	$h = 1$	$h = 2$	$h = 4$
β_h (high-capital slope)	0.449	0.979	2.177
Bootstrap p -value	0.054	0.008	<0.001
95% CI	[-0.006, 0.854]	[0.216, 1.649]	[0.846, 3.510]
κ_h (low-capital interaction)	2.631	2.762	2.222
Bootstrap p -value	<0.001	0.006	0.254
95% CI	[1.345, 3.987]	[0.785, 5.059]	[-1.332, 6.053]
$\beta_h + \kappa_h$ (low-capital slope)	3.080	3.742	4.399
Bootstrap p -value	<0.001	<0.001	0.004
95% CI	[1.955, 4.342]	[2.028, 5.708]	[1.306, 7.898]
N	119	118	116

Notes: Moving-block residual bootstrap for the main 1995Q1–2024Q4 Test 1B specification, using 999 replications and seed 42. Block length is $L_h = \max(4, h)$. The table reports point estimates, percentile-bootstrap p -values, and 95% percentile confidence intervals.

BO.15.2 Test 2: Dealer Intermediation and Central-Bank Substitution

Test 2A This section reports robustness checks for Test 2A, which asks whether intermediary capital predicts future dealer financing in Treasury and agency MBS markets. The baseline specification uses the available quarterly dealer-financing sample, and a parsimonious set of controls.

First, we estimate the Test 2A specification in exclusion samples, excluding 2007Q4–2009Q1 and 2020Q1–2021Q4. The Treasury financing results are robust across these exclusions: the implied low-capital slope remains positive and statistically significant at all reported horizons. The agency MBS financing results remain weaker. In the MBS specifications, the high-capital slope stays positive, but the implied low-capital slope is small, unstable in sign, and generally insignificant. These sample checks therefore preserve the main distinction from the baseline:

Test 2A is most supportive for Treasury dealer financing, not for agency MBS dealer financing.

Table BO.20: Test 2A – sample robustness, treasury dealer financing

	$h = 1$	$h = 2$	$h = 4$
<i>Excl. 2007Q4–2009Q1</i>			
β_h (high-capital slope)	0.487 (0.582)	0.716 (1.078)	0.403 (1.112)
κ_h (low-capital interaction)	0.629 (0.772)	0.158 (1.419)	1.753 (1.370)
$\beta_h + \kappa_h$ (low-capital slope)	1.116* (0.573)	0.874* (0.471)	2.156** (0.918)
Low-state obs.	37	37	35
Adj. R^2	0.984	0.979	0.964
N	71	68	65
<i>Excl. 2020Q1–2021Q4</i>			
β_h (high-capital slope)	0.982 (0.841)	2.126 (1.565)	6.204** (3.026)
κ_h (low-capital interaction)	0.618 (1.244)	0.873 (2.265)	-1.945 (3.273)
$\beta_h + \kappa_h$ (low-capital slope)	1.600** (0.812)	2.999** (1.498)	4.258*** (1.383)
Low-state obs.	39	39	37
Adj. R^2	0.966	0.935	0.883
N	69	66	63

Notes: Each block reports the h -step predictive specification from Test 2A. Rows are omitted when date t or the forecasted quarter $t + h$ intersects the excluded window. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Table BO.21: Test 2A – sample robustness, agency MBS dealer financing

	$h = 1$	$h = 2$	$h = 4$
<i>Excl. 2007Q4–2009Q1</i>			
β_h (high-capital slope)	0.181** (0.073)	0.329*** (0.119)	0.744*** (0.233)
κ_h (low-capital interaction)	-0.337** (0.168)	-0.575** (0.225)	-1.036*** (0.374)
$\beta_h + \kappa_h$ (low-capital slope)	-0.156 (0.114)	-0.245* (0.143)	-0.292 (0.191)
Low-state obs.	37	37	35
Adj. R^2	0.963	0.929	0.868
N	71	68	65
<i>Excl. 2020Q1–2021Q4</i>			
β_h (high-capital slope)	0.185** (0.073)	0.313** (0.123)	0.932*** (0.225)
κ_h (low-capital interaction)	-0.188 (0.181)	-0.112 (0.297)	-0.862** (0.351)
$\beta_h + \kappa_h$ (low-capital slope)	-0.003 (0.134)	0.201 (0.231)	0.070 (0.217)
Low-state obs.	39	39	37
Adj. R^2	0.949	0.907	0.834
N	69	66	63

Notes: Each block reports the h -step predictive specification from Test 2A. Rows are omitted when date t or the forecasted quarter $t + h$ intersects the excluded window. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Second, we add broader macro-financial controls sequentially: the federal funds rate, the term spread, and then the full set that combines the federal funds rate, term spread, and credit spread. The Treasury financing low-capital slope remains positive and statistically significant across the wider-control specifications. The agency MBS results again do not deliver a robust positive low-capital slope: the high-capital slope is positive, but the low-capital interaction is often negative and the implied low-capital slope is close to zero or negative. We therefore treat the wider-control results as reinforcing the Treasury evidence while confirming that the MBS financing evidence is not robust.

Table BO.22: Test 2A – control robustness, treasury dealer financing

	$h = 1$	$h = 2$	$h = 4$
<i>Baseline + Fed Funds</i>			
β_h (high-capital slope)	1.228 (0.927)	2.425 (1.634)	6.136* (3.143)
κ_h (low-capital interaction)	0.292 (1.169)	0.480 (2.024)	-2.658 (3.266)
$\beta_h + \kappa_h$ (low-capital slope)	1.519** (0.649)	2.904** (1.202)	3.478*** (1.174)
Low-state obs.	43	43	41
Adj. R^2	0.968	0.945	0.890
N	78	76	75
<i>Baseline + term spread</i>			
β_h (high-capital slope)	1.130 (0.893)	2.162 (1.567)	6.000** (3.035)
κ_h (low-capital interaction)	0.469 (1.144)	0.985 (2.153)	-2.298 (3.146)
$\beta_h + \kappa_h$ (low-capital slope)	1.599** (0.671)	3.147** (1.351)	3.701*** (1.211)
Low-state obs.	43	43	41
Adj. R^2	0.968	0.942	0.890
N	78	76	75
<i>Full set</i>			
β_h (high-capital slope)	1.205 (0.885)	2.338 (1.556)	5.980** (2.921)
κ_h (low-capital interaction)	0.652 (1.302)	1.379 (2.601)	-0.622 (3.818)
$\beta_h + \kappa_h$ (low-capital slope)	1.857* (1.008)	3.717* (2.106)	5.359** (2.479)
Low-state obs.	43	43	41
Adj. R^2	0.967	0.943	0.890
N	78	76	75

Notes: Each block adds the listed macro-financial control to the baseline Test 2A control vector. The full set adds the federal funds rate, term spread, and credit spread. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Table BO.23: Test 2A – control robustness, agency MBS dealer financing

	$h = 1$	$h = 2$	$h = 4$
<i>Baseline + Fed Funds</i>			
β_h (high-capital slope)	0.182*** (0.067)	0.281*** (0.097)	0.846*** (0.212)
κ_h (low-capital interaction)	-0.165 (0.161)	-0.106 (0.257)	-0.850*** (0.306)
$\beta_h + \kappa_h$ (low-capital slope)	0.017 (0.117)	0.175 (0.199)	-0.004 (0.182)
Low-state obs.	43	43	41
Adj. R^2	0.957	0.929	0.861
N	78	76	75
<i>Baseline + term spread</i>			
β_h (high-capital slope)	0.185*** (0.070)	0.307*** (0.112)	0.893*** (0.224)
κ_h (low-capital interaction)	-0.168 (0.159)	-0.128 (0.254)	-0.926*** (0.342)
$\beta_h + \kappa_h$ (low-capital slope)	0.018 (0.117)	0.179 (0.194)	-0.033 (0.226)
Low-state obs.	43	43	41
Adj. R^2	0.956	0.922	0.859
N	78	76	75
<i>Full set</i>			
β_h (high-capital slope)	0.164** (0.069)	0.255** (0.110)	0.770*** (0.243)
κ_h (low-capital interaction)	-0.371** (0.166)	-0.402* (0.242)	-1.126*** (0.435)
$\beta_h + \kappa_h$ (low-capital slope)	-0.207 (0.135)	-0.147 (0.202)	-0.355 (0.361)
Low-state obs.	43	43	41
Adj. R^2	0.958	0.933	0.877
N	78	76	75

Notes: Each block adds the listed macro-financial control to the baseline Test 2A control vector. The full set adds the federal funds rate, term spread, and credit spread. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Finally, we report moving-block residual bootstrap inference for the main Test 2A specification. The bootstrap inference supports the Treasury low-capital slope at all reported horizons, while the low-capital interaction itself remains imprecisely estimated. For agency MBS financing, the bootstrap confirms that the high-capital slope is positive, but it does not support a positive low-capital slope; at the four-quarter horizon the low-capital interaction is negative. The bootstrap checks therefore support the qualitative interpretation of the baseline tables: the private balance-sheet absorption channel is clearer for Treasury dealer financing than for agency MBS dealer financing.

Table BO.24: Test 2A – bootstrap inference, treasury dealer financing

	$h = 1$	$h = 2$	$h = 4$
β_h (high-capital slope)	1.093	2.106	6.094
Bootstrap p -value	0.142	0.056	0.004
95% CI	[-0.455, 2.399]	[-0.071, 4.285]	[2.775, 9.452]
κ_h (low-capital interaction)	0.418	0.790	-2.610
Bootstrap p -value	0.697	0.667	0.350
95% CI	[-1.821, 2.862]	[-2.764, 4.348]	[-7.568, 2.545]
$\beta_h + \kappa_h$ (low-capital slope)	1.511	2.896	3.484
Bootstrap p -value	0.030	0.008	0.018
95% CI	[0.197, 3.085]	[0.891, 5.184]	[0.548, 6.778]
N	78	76	75

Notes: Moving-block residual bootstrap for the main Test 2A specification, using 999 replications and seed 42. Block length is $L_h = \max(4, h)$. The table reports point estimates, percentile-bootstrap p -values, and 95% percentile confidence intervals.

Table BO.25: Test 2A – bootstrap inference, agency MBS dealer financing

	$h = 1$	$h = 2$	$h = 4$
β_h (high-capital slope)	0.184	0.300	0.892
Bootstrap p -value	0.020	0.048	0.002
95% CI	[0.025, 0.354]	[0.011, 0.603]	[0.435, 1.336]
κ_h (low-capital interaction)	-0.167	-0.129	-0.894
Bootstrap p -value	0.250	0.619	0.018
95% CI	[-0.457, 0.116]	[-0.613, 0.361]	[-1.658, -0.110]
$\beta_h + \kappa_h$ (low-capital slope)	0.016	0.171	-0.002
Bootstrap p -value	0.889	0.244	0.983
95% CI	[-0.175, 0.192]	[-0.122, 0.448]	[-0.438, 0.491]
N	78	76	75

Notes: Moving-block residual bootstrap for the main Test 2A specification, using 999 replications and seed 42. Block length is $L_h = \max(4, h)$. The table reports point estimates, percentile-bootstrap p -values, and 95% percentile confidence intervals.

Test 2B This section reports robustness checks for Test 2B, which asks whether central-bank risk-bearing rises when dealer intermediation contracts and when the intermediary-capital shortfall is larger. The baseline specification uses the post-2012 dealer gross-position sample, and the same parsimonious controls as Test 2A.

First, we estimate the Test 2B specification in an exclusion sample that drops 2020Q1–2021Q4. Because dealer gross positions are available only in the post-2012 sample, this is the relevant sample-exclusion check for the baseline gross-position specification; the GFC transition window is not observed in these data. Excluding COVID strengthens the gross-position evidence. The shortfall coefficient ϕ remains positive in both markets, and the dealer-intermediation coefficient β is negative and statistically significant for both Treasury and agency MBS gross positions. This pattern is consistent with the substitution interpretation, while also showing that the baseline estimates are sensitive to the COVID period.

Table BO.26: Test 2B – sample robustness, treasury and agency MBS dealer gross position

	Treasury gross positions	Agency MBS gross positions
<i>Excl. 2020Q1–2021Q4</i>		
ϕ (capital shortfall)	0.632* (0.356)	0.612* (0.315)
β (dealer intermediation)	-0.889*** (0.326)	-0.412* (0.222)
Adj. R^2	0.686	0.792
N	40	40

Notes: The table reports the baseline gross-position Test 2B specification after excluding 2020Q1–2021Q4. The dependent variable is the quarterly change in Fed risk-bearing to GDP. Dealer intermediation, capital shortfall, and controls are lagged one quarter. HAC standard errors with four lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Second, we add broader macro-financial controls sequentially: the federal funds rate, the term spread, and then the full set that combines the federal funds rate, term spread, and credit spread. The shortfall coefficient ϕ remains positive and statistically significant in both markets across these wider-control specifications. The dealer-intermediation coefficient β is negative for Treasury gross positions but less precisely estimated, and it is close to zero for agency MBS gross positions. Thus, the wider-control checks support the shortfall margin more strongly than the direct dealer-substitution coefficient.

Table BO.27: Test 2B – control robustness, treasury and agency MBS dealer gross position

	Treasury gross positions	Agency MBS gross positions
<i>Baseline + Fed Funds</i>		
ϕ (capital shortfall)	6.733** (2.664)	3.939** (1.943)
β (dealer intermediation)	-2.087 (1.419)	0.028 (0.595)
Adj. R^2	0.598	0.360
N	48	48
<i>Baseline + term spread</i>		
ϕ (capital shortfall)	7.272** (2.847)	4.195** (2.027)
β (dealer intermediation)	-1.951 (1.284)	-0.001 (0.552)
Adj. R^2	0.606	0.359
N	48	48
<i>Full set</i>		
ϕ (capital shortfall)	9.160*** (2.793)	5.163*** (1.948)
β (dealer intermediation)	-2.014 (1.276)	-0.104 (0.536)
Adj. R^2	0.637	0.466
N	48	48

Notes: Each block adds the listed macro-financial control to the baseline gross-position Test 2B control vector. The full set adds the federal funds rate, term spread, and credit spread. HAC standard errors with four lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Third, we report moving-block residual bootstrap inference for the baseline gross-position specifications. For compactness, the bootstrap table focuses on the two coefficients of interest, ϕ and β . The bootstrap confidence intervals for ϕ are strictly positive in both markets, confirming the evidence that central-bank risk-bearing rises when the intermediary-capital shortfall is larger. By contrast, the bootstrap confidence intervals for β include zero in both markets. The bootstrap evidence therefore supports the capital-shortfall part of Test 2B more strongly than the dealer-substitution coefficient itself.

Table BO.28: Test 2B – bootstrap inference, treasury and agency MBS dealer gross position

	Treasury gross positions	Agency MBS gross positions
ϕ (capital shortfall)	6.497	3.886
Bootstrap p -value	<0.001	<0.001
95% CI	[3.898, 9.469]	[2.264, 5.533]
β (dealer intermediation)	-2.731	-0.137
Bootstrap p -value	0.150	0.779
95% CI	[-6.372, 0.922]	[-1.179, 0.882]
N	48	48

Notes: Moving-block residual bootstrap for the baseline gross-position Test 2B specifications, using 999 replications and seed 42. Block length is four. The table reports point estimates, percentile-bootstrap p -values, and 95% percentile confidence intervals.

Finally, we replace dealer gross positions with dealer financing to obtain a longer-sample robustness check. The financing-based specification is useful because it extends the sample back to the mid-2000s, but it is not the preferred balance-sheet absorption specification. In this longer sample, the shortfall coefficient ϕ is positive and statistically significant in both Treasury and agency MBS markets. The dealer-intermediation coefficient β is negative for agency MBS financing but positive for Treasury financing, consistent with the concern that Treasury financing is contaminated by emergency market-functioning and client-demand episodes. We therefore view this table as a useful robustness check on the shortfall margin, not as a replacement for the gross-position baseline.

Table BO.29: Test 2B – regressor robustness, treasury and agency MBS dealer financing

	Treasury financing	Agency MBS financing
α	0.043* (0.025)	0.021** (0.009)
ρ	0.702*** (0.207)	0.670*** (0.124)
ϕ	4.376** (1.940)	2.305*** (0.791)
β	0.409** (0.173)	-0.694* (0.410)
Controls	Yes	Yes
Adj. R^2	0.458	0.462
N	77	77

Notes: This is a financing-based longer-sample robustness check, not the preferred balance-sheet absorption specification. The dependent variable is the quarterly change in Fed risk-bearing to GDP. Dealer financing, capital shortfall, and controls are lagged one quarter. Controls are output gap, productivity growth, and inflation. HAC standard errors with four lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

BO.15.3 Test 3: Pre-QE Intermediation Gap as a Sufficient Statistic

Test 3A This section reports robustness checks for Test 3A, which relates the exact pre-QE intermediation gap to observed slack measures. The baseline specification uses the full pre-QE quarterly sample 1970Q1–2007Q4. We consider two shorter pre-QE windows. First, we start the sample in 1995Q1, aligning the exercise with the modern quarterly window used in the baseline Test 1A specification while keeping only pre-QE observations. Second, we start in 1997Q1, aligning the quarterly test with the annual data window used to calibrate ϵ . In both cases, the sample ends in 2007Q4.

Table BO.30: Test 3A – sample robustness, output gap

	1995Q1–2007Q4		1997Q1–2007Q4	
	Linear	Quadratic	Linear	Quadratic
α	0.004 (0.003)	0.004 (0.003)	0.005 (0.004)	0.005 (0.004)
β_1	0.229 (0.148)	-0.301 (0.822)	0.256* (0.132)	0.111 (1.284)
β_2	—	18.098 (25.155)	—	4.682 (40.101)
$\beta_1 + 2\beta_2\bar{b}_{2007Q4}^{pre}$	—	0.894 (0.853)	—	0.421 (1.379)
p -value of $H_0 : \beta_1 = \beta_2 = 0$	—	0.033	—	0.126
Adj. R^2	-0.011	-0.028	-0.013	-0.037
N	52	52	44	44

Notes: The dependent variable is the output gap. The table reports pre-QE sample-start robustness checks for Test 3A. All samples end in 2007Q4. Both specifications use $\bar{b}_t^{pre} = b_t^{pre} - \bar{b}_t^{pre}$, where $b_t^{pre} = b_u(\hat{w}_t)$ and the mean is computed within the corresponding estimation sample. The quadratic columns also include $(\bar{b}_t^{pre})^2$. The marginal-effect row evaluates $\beta_1 + 2\beta_2\bar{b}_t^{pre}$ in 2007Q4 for the quadratic columns. The reported p -value tests $H_0 : \beta_1 = \beta_2 = 0$ for the quadratic columns. HAC standard errors with four lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Table BO.31: Test 3A – sample robustness, negative unemployment gap

	1995Q1–2007Q4		1997Q1–2007Q4	
	Linear	Quadratic	Linear	Quadratic
α	0.001 (0.002)	0.001 (0.002)	0.002 (0.002)	0.002 (0.002)
β_1	0.020 (0.076)	-0.153 (0.360)	0.020 (0.072)	-0.025 (0.508)
β_2	—	5.919 (10.449)	—	1.461 (15.245)
$\beta_1 + 2\beta_2 \widehat{b}_{2007Q4}^{pre}$	—	0.238 (0.338)	—	0.071 (0.508)
p -value of $H_0 : \beta_1 = \beta_2 = 0$	—	0.564	—	0.934
Adj. R^2	-0.020	-0.039	-0.024	-0.048
N	52	52	44	44

Notes: The dependent variable is the negative unemployment gap. The table reports pre-QE sample-start robustness checks for Test 3A. All samples end in 2007Q4. Both specifications use $\widehat{b}_t^{pre} = b_t^{pre} - b_t^{pre}$, where $b_t^{pre} = b_u(\hat{\omega}_t)$ and the mean is computed within the corresponding estimation sample. The quadratic columns also include $(\widehat{b}_t^{pre})^2$. The marginal-effect row evaluates $\beta_1 + 2\beta_2 \widehat{b}_t^{pre}$ in 2007Q4 for the quadratic columns. The reported p -value tests $H_0 : \beta_1 = \beta_2 = 0$ for the quadratic columns. HAC standard errors with four lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

The shorter-window checks qualify the baseline evidence. For the output gap, the 1995Q1–2007Q4 and 1997Q1–2007Q4 linear specifications do not reproduce the negative baseline relation; the estimated slopes are positive, and the adjusted R^2 values are close to zero or negative. The quadratic specifications are also imprecise. Although the 1995Q1 quadratic output-gap specification rejects the joint null, the implied 2007Q4 marginal effect is positive rather than negative, so this rejection does not support the model’s sign prediction. For the sign-reversed unemployment gap, the shorter-window estimates are small, imprecise, and the joint tests do not reject.

We therefore interpret Test 3A as evidence about the low-frequency pre-QE mapping over the full pre-QE state space, not as a relation that can be precisely recovered from the short late-pre-crisis window alone. This is consistent with the purpose of the test: the exact pre-QE gap is a calibrated level object, and the shorter windows remove much of the variation in intermediary capital and in the exact gap that identifies the full-sample relation. The robustness checks are therefore useful qualifications rather than alternative preferred specifications.

Test 3B This section reports robustness checks for Test 3B, which asks whether the exact pre-QE gap proxy has incremental predictive content beyond raw intermediary capital. The baseline specification uses the whole 1970Q1–2024Q4 sample, omitting the 2008Q1–2009Q1 transition quarters.

First, we estimate the Test 3B specification in alternative sample windows and exclusion samples: the 1995Q1–2024Q4 sample, the 1997Q1–2024Q4 sample, the whole 1970Q1–2024Q4 sample excluding 2007Q4–2009Q1, and the whole 1970Q1–2024Q4 sample excluding 2020Q1–2021Q4. The regime difference coefficient remains mostly statistically insignificant.

Table BO.32: Test 3B – sample robustness, exact intermediation gap

	$h = 1$	$h = 2$	$h = 4$
<i>1995Q1–2024Q4</i>			
β_h^{pre} (exact gap, pre-QE)	-0.431* (0.239)	-0.589* (0.304)	-1.186*** (0.423)
β_h^{QE} (exact gap, QE-active)	-0.101 (0.104)	-0.096 (0.093)	-0.070 (0.055)
$\beta_h^{QE} - \beta_h^{pre}$ (regime difference)	0.331 (0.258)	0.493 (0.318)	1.116*** (0.426)
Wald p-value: $\beta_h^{QE} = \beta_h^{pre}$	0.200	0.120	0.009
N	113	111	107
<i>1997Q1–2024Q4</i>			
β_h^{pre} (exact gap, pre-QE)	-0.049 (0.148)	5.244*** (1.937)	4.589 (3.336)
β_h^{QE} (exact gap, QE-active)	-0.101 (0.104)	-0.096 (0.093)	-0.070 (0.055)
$\beta_h^{QE} - \beta_h^{pre}$ (regime difference)	-0.052 (0.181)	-5.340*** (1.940)	-4.658 (3.337)
Wald p-value: $\beta_h^{QE} = \beta_h^{pre}$	0.774	0.006	0.163
N	105	103	99
<i>Excl. 2007Q4–2009Q1</i>			
β_h^{pre} (exact gap, pre-QE)	-0.037*** (0.013)	-0.061** (0.025)	-0.085** (0.037)
β_h^{QE} (exact gap, QE-active)	-0.101 (0.104)	-0.096 (0.093)	-0.070 (0.055)
$\beta_h^{QE} - \beta_h^{pre}$ (regime difference)	-0.063 (0.105)	-0.035 (0.096)	0.015 (0.066)
Wald p-value: $\beta_h^{QE} = \beta_h^{pre}$	0.545	0.715	0.818
N	212	210	206
<i>Excl. 2020Q1–2021Q4</i>			
β_h^{pre} (exact gap, pre-QE)	-0.037*** (0.013)	-0.062** (0.025)	-0.085** (0.037)
β_h^{QE} (exact gap, QE-active)	0.005 (0.022)	0.011 (0.029)	-0.030 (0.029)
$\beta_h^{QE} - \beta_h^{pre}$ (regime difference)	0.042 (0.026)	0.073* (0.038)	0.055 (0.047)
Wald p-value: $\beta_h^{QE} = \beta_h^{pre}$	0.103	0.059	0.241
N	204	201	195

Notes: Each block re-estimates the fully saturated pooled Test 3B interaction regression. For excluded periods, rows are omitted when date t or the forecasted quarter $t + h$ intersects the excluded window. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Second, we report moving-block residual bootstrap inference for the main Test 3B specification. The bootstrap confirms the main pre-QE result. The bootstrap confidence intervals for

β_h^{pre} are strictly negative for all horizons, confirming the evidence that the exact gap has pre-QE incremental predictive content. By contrast, the bootstrap confidence intervals for $\beta_h^{QE} - \beta_h^{pre}$ include zero for all horizons, which confirms the lack of a statistically sharp regime-difference result.

Table BO.33: Test 3B – bootstrap inference, exact intermediation gap

	$h = 1$	$h = 2$	$h = 4$
β_h^{pre} (exact gap, pre-QE)	-0.037	-0.062	-0.085
Bootstrap p -value	0.010	0.010	0.008
95% CI	[-0.066, -0.009]	[-0.107, -0.017]	[-0.156, -0.018]
β_h^{QE} (exact gap, QE-active)	-0.101	-0.096	-0.070
Bootstrap p -value	0.048	0.196	0.553
95% CI	[-0.190, -0.000]	[-0.246, 0.053]	[-0.320, 0.154]
$\beta_h^{QE} - \beta_h^{pre}$ (regime difference)	-0.063	-0.035	0.015
Bootstrap p -value	0.250	0.657	0.895
95% CI	[-0.159, 0.039]	[-0.193, 0.128]	[-0.255, 0.236]
N	213	211	207

Notes: Moving-block residual bootstrap for the main pooled Test 3B specification, using 999 replications and seed 42. Block length is $L_h = \max(4, h)$. The table reports point estimates, percentile-bootstrap p -values, and 95% percentile confidence intervals.

Finally, we report a CAPE benchmark comparison. This is not a control robustness check; it asks whether the exact intermediation-gap proxy is merely repackaging a familiar stock-market valuation predictor. We therefore estimate a pre-QE horse race between $\hat{b}_t^{pre}$ and the cyclically adjusted price-earnings ratio, always controlling for the current output gap. The CAPE horse race supports the distinctiveness of the exact gap proxy. The coefficient on $\hat{b}_t^{pre}$ remains negative and statistically significant at all horizons, with estimates of -0.037 , -0.063 , and -0.092 for $h = 1, 2, 4$, while CAPE is economically small and statistically insignificant. This suggests that the pre-QE exact gap is not simply a standard valuation-ratio predictor.

Table BO.34: Test 3B – horse race, cyclically adjusted price-earnings ratio

	$h = 1$	$h = 2$	$h = 4$
β_h (exact gap)	-0.037*** (0.013)	-0.063*** (0.023)	-0.092*** (0.035)
f_h (CAPE)	-0.000 (0.000)	-0.000 (0.000)	-0.000 (0.000)
ρ_h (current output gap)	0.885*** (0.042)	0.737*** (0.076)	0.391*** (0.122)
Adj. R^2	0.860	0.693	0.381
N	151	150	148

Notes: This is a benchmark-predictor comparison, not a baseline control specification. The dependent variable is the output gap h quarters ahead in the pre-QE sample ending in 2007Q4. The specification includes the exact pre-QE gap proxy, Shiller P/E as the CAPE measure, and the current output gap. HAC standard errors with $\max(h, 4)$ lags are reported in parentheses. ***, **, * denote significance at 1%, 5%, 10%.

Appendix C

The Econometrics of Portfolio Sorts

Nicolo Ceneda, Valentina Corradi and Walter Distaso

C.1 Proofs

C.1.1 Theorem 1

C.1.1.1 Part (i)

Rewriting the test statistic in Definition 3.1:

$$\begin{aligned} Z_{T,N} &= \frac{N}{\sqrt{T-\mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq S^{q(1)} \right\} - \mathbb{1} \left\{ S^{q(N-1)} < S_{i,\tau(t)} \right\} \right) \\ &\quad + \frac{N}{\sqrt{T-\mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(1)} \right\} - \mathbb{1} \left\{ S_{i,\tau(t)} \leq S^{q(1)} \right\} \right) \\ &\quad - \frac{N}{\sqrt{T-\mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ S_{\tau(t)}^{p(N-1)} < S_{i,\tau(t)} \right\} - \mathbb{1} \left\{ S^{q(N-1)} < S_{i,\tau(t)} \right\} \right) \\ &= I_{T,N} + II_{T,N} + III_{T,N}. \end{aligned}$$

First, under the null hypothesis in equation (3.1) and given assumptions 3.1.(i)–(iii), by the central limit theorem (see Rio et al. (2017), Theorem 4.2):

$$\begin{aligned} I_{T,N} &= \frac{N}{\sqrt{T-\mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(r_{i,t} \mathbb{1} \left\{ S_{i,\tau(t)} \leq S^{q(1)} \right\} - \mu^{q(1)} \right) \\ &\quad - \frac{N}{\sqrt{T-\mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(r_{i,t} \mathbb{1} \left\{ S^{q(N-1)} < S_{i,\tau(t)} \right\} - \mu^{q(N)} \right) \end{aligned}$$

$$\xrightarrow{d} N(0, V).$$

Second, rewriting $II_{T,N}$:

$$\begin{aligned} II_{T,N} &= \frac{\mathcal{N}}{\sqrt{T-\mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(1)} \right\} - F \left(S_{\tau(t)}^{p(1)} \right) \right) \right. \\ &\quad \left. - \left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq S^{q(1)} \right\} - F \left(S^{q(1)} \right) \right) \right) \\ &\quad + \frac{\mathcal{N}}{\sqrt{T-\mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(F \left(S_{\tau(t)}^{p(1)} \right) - F \left(S^{q(1)} \right) \right) \\ &= II_{T,N}^A + II_{T,N}^B, \end{aligned}$$

where $F(x)$ denotes the CDF of $S_{i,\tau(t)}$ evaluated at x . First, $II_{T,N}^A = o_p(1)$ because of stochastic equicontinuity. Moreover, let $f(x)$ denote the PDF of $S_{i,\tau(t)}$ evaluated at x . By a first-order expansion around $S^{q(1)}$ and given assumption 3.1(iv):

$$\begin{aligned} II_{T,N}^B &= \frac{\mathcal{N}}{\sqrt{T-\mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} f \left(S^{q(1)} \right) \left(S_{\tau(t)}^{p(1)} - S^{q(1)} \right) + o_p(1) \\ &= \frac{\mathcal{N}}{\sqrt{T-\mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \mathbb{E} \left(r_{i,t} f \left(S^{q(1)} \right) \right) \left(S_{\tau(t)}^{p(1)} - S^{q(1)} \right) \\ &\quad + \frac{\mathcal{N}}{\sqrt{T-\mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(r_{i,t} f \left(S^{q(1)} \right) - \mathbb{E} \left(r_{i,t} f \left(S^{q(1)} \right) \right) \right) \left(S_{\tau(t)}^{p(1)} - S^{q(1)} \right) + o_p(1) \\ &= \mathbb{E} \left(r_{i,t} f \left(S^{q(1)} \right) \right) \frac{\mathcal{N}}{\sqrt{T-\mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \left(S_{\tau(t)}^{p(1)} - S^{q(1)} \right) \\ &\quad + \frac{\mathcal{N}}{\sqrt{T-\mathcal{T}}} \sum_{t=\mathcal{T}+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(r_{i,t} f \left(S^{q(1)} \right) - \mathbb{E} \left(r_{i,t} f \left(S^{q(1)} \right) \right) \right) \left(S_{\tau(t)}^{p(1)} - S^{q(1)} \right) + o_p(1) \\ &= O_p \left(N^{-(1-\rho)/2} \right) + o_p(1) \\ &= o_p(1). \end{aligned}$$

Third, using similar arguments as above, $III_{T,N} = o_p(1)$ as well.

C.1.1.2 Part (ii)

Under the alternative hypothesis in equation (3.2), $Z_{T,N}$ is the sum of $T - \mathcal{T}$ non-zero differences divided by $\sqrt{T - \mathcal{T}}$ and diverges at the rate $\sqrt{T - \mathcal{T}}$. After dividing by $\sqrt{T - \mathcal{T}}$, the

result follows.

C.1.2 Theorem 2

C.1.2.1 Part (i)

As in Theorem 3.1, $Z_{T,N} = I_{T,N} + II_{T,N} + III_{T,N}$. First, under the null hypothesis 3.1 and given Assumption 3.1.(i)–(iii), by the central limit theorem, $I_{T,N} \xrightarrow{d} N(0, V)$. Second, $II_{T,N} = II_{T,N}^A + II_{T,N}^B$. In particular, $II_{T,N}^A = o_p(1)$ because of stochastic equicontinuity. On the other hand, under Assumption 3.3 in place of 3.1.(iv), $\rho = 1$ and $II_{T,N}^B$ is no longer $o_p(1)$, since $\text{var}\left(S_{\tau(t)}^{p(1)} - S_{\tau(t)}^{q(1)}\right)$ is bounded away from zero. In fact, given Assumptions 3.1.(i) and 3.3,

$$II_{T,N}^B = \mathbb{E}\left(r_{i,t} f\left(S_{\tau(t)}^{q(1)}\right)\right) \frac{N}{\sqrt{T-R}} \sum_{t=R+1}^T \left(S_{\tau(t)}^{p(1)} - S_{\tau(t)}^{q(1)}\right)$$

satisfies a central limit theorem and the statement follows. Finally, $III_{T,N}$ can be treated analogously.

C.1.2.2 Part (ii)

The reasoning is analogous to that in Theorem 3.1.

C.1.3 Lemma 1

We derive Lemma C.1, which will be used in the derivation of Theorem 3.3 and Corollary 3.1.

Lemma C.1 *Let either Assumptions 3.1, 3.2, 3.4, 3.5 hold, or 3.1.(i)–(iii), 3.2–3.5 hold:*

$$\begin{aligned} \frac{N}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} & \left(\mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(1)} < S_{i,\tau(t)} \right\} \right. \\ & \left. - \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(N-1),U} < \widehat{S}_{i,\tau(t),R} \right\} \mathbb{1} \left\{ S_{i,\tau(t)} \leq \widehat{S}_{\tau(t),R}^{p(N-1)} \right\} \right) = o_{a.s.}(1). \end{aligned}$$

Proof of Lemma 1 Consider the first portfolio. From equation (3.3):

$$\begin{aligned} & \frac{N}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(1)} < S_{i,\tau(t)} \right\} \\ &= \frac{N}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \mathbb{1} \left\{ S_{i,\tau(t)} + \left(\widehat{S}_{i,\tau(t),R} - S_{i,\tau(t)} \right) \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(1)} < S_{i,\tau(t)} \right\} \end{aligned}$$

$$\begin{aligned}
&\leq \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \mathbb{1} \left\{ S_{i,\tau(t)} \leq \widehat{S}_{\tau(t),R}^{p(1),L} + c_{1,\tau(t),R} \right\} \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(1)} < S_{i,\tau(t)} \right\} \\
&= o_{a.s.}(1)
\end{aligned}$$

An analogous result holds for the last portfolio.

C.1.4 Theorem 3

C.1.4.1 Part (i)

Rewriting the test statistic in Definition 3.2 and using Lemma C.1:

$$\begin{aligned}
&\widehat{Z}_{T,N,R} \\
&= \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} \left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq \widehat{S}_{\tau(t),R}^{p(1)} \right\} + \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(1)} < S_{i,\tau(t)} \right\} \right) \\
&\quad - \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(N-1),U} < \widehat{S}_{i,\tau(t),R} \right\} \left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq \widehat{S}_{\tau(t),R}^{p(N-1)} \right\} + \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(N-1)} < S_{i,\tau(t)} \right\} \right) \\
&= \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} \mathbb{1} \left\{ S_{i,\tau(t)} \leq \widehat{S}_{\tau(t),R}^{p(1)} \right\} \\
&\quad - \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(N-1),U} < \widehat{S}_{i,\tau(t),R} \right\} \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(N-1)} < S_{i,\tau(t)} \right\} + o_{a.s.}(1) \\
&= \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} \mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(1)} \right\} \right. \\
&\quad \left. - \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(N-1),U} < \widehat{S}_{i,\tau(t),R} \right\} \mathbb{1} \left\{ S_{\tau(t)}^{p(N-1)} < S_{i,\tau(t)} \right\} \right) \\
&\quad + \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} \left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq \widehat{S}_{\tau(t),R}^{p(1)} \right\} - \mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(1)} \right\} \right) \\
&\quad - \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(N-1),U} < \widehat{S}_{i,\tau(t),R} \right\} \left(\mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(N-1)} < S_{i,\tau(t)} \right\} - \mathbb{1} \left\{ S_{\tau(t)}^{p(N-1)} < S_{i,\tau(t)} \right\} \right) \\
&= I_{T,N,R} + II_{T,N,R} + III_{T,N,R}.
\end{aligned}$$

Let $\widehat{u}_{\tau(t),R}^{p(1)} = \widehat{S}_{\tau(t),R}^{p(1)} - S_{\tau(t)}^{p(1)}$. Rewriting $II_{T,N,R}$:

$$\begin{aligned}
&II_{T,N,R} \\
&= \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} \left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(1)} + \widehat{u}_{\tau(t),R}^{p(1)} \right\} - \mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(1)} \right\} \right)
\end{aligned}$$

$$\begin{aligned}
&= \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} \left(\left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(1)} + \widehat{u}_{\tau(t),R}^{p(1)} \right\} - F \left(S_{\tau(t)}^{p(1)} + \widehat{u}_{\tau(t),R}^{p(1)} \right) \right) \right. \\
&\quad \left. - \left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(1)} \right\} - F \left(S_{\tau(t)}^{p(1)} \right) \right) \right) \\
&+ \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} \left(F \left(S_{\tau(t)}^{p(1)} + \widehat{u}_{\tau(t),R}^{p(1)} \right) - F \left(S_{\tau(t)}^{p(1)} \right) \right) \\
&= II_{T,N,R}^A + II_{T,N,R}^B.
\end{aligned}$$

$II_{T,N,R}^A = o_p(1)$ because of stochastic equicontinuity. Moreover, rewriting $II_{T,N,R}^B$:

$$\begin{aligned}
II_{T,N,R}^B &= \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} f \left(S_{\tau(t)}^{p(1)} + \widehat{u}_{\tau(t),R}^{p(1)} \right) \left(\widehat{S}_{\tau(t),R}^{p(1)} - S_{\tau(t)}^{p(1)} \right) \\
&= \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} E \left(r_{i,t} \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} f \left(S_{\tau(t)}^{p(1)} + \widehat{u}_{\tau(t),R}^{p(1)} \right) \right) \left(\widehat{S}_{\tau(t),R}^{p(1)} - S_{\tau(t)}^{p(1)} \right) \\
&+ \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(r_{i,t} \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} f \left(S_{\tau(t)}^{p(1)} + \widehat{u}_{\tau(t),R}^{p(1)} \right) \right. \\
&\quad \left. - E \left(r_{i,t} \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} f \left(S_{\tau(t)}^{p(1)} + \widehat{u}_{\tau(t),R}^{p(1)} \right) \right) \right) \left(\widehat{S}_{\tau(t),R}^{p(1)} - S_{\tau(t)}^{p(1)} \right) \\
&= \frac{\mathcal{N}m_1}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{p(1)} - S_{\tau(t)}^{p(1)} \right) + o_p(1)
\end{aligned}$$

where $m_1 = E \left(r_{i,t} \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} f \left(S_{\tau(t)}^{p(1)} + \widehat{u}_{\tau(t),R}^{p(1)} \right) \right)$. Using similar arguments as above for $III_{T,N,R}$, it follows that:

$$\begin{aligned}
\widehat{Z}_{T,N,R} &= \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(1),L} \right\} \mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(1)} \right\} \right. \\
&\quad \left. - \mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(N-1),U} < \widehat{S}_{i,\tau(t),R} \right\} \mathbb{1} \left\{ S_{\tau(t)}^{p(N-1)} < S_{i,\tau(t)} \right\} \right) \\
&+ \frac{\mathcal{N}m_1}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{p(1)} - S_{\tau(t)}^{p(1)} \right) - \frac{\mathcal{N}m_N}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{p(N-1)} - S_{\tau(t)}^{p(N-1)} \right) \\
&= Z_{T,N} - \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ \widehat{S}_{\tau(t),R}^{p(1),L} < \widehat{S}_{i,\tau(t),R} \right\} \mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(1)} \right\} \right. \\
&\quad \left. - \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{p(N-1),U} \right\} \mathbb{1} \left\{ S_{\tau(t)}^{p(N-1)} < S_{i,\tau(t)} \right\} \right) \\
&+ \frac{\mathcal{N}m_1}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{p(1)} - S_{\tau(t)}^{p(1)} \right) - \frac{\mathcal{N}m_N}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{p(N-1)} - S_{\tau(t)}^{p(N-1)} \right)
\end{aligned}$$

Under the null hypothesis 3.1, given assumption 3.6, and recalling the definition of $C_{1,\tau(t),R}$ and

$C_{N,\tau(t),R}$, the second term in the last equality reads as:

$$\begin{aligned} & \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \left(\sum_{i \in C_{1,\tau(t),R}} (r_{i,t} - \mu_i) - \sum_{i \in C_{N,\tau(t),R}} (r_{i,t} - \mu_i) \right) \\ & + \frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \left(\sum_{i \in C_{1,\tau(t),R}} \mu_i - \sum_{i \in C_{N,\tau(t),R}} \mu_i \right) = o_p(1). \end{aligned}$$

C.1.4.2 Part (ii)

Under the alternative hypothesis 3.2, given assumption 3.6, and recalling the definition of $C_{1,\tau(t),R}$ and $C_{N,\tau(t),R}$, the second term in the last equality reads as:

$$\frac{\mathcal{N}}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \left(\sum_{i \in C_{1,\tau(t),R}} \mu_i - \sum_{i \in C_{N,\tau(t),R}} \mu_i \right).$$

C.1.5 Corollary 1

C.1.5.1 Part (i)

By Theorem 3.3, under H_0 ,

$$\widehat{Z}_{T,N,R} - Z_{T,N} = \frac{\mathcal{N}m_1}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{p(1)} - S_{\tau(t)}^{p(1)} \right) - \frac{\mathcal{N}m_N}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{p(N-1)} - S_{\tau(t)}^{p(N-1)} \right) + o_p(1).$$

Consider only the first quantile, as the argument for the last quantile is identical. The first sample quantile and the corresponding order statistic are asymptotically equivalent, hence:

$$\frac{\mathcal{N}m_1}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{p(1)} - S_{\tau(t)}^{p(1)} \right) = \frac{\mathcal{N}m_1}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{q(1)} - S_{\tau(t)}^{q(1)} \right) + o_p(1),$$

where $\widehat{S}_{\tau(t),R}^{q(1)}$ and $S_{\tau(t)}^{q(1)}$ are the estimators of the first quantile based on the estimated and true sorting variable, i.e.

$$\widehat{S}_{\tau(t),R}^{q(1)} = \arg \min_x \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\frac{1}{\mathcal{N}} - \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq x \right\} \right) \left(\widehat{S}_{i,\tau(t),R} - x \right).$$

$$S_{\tau(t)}^{q(1)} = \arg \min_x \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\frac{1}{\mathcal{N}} - \mathbb{1} \left\{ S_{i,\tau(t)} \leq x \right\} \right) \left(S_{i,\tau(t)} - x \right).$$

Since $\widehat{S}_{\tau(t),R}^{q(1)}$ and $S_{\tau(t)}^{q(1)}$ satisfy the equations above, the pseudo-score first-order conditions are:

$$\begin{aligned}
0 &= \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\frac{1}{N} - \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{q(1)} \right\} \right) \\
&= \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\frac{1}{N} - \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq S_{\tau(t)}^{q(1)} \right\} \right) \\
&\quad - \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{q(1)} \right\} - \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq S_{\tau(t)}^{q(1)} \right\} \right) \\
&= \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\frac{1}{N} - \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq S_{\tau(t)}^{q(1)} \right\} \right) \\
&\quad - \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\left(\mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq \widehat{S}_{\tau(t),R}^{q(1)} \right\} - F \left(\widehat{S}_{\tau(t),R}^{q(1)} \right) \right) - \left(\mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq S_{\tau(t)}^{q(1)} \right\} - F \left(S_{\tau(t)}^{q(1)} \right) \right) \right) \\
&\quad - \left(F \left(\widehat{S}_{\tau(t),R}^{q(1)} \right) - F \left(S_{\tau(t)}^{q(1)} \right) \right) \\
&= \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\frac{1}{N} - \mathbb{1} \left\{ \widehat{S}_{i,\tau(t),R} \leq S_{\tau(t)}^{q(1)} \right\} \right) - I_N^A - II_N^A.
\end{aligned}$$

$$\begin{aligned}
0 &= \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\frac{1}{N} - \mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{q(1)} \right\} \right) \\
&= \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\frac{1}{N} - \mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{q(1)} \right\} \right) \\
&\quad - \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{q(1)} \right\} - \mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{q(1)} \right\} \right) \\
&= \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\frac{1}{N} - \mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{q(1)} \right\} \right) \\
&\quad - \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{q(1)} \right\} - \mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{q(1)} \right\} \right) - \left(F \left(S_{\tau(t)}^{q(1)} \right) - F \left(S_{\tau(t)}^{q(1)} \right) \right) \right) \\
&\quad - \left(F \left(S_{\tau(t)}^{q(1)} \right) - F \left(S_{\tau(t)}^{q(1)} \right) \right) \\
&= \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\frac{1}{N} - \mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{q(1)} \right\} \right) - I_N^B - II_N^B.
\end{aligned}$$

First, $I_N^A = o_p\left(\left|\widehat{S}_{\tau(t),R}^{q(1)} - S^{q(1)}\right|\right)$ and $I_N^B = o_p\left(\left|S_{\tau(t)}^{q(1)} - S^{q(1)}\right|\right)$ because of stochastic equicontinuity. Moreover, by a first-order expansion around $S^{q(1)}$:

$$II_N^A = f\left(S^{q(1)}\right)\left(\widehat{S}_{\tau(t),R}^{q(1)} - S^{q(1)}\right) + o_p\left(\left|\widehat{S}_{\tau(t),R}^{q(1)} - S^{q(1)}\right|\right).$$

$$II_N^B = f\left(S^{q(1)}\right)\left(S_{\tau(t)}^{q(1)} - S^{q(1)}\right) + o_p\left(\left|S_{\tau(t)}^{q(1)} - S^{q(1)}\right|\right).$$

Substituting back and rearranging:

$$\widehat{S}_{\tau(t),R}^{q(1)} - S^{q(1)} = \frac{1}{f\left(S^{q(1)}\right)} \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\frac{1}{N} - \mathbb{1}\left\{\widehat{S}_{i,\tau(t),R} \leq S^{q(1)}\right\} \right) + o_p\left(\left|\widehat{S}_{\tau(t),R}^{q(1)} - S^{q(1)}\right|\right),$$

$$S_{\tau(t)}^{q(1)} - S^{q(1)} = \frac{1}{f\left(S^{q(1)}\right)} \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\frac{1}{N} - \mathbb{1}\left\{S_{i,\tau(t)} \leq S^{q(1)}\right\} \right) + o_p\left(\left|S_{\tau(t)}^{q(1)} - S^{q(1)}\right|\right).$$

Subtracting the two expansions:

$$\widehat{S}_{\tau(t),R}^{q(1)} - S_{\tau(t)}^{q(1)} = \frac{1}{f\left(S^{q(1)}\right)} \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\mathbb{1}\left\{S_{i,\tau(t)} \leq S^{q(1)}\right\} - \mathbb{1}\left\{\widehat{S}_{i,\tau(t),R} \leq S^{q(1)}\right\} \right) + r_{\tau(t),R},$$

where $r_{\tau(t),R} = o_p\left(\left|\widehat{S}_{\tau(t),R}^{q(1)} - S^{q(1)}\right|\right) + o_p\left(\left|S_{\tau(t)}^{q(1)} - S^{q(1)}\right|\right)$ and $\frac{N}{\sqrt{T-R}} \sum_{t=R+1}^T r_{\tau(t),R} = o_p(1)$. Next, rewriting:

$$\begin{aligned} & \widehat{S}_{\tau(t),R}^{q(1)} - S_{\tau(t)}^{q(1)} \\ &= \frac{1}{f\left(S^{q(1)}\right)} \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\mathbb{1}\left\{S_{i,\tau(t)} \leq S^{q(1)}\right\} - \mathbb{1}\left\{S_{i,\tau(t)} \leq S^{q(1)} - \left(\widehat{S}_{i,\tau(t),R} - S_{i,\tau(t)}\right)\right\} \right) + r_{\tau(t),R} \\ &= \frac{1}{f\left(S^{q(1)}\right)} \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\left(\mathbb{1}\left\{S_{i,\tau(t)} \leq S^{q(1)}\right\} - F\left(S^{q(1)}\right) \right) \right. \\ & \quad \left. - \left(\mathbb{1}\left\{S_{i,\tau(t)} \leq S^{q(1)} - \left(\widehat{S}_{i,\tau(t),R} - S_{i,\tau(t)}\right)\right\} - F\left(S^{q(1)} - \left(\widehat{S}_{i,\tau(t),R} - S_{i,\tau(t)}\right)\right) \right) \right) \\ & \quad + \frac{1}{f\left(S^{q(1)}\right)} \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(F\left(S^{q(1)}\right) - F\left(S^{q(1)} - \left(\widehat{S}_{i,\tau(t),R} - S_{i,\tau(t)}\right)\right) \right) + r_{\tau(t),R} \\ &= I_N^C + II_N^C + r_{\tau(t),R}. \end{aligned}$$

First, by stochastic equicontinuity and Assumption 3.5, I_N^C is negligible after summing over t and multiplying by $N/\sqrt{T-R}$. Moreover, by a first-order expansion around $S^{q(1)}$:

$$II_N^C = \frac{1}{f(S^{q(1)})} \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} f(S^{q(1)}) \left(\widehat{S}_{i,\tau(t),R} - S_{i,\tau(t)} \right) + o_p \left(\left| \widehat{S}_{i,\tau(t),R} - S_{i,\tau(t)} \right| \right).$$

Hence:

$$\widehat{S}_{\tau(t),R}^{q(1)} - S_{\tau(t)}^{q(1)} = \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\widehat{S}_{i,\tau(t),R} - S_{i,\tau(t)} \right) + \widetilde{r}_{\tau(t),R},$$

where $\frac{Nm_1}{\sqrt{T-R}} \sum_{t=R+1}^T \widetilde{r}_{\tau(t),R} = o_p(1)$. It follows that:

$$\frac{Nm_1}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{q(1)} - S_{\tau(t)}^{q(1)} \right) = \frac{Nm_1}{\sqrt{T-R}} \sum_{t=R+1}^T \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} \left(\widehat{S}_{i,\tau(t),R} - S_{i,\tau(t)} \right) + o_p(1).$$

Because $\widehat{S}_{i,\tau(t),R}$ is estimated from a window of length R , the cross-sectional average estimation error is of order $O_p(R^{-1/2})$. Therefore:

$$\frac{Nm_1}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{q(1)} - S_{\tau(t)}^{q(1)} \right) = O_p \left(\sqrt{\frac{T-R}{R}} \right).$$

The same argument applies to the last quantile.

Part (a) If $(T-R)/R \rightarrow 0$, then the two quantile-error terms are $o_p(1)$, and Theorem 3.3 implies:

$$\widehat{Z}_{T,N,R} = Z_{T,N} + o_p(1).$$

Parts (b) and (c) If instead $(T-R)/R \rightarrow \pi$ with $0 < \pi < \infty$, then the same terms are $O_p(1)$ and contribute to the asymptotic variance. Under Assumption 3.1(iv), Theorem 3.1 yields the limit variance V for $Z_{T,N}$, and therefore:

$$\widehat{Z}_{T,N,R} \xrightarrow{d} N(0, \Omega), \quad \Omega = V + V_{S^{q(1)}} + V_{S^{q(N)}} + \text{covs},$$

and the variance terms are those generated by the first- and last-quantile measurement-error components:

$$\frac{\mathcal{N}m_1}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{q(1)} - S_{\tau(t)}^{q(1)} \right), \quad \frac{\mathcal{N}m_N}{\sqrt{T-R}} \sum_{t=R+1}^T \left(\widehat{S}_{\tau(t),R}^{q(N-1)} - S_{\tau(t)}^{q(N-1)} \right).$$

Under Assumption 3.3, Theorem 3.2 gives the variance V^* for $Z_{T,N}$, so that

$$\widehat{Z}_{T,N,R} \xrightarrow{d} N(0, \Omega^*), \quad \Omega^* = V^* + V_{Sq(1)} + V_{Sq(N)} + \text{covs.}$$

C.1.5.2 Part (ii)

This follows immediately from Theorem 3.3, part (ii).

C.1.6 Theorem 4

C.1.6.1 Part (i)

We restrict attention to the case $\frac{T-R}{R} \rightarrow \pi$ and $\frac{T_\delta-R_\delta}{R_\delta} \rightarrow \pi$ where $0 < \pi < \infty$. For the k -th overlapping block, define the infeasible subsample counterpart of $\widehat{Z}_{T,N,R}$ by

$$Z_{T_\delta,N}^{(k)} = \frac{\mathcal{N}}{\sqrt{T_\delta-R_\delta}} \sum_{t=k+R_\delta}^{k+T_\delta-1} \frac{1}{N_{\tau(t)}} \sum_{i=1}^{N_{\tau(t)}} r_{i,t} \left(\mathbb{1} \left\{ S_{i,\tau(t)} \leq S_{\tau(t)}^{p(1)} \right\} - \mathbb{1} \left\{ S_{\tau(t)}^{p(N-1)} < S_{i,\tau(t)} \right\} \right).$$

By the same argument as in the proof of Theorem 3.3, with T_δ and R_δ in place of T and R , we obtain for every k :

$$\begin{aligned} \widehat{Z}_{T_\delta,N,R_\delta}^{(k)} &= Z_{T_\delta,N}^{(k)} + \frac{\mathcal{N}m_1}{\sqrt{T_\delta-R_\delta}} \sum_{t=k+R_\delta}^{k+T_\delta-1} \left(\widehat{S}_{\tau(t),R_\delta}^{p(1)} - S_{\tau(t)}^{p(1)} \right) \\ &\quad - \frac{\mathcal{N}m_N}{\sqrt{T_\delta-R_\delta}} \sum_{t=k+R_\delta}^{k+T_\delta-1} \left(\widehat{S}_{\tau(t),R_\delta}^{p(N-1)} - S_{\tau(t)}^{p(N-1)} \right) + o_p(1). \end{aligned}$$

By the same argument as in the proof of Corollary 3.1, the two quantile-error terms can be rewritten with $q(1)$ and $q(N-1)$ in place of $p(1)$ and $p(N-1)$, respectively.

Under H_0 , Corollary 3.1(i)(b), applied to a block of length T_δ , together with Assumption 3.7, implies that under assumptions 3.1, 3.2, 3.4, 3.5:

$$\widehat{Z}_{T_\delta,N,R_\delta}^{(k)} \xrightarrow{d} N(0, \Omega).$$

Likewise, Corollary 3.1(i)(c), again applied to a block of length T_δ , together with Assumption 3.7, implies that under assumptions 3.1.(i)–(iii), 3.2-3.5:

$$\widehat{Z}_{T_\delta, N, R_\delta}^{(k)} \xrightarrow{d} N(0, \Omega^*).$$

Hence, for each k , the subsampled statistic $\widehat{Z}_{T_\delta, N, R_\delta}^{(k)}$ has the same continuous limiting law as $\widehat{Z}_{T, N, R}$ under H_0 . Moreover, since $R_\delta = \lfloor RT_\delta/T \rfloor$, $0 \leq \frac{RT_\delta}{T} - R_\delta < 1$. Therefore,

$$\begin{aligned} \frac{T_\delta - R_\delta}{T - R} &= \frac{T_\delta - RT_\delta/T}{T - R} + \frac{RT_\delta/T - R_\delta}{T - R} \\ &= \frac{T_\delta}{T} + \frac{RT_\delta/T - R_\delta}{T - R} \rightarrow 0, \end{aligned}$$

because $T_\delta/T \rightarrow 0$ and $0 < \pi < \infty$ implies $T - R \rightarrow \infty$. In addition, $T_\delta = \lfloor T^\delta \rfloor$ with $0 < \delta < 1$ implies $T_\delta \rightarrow \infty$, and therefore $K = T - T_\delta + 1 \rightarrow \infty$. Theorem 2.1 in Politis et al. (1997) applies and yields consistency of the empirical subsampling distribution. Consequently,

$$\lim_{T, N, R \rightarrow \infty} \Pr \left(c_{T_\delta, K, \alpha/2} \leq \widehat{Z}_{T, N, R} \leq c_{T_\delta, K, 1-\alpha/2} \right) = 1 - \alpha.$$

C.1.6.2 Part (ii)

By Corollary 3.1(ii), under H_A there exists $\eta > 0$ such that

$$\lim_{T, N, R \rightarrow \infty} \Pr \left(\frac{1}{\sqrt{T - R}} \left| \widehat{Z}_{T, N, R} \right| \geq \eta \right) = 1.$$

Applying the same argument as in Theorem 3.3(ii) to a block of length T_δ shows that, for each k ,

$$\widehat{Z}_{T_\delta, N, R_\delta}^{(k)} = O_p \left(\sqrt{T_\delta - R_\delta} \right).$$

Hence the empirical critical values satisfy

$$c_{T_\delta, K, \alpha/2} = O_p \left(\sqrt{T_\delta - R_\delta} \right), \quad c_{T_\delta, K, 1-\alpha/2} = O_p \left(\sqrt{T_\delta - R_\delta} \right).$$

Since $(T_\delta - R_\delta)/(T - R) \rightarrow 0$, it follows that

$$\frac{c_{T_\delta, K, \alpha/2}}{\sqrt{T - R}} \xrightarrow{p} 0, \quad \frac{c_{T_\delta, K, 1-\alpha/2}}{\sqrt{T - R}} \xrightarrow{p} 0.$$

Combining the last display with Corollary 3.1(ii) gives

$$\lim_{T,N,R \rightarrow \infty} \Pr \left(\left(\widehat{Z}_{T,N,R} < c_{T_\delta, K, \alpha/2} \right) \cup \left(c_{T_\delta, K, 1-\alpha/2} < \widehat{Z}_{T,N,R} \right) \right) = 1.$$

C.1.7 Proposition B

C.1.7.1 Part (i)

Given Assumptions B1 and B2, by the Law of the Iterated Logarithm for an ergodic strongly mixing process, for all $i = 1, \dots, N_{\tau(t)}$ and $t = R, \dots, T$:

$$\limsup_{R \rightarrow \infty} \frac{1}{\sqrt{2R \ln \ln R}} \sum_{j=\tau(t)-R+1}^{\tau(t)} (r_{M,j} - \bar{r}_{M,R}) \epsilon_{i,j} = v_{i,\tau(t)}$$

and a.s.

$$\liminf_{R \rightarrow \infty} \frac{1}{\sqrt{2R \ln \ln R}} \sum_{j=\tau(t)-R+1}^{\tau(t)} (r_{M,j} - \bar{r}_{M,R}) \epsilon_{i,j} = -v_{i,\tau(t)}$$

where

$$v_{i,\tau(t)}^2 = \text{avar} \left(\frac{1}{\sqrt{R}} \sum_{j=\tau(t)-R+1}^{\tau(t)} (r_{M,j} - \mathbb{E}(r_{M,j})) \epsilon_{i,j} \right)$$

Also, assumptions B1 and B2 ensure that for all $i = 1, \dots, N_{\tau(t)}$, and each $t = R, \dots, T$, $\widehat{A}_{i,R,\tau(t)}^{-1} - A_{i,\tau(t)}^{-1} = o_{a.s.}(1)$, where $\widehat{A}_{i,R,\tau(t)}$ was defined in (3.5). Hence,

$$\begin{aligned} & \sqrt{\frac{R}{2R \ln \ln R}} \left(\widehat{\beta}_{i,\tau(t),R} - \beta_{i,\tau(t)} \right) \\ &= A_{i,\tau(t)}^{-1} \frac{1}{\sqrt{2R \ln \ln R}} \sum_{j=\tau(t)-R+1}^{\tau(t)} (r_{M,j} - \bar{r}_{M,R}) \epsilon_{i,j} \\ & \quad + \left(\widehat{A}_{i,R,\tau(t)}^{-1} - A_{i,\tau(t)}^{-1} \right) \frac{1}{\sqrt{2R \ln \ln R}} \sum_{j=\tau(t)-R+1}^{\tau(t)} (r_{M,j} - \bar{r}_{M,R}) \epsilon_{i,j} \\ &= A_{i,\tau(t)}^{-1} \frac{1}{\sqrt{2R \ln \ln R}} \sum_{j=\tau(t)-R+1}^{\tau(t)} (r_{M,j} - \bar{r}_{M,R}) \epsilon_{i,j} + o_{a.s.}(1) \\ &= A_{i,\tau(t)}^{-1} \frac{1}{\sqrt{2R \ln \ln R}} \sum_{j=\tau(t)-R+1}^{\tau(t)} (r_{M,j} - \mathbb{E}(r_{M,j})) \epsilon_{i,j} + o_{a.s.}(1). \end{aligned}$$

Hence, a.s.

$$\limsup_{R \rightarrow \infty} \sqrt{\frac{R}{2R \ln \ln R}} \left(\widehat{\beta}_{i,\tau(t),R} - \beta_{i,\tau(t)} \right) = \sqrt{A_{i,\tau(t)}^{-1} v_{i,\tau(t)}^2 A_{i,\tau(t)}^{-1}} = \sigma_{i,\tau(t)}$$

and analogously, a.s.

$$\liminf_{R \rightarrow \infty} \sqrt{\frac{R}{2R \ln \ln R}} \left(\widehat{\beta}_{i,\tau(t),R} - \beta_{i,\tau(t)} \right) = -\sqrt{A_{i,\tau(t)}^{-1} v_{i,\tau(t)}^2 A_{i,\tau(t)}^{-1}} = -\sigma_{i,\tau(t)}$$

It then follows from Lemma 2.1 in [Corradi \(1999\)](#) that for all $i = 1, \dots, N_{\tau(t)}$, and each $t = R, \dots, T$,

$\widehat{v}_{i,R,\tau(t)}^2 - v_{i,\tau(t)}^2 = o_{a.s.}(1)$, and thus $\widehat{\sigma}_{R,i,\tau(t)}^2 - \sigma_{R,i}^2 = o_{a.s.}(1)$. Hence,

$$\frac{\widehat{c}_{LIL,i,\tau(t),R}^L}{c_{LIL,i,R}^L} \xrightarrow{a.s.} 1,$$

and so the statement in the Lemma follows.

C.1.7.2 Part (ii)

Given part (i), by the same argument as in Lemma C.1, replacing $c_{i,\tau(t),R}^L$ and $c_{i,\tau(t),R}^U$ with $\widehat{c}_{LIL,i,\tau(t),R}^L$ and $\widehat{c}_{LIL,i,\tau(t),R}^U$ respectively, the statement follow.

C.1.8 Proposition B'

C.1.8.1 Part (i)

Given assumptions B1' and B2', $\frac{1}{\sqrt{2R \ln \ln R}} \sum_{s=\tau(t)-R+1}^{\tau(t)} (X_{i,s} - \bar{X}_R) u_{i,t}$ satisfies a LIL, and so for $j = 0, 1$, a.s.

$$\limsup_{R \rightarrow \infty} \sqrt{\frac{R}{2R \ln \ln R}} \left(\widehat{\beta}_{i,j,\tau(t),R} - \beta_{j,i,\tau(t)} \right) = \sigma_{\beta_{j,i,\tau(t)}}$$

and analogously for the liminf. Then, conditionally on $W_{i,\tau(t)}$, a.s.,

$$\limsup_{R \rightarrow \infty} \sqrt{\frac{R}{2R \ln \ln R}} \left(\widehat{\beta}_{i,\tau(t),R} (W_{i,\tau(t)}) - \beta_{i,\tau(t)} (W_{i,\tau(t)}) \right) \leq \sigma_{\beta_{0,i,\tau(t)}} + \sigma_{\beta_{1,i,\tau(t)}} |W_{i,\tau(t)}|$$

$$\liminf_{R \rightarrow \infty} \sqrt{\frac{R}{2R \ln \ln R}} \left(\widehat{\beta}_{i,\tau(t),R} (W_{i,\tau(t)}) - \beta_{i,\tau(t)} (W_{i,\tau(t)}) \right) \geq -\sigma_{\beta_{0,i,\tau(t)}} - \sigma_{\beta_{1,i,\tau(t)}} |W_{i,\tau(t)}|$$

Finally, since for $j = 0, 1$ $\widehat{\sigma}_{R, \beta_{j,i}, \tau(t)}^2 = \sigma_{\beta_{j,i}, \tau(t)}^2$ the statement follows.

C.1.8.2 Part (ii)

By the same argument as in Lemma C.1, replacing $c_{i, \tau(t), R}^L$ and $c_{i, \tau(t), R}^U$ with $\pm \sigma_{\beta_{0,i}, \tau(t)} \pm \sigma_{\beta_{1,i}, \tau(t)} |W_{i, \tau(t)}|$, the statement follow.

C.1.9 Theorem B1

C.1.9.1 Part (i)

Given Assumptions B1' and B2', Assumption 3.5 is satisfied with

$$c_{\tau(t), R}^U = -c_{\tau(t), R}^L = \sup_{i \leq N_{\tau(t)}} \sigma_{i, \tau(t)} \sqrt{\frac{2 \ln \ln R}{R}},$$

for the unconditional case, and with

$$c_{\tau(t), R}^U = -c_{\tau(t), R}^L = \sup_{i \leq N_{\tau(t)}} \sigma_{\beta_{0,i}, \tau(t)} \sqrt{\frac{2 \ln \ln R}{R}} + \sup_{i \leq N_{\tau(t)}} \sigma_{\beta_{1,i}, \tau(t)} \sqrt{\frac{2 \ln \ln R}{R}} |W_{i, \tau(t)}|,$$

for the conditional case. Given part (ii) of Propositions B and B', the statement in Lemma C.1 follows. Then the statements in (i) and (ii) follow the same argument as in Theorem 3.3 and in Corollary 3.1.

C.1.9.2 Part (ii)

By the same argument as in part (i) of Theorem 3.3, the statement follow.

C.2 Simulations for Classification Errors

C.2.1 True (Unobservable) Sorting Variable

Tables C.1-C.6 show summary statistics for the market beta, for different values of E , when stocks are sorted into N portfolios. Across these tables, two patterns stand out. First, as E increases, and hence the estimation window $R = 21E$ becomes longer, the portfolio distributions become markedly tighter and less skewed, especially in the tails: for $N = 10$, the standard deviation falls from 0.27 to 0.12 in P1 and from 0.35 to 0.22 in P10 when moving from $E = 12$ to $E = 60$, while the corresponding tail ranges shrink from $[-2.32, 0.13]$ to $[0.12, 0.26]$ in P1 and

from $[0.35, 4.12]$ to $[0.22, 2.68]$ in P10. This is consistent with more precise beta estimates and therefore smaller classification errors. Second, reducing the number of portfolios from $N = 10$ to $N = 5$ pools adjacent deciles into broader groups, which smooths the cross-sectional pattern in means and medians and makes the extreme portfolios less isolated, but increases dispersion in the intermediate portfolios because each group spans a wider range of betas. Overall, larger R reduces estimation noise, whereas smaller N trades off finer sorting for more aggregated portfolios.

Table C.1: Summary statistics of the estimated betas across portfolios ($E = 12, N = 10$).

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Std. Dev.	0.27	0.06	0.05	0.04	0.04	0.04	0.05	0.06	0.09	0.35
Min	-2.32	0.13	0.34	0.50	0.64	0.78	0.92	1.09	1.30	1.61
Mean	-0.09	0.24	0.42	0.57	0.71	0.85	1.00	1.19	1.44	1.99
Median	-0.02	0.24	0.42	0.57	0.71	0.85	1.00	1.18	1.43	1.89
Max	0.13	0.34	0.50	0.64	0.78	0.92	1.09	1.30	1.61	4.12

Table C.2: Summary statistics of the estimated betas across portfolios ($E = 24, N = 10$).

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Std. Dev.	0.18	0.05	0.04	0.04	0.04	0.04	0.04	0.05	0.08	0.28
Min	-1.31	0.20	0.38	0.53	0.66	0.79	0.92	1.07	1.26	1.54
Mean	0.03	0.29	0.46	0.59	0.72	0.85	0.99	1.16	1.39	1.86
Median	0.07	0.29	0.46	0.59	0.72	0.85	0.99	1.16	1.38	1.79
Max	0.20	0.38	0.53	0.66	0.79	0.92	1.07	1.26	1.54	3.27

Table C.3: Summary statistics of the estimated betas across portfolios ($E = 60, N = 10$).

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Std. Dev.	0.12	0.05	0.04	0.03	0.03	0.03	0.04	0.05	0.07	0.22
Min	-0.57	0.26	0.43	0.56	0.68	0.80	0.92	1.05	1.21	1.45
Mean	0.13	0.34	0.50	0.62	0.74	0.86	0.98	1.12	1.32	1.71
Median	0.15	0.35	0.50	0.62	0.74	0.86	0.98	1.12	1.31	1.65
Max	0.26	0.43	0.56	0.68	0.80	0.92	1.05	1.21	1.45	2.68

Table C.4: Summary statistics of the estimated betas across portfolios ($E = 12, N = 5$).

	P1	P2	P3	P4	P5
Std. Dev.	0.26	0.09	0.08	0.11	0.38
Min	-2.32	0.34	0.64	0.92	1.30
Mean	0.07	0.49	0.78	1.09	1.71
Median	0.13	0.50	0.78	1.09	1.61
Max	0.34	0.64	0.92	1.30	4.12

Table C.5: Summary statistics of the estimated betas across portfolios ($E = 24$, $N = 5$).

	P1	P2	P3	P4	P5
Std. Dev.	0.19	0.08	0.08	0.10	0.31
Min	-1.31	0.38	0.66	0.92	1.26
Mean	0.16	0.52	0.79	1.08	1.62
Median	0.20	0.53	0.79	1.07	1.54
Max	0.38	0.66	0.92	1.26	3.27

Table C.6: Summary statistics of the estimated betas across portfolios ($E = 60$, $N = 5$).

	P1	P2	P3	P4	P5
Std. Dev.	0.14	0.07	0.07	0.08	0.26
Min	-0.57	0.43	0.68	0.92	1.21
Mean	0.23	0.56	0.80	1.05	1.51
Median	0.26	0.56	0.80	1.05	1.45
Max	0.43	0.68	0.92	1.21	2.68

C.2.2 Classification Errors

Tables C.7-C.18 show the contamination error and the information loss error with and without trimming, for different values of E , when stocks are sorted into N portfolios. Since larger E corresponds to a larger estimation window R , the tables display a clear pattern: both contamination and information-loss errors fall as R increases and as the number of portfolios is reduced from $N = 10$ to $N = 5$. In the untrimmed case, the overall error for the first portfolio declines from 32.92% to 19.60% when $N = 10$ and from 23.00% to 12.38% when $N = 5$ as E rises from 12 to 60; the corresponding figures for the last portfolio fall from 24.83% to 13.40% and from 19.44% to 10.38%, respectively. The same ranking is preserved in the trimmed tables. Moreover, as E becomes larger, the remaining errors are increasingly concentrated in adjacent portfolios, while contamination from more distant portfolios becomes negligible; reducing N has a similar effect, because wider portfolio bins make the portfolio cutoffs less sensitive to beta-estimation noise.

Table C.7: Contamination and information loss errors without trimming (in %, $E = 12$, $N = 10$).

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Contamination P1	32.92	19.19	6.59	3.09	1.69	0.99	0.61	0.37	0.23	0.16
Information loss P1	32.92	22.46	5.35	2.12	1.12	0.68	0.44	0.31	0.23	0.22
Contamination P10	0.22	0.16	0.21	0.30	0.45	0.74	1.46	3.79	17.51	24.83
Information loss P10	0.16	0.10	0.12	0.16	0.25	0.44	1.00	3.29	19.32	24.83
Portfolio size	424	424	424	424	423	423	423	423	423	423

Table C.8: Contamination and information loss errors with trimming (in %, $E = 12$, $N = 10$).

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Contamination P1	25.83	13.96	5.37	2.68	1.52	0.92	0.58	0.37	0.24	0.19
Information loss P1	25.83	2.93	0.17	0.00	0.00	0.00	0.00	0.00	0.01	0.23
Contamination P10	0.23	0.16	0.19	0.25	0.36	0.56	1.00	2.23	8.86	13.84
Information loss P10	0.20	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.52	13.84
Portfolio size	276	46	13	3	0	0	0	1	13	259
Trimmed quantity	148	378	411	421	423	423	423	422	410	164
Trimmed percentage	34.93	89.08	97.02	99.31	99.88	99.96	99.99	99.75	96.81	38.78

Table C.9: Contamination and information loss errors without trimming (in %, $E = 24$, $N = 10$).

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Contamination P1	27.52	18.51	4.91	1.99	1.01	0.53	0.27	0.15	0.08	0.05
Information loss P1	27.52	20.91	3.77	1.32	0.64	0.35	0.22	0.14	0.09	0.07
Contamination P10	0.07	0.07	0.10	0.14	0.23	0.38	0.76	2.28	15.27	19.29
Information loss P10	0.05	0.03	0.04	0.06	0.09	0.17	0.43	1.78	16.64	19.29
Portfolio size	393	393	392	392	392	392	392	392	392	392

Table C.10: Contamination and information loss errors with trimming (in %, $E = 24$, $N = 10$).

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Contamination P1	20.25	12.50	3.99	1.75	0.93	0.50	0.27	0.16	0.09	0.06
Information loss P1	20.25	2.64	0.12	0.01	0.00	0.00	0.00	0.00	0.01	0.08
Contamination P10	0.08	0.06	0.09	0.11	0.18	0.29	0.51	1.28	6.63	9.23
Information loss P10	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.01	1.14	9.23
Portfolio size	253	46	22	10	4	2	1	5	34	257
Trimmed quantity	140	346	371	382	388	391	391	387	358	134
Trimmed percentage	35.65	88.26	94.42	97.44	98.99	99.59	99.82	98.70	91.38	34.29

Table C.11: Contamination and information loss errors without trimming (in %, $E = 60$, $N = 10$).

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Contamination P1	19.60	15.84	2.42	0.78	0.33	0.14	0.06	0.03	0.02	0.00
Information loss P1	19.60	16.85	1.78	0.52	0.22	0.11	0.06	0.03	0.02	0.01
Contamination P10	0.01	0.01	0.02	0.04	0.07	0.11	0.24	0.89	12.02	13.40
Information loss P10	0.00	0.00	0.00	0.01	0.01	0.03	0.09	0.57	12.68	13.40
Portfolio size	317	316	316	316	316	316	316	316	316	316

Table C.12: Contamination and information loss errors with trimming (in %, $E = 60$, $N = 10$).

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
Contamination P1	11.93	8.94	1.82	0.65	0.29	0.13	0.05	0.03	0.02	0.01
Information loss P1	11.93	2.88	0.09	0.01	0.00	0.00	0.00	0.00	0.00	0.01
Contamination P10	0.01	0.01	0.02	0.03	0.05	0.08	0.16	0.47	4.13	4.96
Information loss P10	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.02	1.96	4.96
Portfolio size	209	64	40	22	13	8	7	19	69	225
Trimmed quantity	107	253	276	295	303	308	309	297	247	90
Trimmed percentage	33.92	79.88	87.28	93.15	95.75	97.43	97.73	93.96	78.22	28.62

Table C.13: Contamination and information loss errors without trimming (in %, $E = 12$, $N = 5$).

	P1	P2	P3	P4	P5
Contamination P1	23.00	17.81	3.68	1.14	0.36
Information loss P1	23.00	18.73	2.89	0.94	0.43
Contamination P5	0.43	0.77	2.35	15.88	19.44
Information loss P5	0.36	0.56	1.93	16.58	19.44
Portfolio size	847	847	847	847	846

Table C.14: Contamination and information loss errors with trimming (in %, $E = 12$, $N = 5$).

	P1	P2	P3	P4	P5
Contamination P1	16.81	12.36	3.05	1.04	0.36
Information loss P1	16.81	5.01	0.30	0.18	0.36
Contamination P5	0.39	0.60	1.54	8.22	10.75
Information loss P5	0.39	0.16	0.19	3.14	10.75
Portfolio size	622	226	113	172	596
Trimmed quantity	225	621	734	675	251
Trimmed percentage	26.60	73.28	86.71	79.71	29.60

Table C.15: Contamination and information loss errors without trimming (in %, $E = 24$, $N = 5$).

	P1	P2	P3	P4	P5
Contamination P1	18.23	15.37	2.20	0.52	0.13
Information loss P1	18.23	15.87	1.71	0.47	0.18
Contamination P5	0.18	0.39	1.33	13.24	15.14
Information loss P5	0.13	0.23	0.96	13.81	15.14
Portfolio size	785	785	784	784	784

Table C.16: Contamination and information loss errors with trimming (in %, $E = 24$, $\mathcal{N} = 5$).

	P1	P2	P3	P4	P5
Contamination P1	11.76	9.43	1.75	0.46	0.13
Information loss P1	11.76	4.79	0.35	0.15	0.15
Contamination P5	0.16	0.29	0.87	6.07	7.39
Information loss P5	0.13	0.08	0.17	3.67	7.39
Portfolio size	574	259	180	243	580
Trimmed quantity	211	526	604	541	204
Trimmed percentage	26.85	67.01	77.03	68.98	26.06

Table C.17: Contamination and information loss errors without trimming (in %, $E = 60$, $\mathcal{N} = 5$).

	P1	P2	P3	P4	P5
Contamination P1	12.38	11.43	0.81	0.12	0.02
Information loss P1	12.38	11.55	0.66	0.13	0.03
Contamination P5	0.03	0.10	0.47	9.77	10.38
Information loss P5	0.02	0.04	0.26	10.05	10.38
Portfolio size	633	632	632	632	632

Table C.18: Contamination and information loss errors with trimming (in %, $E = 60$, $\mathcal{N} = 5$).

	P1	P2	P3	P4	P5
Contamination P1	6.59	5.88	0.60	0.10	0.02
Information loss P1	6.59	4.42	0.27	0.06	0.03
Contamination P5	0.03	0.08	0.30	3.64	4.05
Information loss P5	0.02	0.02	0.09	3.03	4.05
Portfolio size	495	301	249	283	496
Trimmed quantity	137	332	383	349	136
Trimmed percentage	21.71	52.46	60.62	55.16	21.53