

BOOK VALUE OPTIMISATION, RISK AND REDISTRIBUTION

A THESIS PRESENTED FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY OF IMPERIAL COLLEGE LONDON
AND THE
DIPLOMA OF IMPERIAL COLLEGE
BY
JEAN HERSKOVITS

DEPARTMENT OF MATHEMATICS
IMPERIAL COLLEGE
180 QUEEN'S GATE, LONDON SW7 2BZ

AUGUST 2024

STATEMENT OF ORIGINALITY

I certify that this thesis, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Signed: Jean Herskovits

COPYRIGHT

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.

Book Value Optimisation, Risk and Redistribution

Thesis advisors:

Professor Johannes Muhle-Karbe and Professor Alex Tse

ABSTRACT

Across three different settings, we show that if classical metrics are key to understanding behaviours of financial actors, it is necessary to also examine others, generally considered as secondary, to comprehend actions taken by decision-makers. Firstly, while managers of public companies maximize investors' utility first, they also use redistribution to manage the company's earnings-per-share ratio, stock price, and their own managerial stock options' valuation. We prove that in a frictionless setting, share repurchases and dividends are equivalent in analogy to the classical Modigliani-Miller theorem. We further exhibit extensions in which the utility equivalence of dividends and repurchases is broken. Our analytical differentiation of redistribution methods provides grounding towards explaining the various empirical redistribution behaviours observed in markets. Secondly, we prove and display empirically that typical risk metrics such as Value-at-Risk and Expected Shortfall mechanically grow as the number market observations is reduced. Modern international regulation assumes the opposite and in turn implies the double counting of risk over illiquid assets. We argue that this accounting methodology specification has considerable ramifications, altering the global structure through which banks calibrate their risk limits: regulated banks are now opting for the "Standardised Approach" instead of the "Internal Model Approach". Finally we present a novel sovereign credit grades prediction model, which beats Nomura's previous internal implementations as well as, to our knowledge, all previously published ones. It is the first to attain a sufficient level of accuracy satisfactory to financial institutions. In particular, better modelling of credit grades entails a more precise calibration of credit risk. Not linked to material adjustments in P&L or positions, improved credit grades' modelling can reduce portfolios' uncertainty: changes to the predictions of purely indicative sovereign credit grades have meaningful repercussions over the haircuts and spreads observed in sovereign bond markets.

To my family.

ACKNOWLEDGMENTS

I would like to thank Dr. Eduardo Epperlein for the unceasing backing he provided to this studentship for over a decade. I also thank Dr. Eduardo Epperlein for the support and guidance he delivered in leading the developments presented in Chapter 2.

For their advice, help and comments, I would like to thank my colleagues Kenichi Kurasawa, Xaver Sailer, John Sleath and Lolis Vasilis at Nomura, Luca Gerolla at Imperial College London and former colleague Kyriakos Chourdakis.

I would like to thank Alfonso de la Torre for the economic expertise he consistently brought in creating the model presented in Chapter 3, as well as my former manager Dr. Alan Smillie for initiating this project. I also acknowledge the valuable directions that Professor Lukas Gonon, Professor Johannes Ruf and Professor Cristopher Salvi have contributed to this chapter. I would like to thank Niklas Walter for his help.

Finally, I would like to thank my mathematical finance PhD colleagues at Imperial College London for the help and support they have provided continuously over the last four years.

LIST OF FIGURES

1.1	First Public Shares, Dutch East India Company Foundational Charter.	8
1.2	Baseline Model Equilibrium Behaviour for varying expected growth rate μ .	45
1.3	Baseline model equilibrium behaviour for varying risk aversion γ .	46
1.4	Heterogeneous investors dividends' model equilibrium for varying risk-aversion γ and growth expectation disagreement $(\mu_1 - \mu_2)$.	58
1.5	Heterogeneous investors dividends' model equilibrium for varying optimistic expected growth rate μ_1 and risk-free rate R_f .	59
1.6	Heterogeneous investors buybacks' model equilibrium for varying risk-aversion γ and optimistic expected growth rate μ_1 .	60
1.7	Endogenized risk-free rate model equilibrium for varying risk-aversion γ .	73
1.8	Endogenized risk-free rate model equilibrium for varying expected growth rate μ .	74
1.9	Typical dividends' equilibrium behaviour for varying discount factor β .	74
1.10	Typical Model behaviour with trading constraints for varying θ_{BB} .	79
1.11	Investor utility Surface for varying risk-aversion γ and ratio bought back θ_{BB} .	79
1.12	Optimal model behaviour with trading restrictions for varying risk-aversion γ .	80
1.13	Equilibrium model behaviour with managers for varying risk-free rate R_f , expected growth rate μ and firm's assets K_0 .	84
A.14	Acceptable Ratios Bought Back (<i>Blue</i>) for Varying γ	87
2.1	Earliest derivatives contracts: futures.	108
2.2	P&L distributions for daily and weekly observation frequencies.	113
2.3	VaR and ES for varying observation period T .	119
2.4	Maximum period incurring a VaR / ES surcharge for varying tail level α .	121
2.5	Portfolio's VaR for varying MtM frequency and correlation level.	123
2.6	Yearly expected number of 1-day 99% VaR exceptions as a function of MtM period T .	125
2.7	Expected number of 99%-VaR exceptions as a function of % of missing risk.	127
A.8	Available equity names' geographical and sectorial repartition.	131
A.9	Distribution across equity names of $(\text{VaR}_{1D}(\alpha, \chi^T) - \text{VaR}_{1D}(\alpha, \chi^1))$. <i>Left</i> : for $T = 2$ and $\alpha = 97.5\%$ – <i>Right</i> : for $T = 5$ and $\alpha = 99.5\%$	132
A.10	Average and quantiles of VaR surcharges, varying MtM frequency T and tail level α .	132
A.11	Equity returns and autocorrelation function for varying MtM frequency.	134
A.12	Available CDS names' geographical and sectorial repartition.	135
A.13	CDS returns' volatility is clustered around 2008.	136

A.14	CDS returns' autocorrelation and standard deviation for varying MtM frequency, 2008 and post-2008.	136
A.15	CDS daily absolute returns, impact of infrequent MtM on VaR, for different MtM frequencies and VaR levels.	137
A.16	Equity autocorrelation levels during crises periods.	138
A.17	For varying MtM frequency, distribution of 99%-VaR surcharges across equities and CDS.	139
A.18	Distribution of impacts from infrequent MtM on single stock portfolios hedged by the market average.	140
A.19	Average VaR and ES surcharges for varying returns reporting threshold (T).	142
A.20	Equity daily log-returns, autocorrelation and infrequent MtM's impact on VaR.	145
A.21	Equity daily relative-returns, autocorrelation and infrequent MtM's impact on VaR.	145
A.22	Equity daily absolute returns, autocorrelation and infrequent MtM's impact on VaR.	146
A.23	CDS returns – Distribution of the impact on 1-Day 1% VaR of bi-monthly price updates.	146
A.24	Distribution of autocorrelations, rolling weekly returns.	147
A.25	Infrequent returns standard deviations, using weekly and monthly rolling returns.	148
A.26	VaR surcharge from infrequent MtM, on weekly and monthly rolling returns.	148
A.27	Equity returns' correlation distribution for varying return type and window.	149
3.1	Standardised Approach, credit risk to default probability table.	159
3.2	Sovereign credit grades' correspondence between agencies.	164
3.3	Structure of the Available Dataset.	168
3.4	Geographical repartition of available countries.	168
3.5	Sovereign credit grade changes and countries' economic development indicators.	172
3.6	Dataset's geographical biases for 2018 & 2019.	176
3.7	Time Series Data Slicing Process.	180

LIST OF TABLES

2.1	Probability of a 97.5% and 99% VaR exception being observed for varying MtM frequency.	124
A.2	Maximum observation periods incurring a VaR surcharge for varying α	133
A.3	Maximum observation periods incurring a VaR surcharge for varying α during stressed market periods.	139
A.4	Maximum observation periods incurring a VaR surcharge on a structured portfolio for varying tail level α	141
A.5	VaR and ES surcharges for small varying threshold $\epsilon(T)$ at different tail levels.	143
A.6	VaR and ES values for original normalized CDS and normal returns.	143
3.1	Typical List of Features used in predicting the evolution of sovereign credit grades.	163
3.2	Distribution of sovereign ratings by Standard & Poors, 2019.	167
3.3	Legacy model typical prediction quality.	167
3.4	Constant Model prediction quality compared to typical legacy model.	174
3.5	Features chosen to correspond to the nine ones currently in use.	175
3.6	Quality of the out-of-sample fit of neural networks over the restrained dataset.	177
3.7	Impact of varying the kernel size WL on number of available entries in the resulting dataset.	180
3.8	For varying maximal filling size, number of available entries in the resulting dataset.	182
3.9	Grade changes distribution of the whole reconstructed dataset.	182
3.10	Multi-Year Neural Network model results outperform the results from the constant model, meeting its benchmark target.	184
3.11	Features chosen by Credit Risk as the most relevant to predict Sovereign Credit Grades.	186
3.12	Average correlation of features' time series to time series of sovereign credit grades.	187
3.13	Average Random Forest performance per class for varying number of individual decision trees in the forest.	188
3.14	By assigning unequal weights to the two kinds of errors, we can shift the model's performance.	188
A.15	Complete List of Available Features concatenated from various data vendors - Part I.	191
A.16	Complete List of Available Features concatenated from various data vendors - Part II.	192

CONTENTS

0	INTRODUCTION	1
1	BUYBACKS, DIVIDENDS, EQUIVALENCE AND DIFFERENCES	6
1	Setup of the baseline model	20
1.1	The firm	21
1.2	The representative investor	22
1.3	Competitive equilibrium under exogenously given distribution strategy	23
1.4	Optimally endogenized dividend payout and buyback strategy .	34
2	Equivalence of dividend and buyback in the baseline model	36
3	Model extensions under which dividend-buyback equivalence remains valid	47
3.1	Investors with heterogeneous beliefs over firm's profitability . . .	47
3.2	Consumption and endogenous interest rate	61
4	Model extensions under which dividend-buyback equivalence breaks down	75
4.1	Investors with trading constraint	75
4.2	Managers with stock options	80
A	Appendix	87
1	Removing Assumption 1.7: negative Δ	87
2	Irrelevance of the order of endogenization in the case with endo- genized risk-free rate	99
2	INFREQUENT MARKING-TO-MARKET DOES NOT REDUCE VaR NOR BACK- TESTING EXCEPTIONS	106
1	Infrequent Marking to Market for a Single Risk Factor	112
2	Infrequent MTM's Effect on Correlated Risk Factors	122
3	VaR Backtesting Exceptions with Infrequent P&L	123
4	VaR and Backtesting under Daily MtM	126
5	Conclusion	127
A	Appendix	129
1	Empirical equities, the effects of infrequent MtM	130
2	Empirical CDS, the effects of infrequent MtM	134
3	Stressed Market Periods, 2008 Financial Crisis and Covid	138
4	VaR of Individual Equities Hedged against Index	140
5	Threshold Based Infrequent Returns	141
6	The impact of changing the returns' structure	144
7	The impact of the returns' window on their correlation structure	148
3	MACHINE LEARNING FOR SOVEREIGN CREDIT GRADE PREDICTION	153
1	Legacy Industry Models and Limitations	163
2	Constant Predictor & Modelling Challenges	167
3	No History Neural Net Model	174

4	Multiple Years Neural Net Model	177
1	Building the Dataset	179
2	Model Results	184
5	Random Forests	185
6	Conclusion	189
A	Appendix	191
1	Available Indicators	191
2	Data Reconstruction	193
3	Class Weights in Random Forests	194
	REFERENCES	220



INTRODUCTION

This thesis is split in three chapters which present research conducted while working in partnership between Imperial College London's Mathematical Finance Section and Nomura International Securities' Risk Methodology Group. The first chapter, a contribution to theoretical Corporate Finance within the field of optimal capital allocation, is an academic comparative analysis of dividend distribution and share repurchases. The second chapter discusses the ramifications of recent financial regulations and how liquidity risk evolves with changing observed market conditions, both analytically and empirically. The third chapter outlines a new method to better predict sovereign credit grades' movements. Each chapter thus relates to a distinct area of research in modern mathematical finance. Yet, all three parts, for different kinds of assets and frameworks, present analytical, numerical and statistical results sharing a common overall theme. Indeed, we study how risk-takers and decision makers can optimise the face value of the assets they manage without impacting their material commitments. That is, their assets' risk levels and valuations can be altered without tangible modifications to the assets' structure, such as payments, investments, reserves or portfolio composition.

More precisely, the first chapter proves that market capitalisation, share price, debt-to-equity and earnings-per-share ratios can be managed by executives through redistribution of assets to investors without modifying the investors' utility, their main concern, but also without impacting the company's investments, cash flows, growth prospects or bankruptcy probability. The second chapter outlines how new international regulation ([BCBS, 2019c](#)) allows for the management of crucial risk measures, the value-at-risk and expected shortfall at an individual risk factor level, without adjusting the positions and exposures to underlying securities. This change in reported risk figures

is not tied to any real change in market conditions, specifically market liquidity and volatility: “real” risk remains unchanged. Finally, the third chapter presents a new method for sovereign credit grades’ prediction. This process enhances the accuracy of credit risk measurement, mainly for sovereign debt but also for corporate bonds. Our methodology change reduces the uncertainty of the portfolios’ valuations. Thus, in the absence of changes to a security’s market value, or to its measurable risk drivers, such as its issuer’s default probability or financial situation, the security’s risk valuation is altered by the better prediction of its future credit grade. Indeed, countries’ credit grades themselves do not matter towards determining government bonds’ yields, solvency status nor does their movements impact banks’ Profit-and-Loss figures (PnL). However, the better calibration of risk from enhanced sovereign grades predictions affects the book value of banks’ positions, and as a consequence, banks’ behaviours are also impacted.

BUYBACKS, DIVIDENDS, EQUIVALENCE AND DIFFERENCES

The first chapter, developed in collaboration with Professor Muhle-Karbe¹ and Professor Tse², compares the impact of dividends’ distribution and share repurchases on firms’ equilibrium valuations and optimal behaviours. Given a publicly owned firm, its managers can decide to redistribute some, or all, of the firm’s assets at time zero to their risk-averse shareholders. As the firm’s investment after redistribution is risky, the redistribution allows the firm’s management to set their firm’s level of investment. Investors can also buy and sell the company’s stock to rebalance their exposure to the firm’s risky technology ventures.

The Modigliani-Miller theorem states that, in a perfect market, the growth prospects of a company define its valuation, not its financing methods, such as debt and equity issuance (Modigliani and Miller, 1961, 1958, 1963). We confirm that the Modigliani-Miller theorem applies to our baseline model. Indeed, we show that within a friction-less market with no information asymmetry in a one-period setting, investors are not affected by the redistribution method chosen by the firm. That is, the investors’ optimal expected utility, that executives are legally mandated to maximise, does not depend on whether the firm’s management opts for dividends’ distribution or common stock repurchases. We further prove that there exists only one optimal dividends strategy, and one optimal repurchases strategy, which both redistribute the same ratio of the company’s assets back to its investors. While these results are first proven with a single representative investor and an exogenous risk-free rate, we prove that our results are robust to investors’ heterogeneity and to multi-period investor consumption. Conversely, taking into account either managerial preferences or investors not willing

¹Imperial College London, Mathematical Finance Section.

²University College London, Departement of Mathematics.

or able to reduce their ownership in the firm, the equivalence between dividends and repurchases is broken. Majority shareholders benefit from the distribution of dividends and managers prefer buying back company stock.

Furthermore, even for setups in which investors' optimal utility is independent of the redistribution method implemented, the choice of redistribution method allows for the management of the company's stock price, market capitalisation and earnings-per-share ratio. As mentioned, the dividends and repurchases optimal strategies redistribute the same amount to investors. Thus, our analysis indicates that to decipher managers' redistribution decisions, one must also consider their influence on firm's indicators which do not directly impact tangible capital flows and legally binding investors' utilities. In other words, executives use redistribution as a tool to manage the company's book value, both from earnings and market value perspectives. As such, redistribution permits the book value optimisation of publicly traded firms, without affecting their actual structure.

These book value optimisations have significant implications. Indeed, stock price itself and earnings-per-share ratio are crucial indicators of a company's prospects. At a systemic level, the widespread use of stock buybacks has motivated political action ([IRS, 2022a](#); [IRS, 2022b](#)), while increased stock prices are taken into account in managerial decisions ([Campello and Graham, 2013](#)).

INFREQUENT MARKING-TO-MARKET DOES NOT REDUCE VaR NOR BACKTESTING EXCEPTIONS

In the second chapter, Doctor Epperlein³ and I study the impact of infrequent trade observations, or low marking-to-market frequency, on value-at-risk (VaR) and expected shortfall (ES) risk measures. For normally distributed returns, we prove that observing less market transactions for a risk factor mechanically increases reported risk figures. We show that this effect stands for large sets of empirical equity and CDS time series. Swapping daily marking for weekly observations, the resulting value-at-risk statistics can double. For expected shortfall especially, the increase in reported risk is maintained even for very scarce observations, e.g. monthly. In essence, rarer observations aggregate small individual price moves in larger ones. Thus, infrequently observed returns follow a distribution with fatter tails, leading to higher reported risk measures.

While the “real” risk carried by the position is unchanged, the risk figures are multiplied. As such, buying third party data, increasing observation frequencies, drives the banks' risk down with no material change to their portfolios. In practice, current “Basel” regulations ([BCBS, 2019c](#)) have a specific provision for risk factors' not observed frequently enough: non-modellable risk factors (NMRF). Under NMRF, for any

³Head of the Risk Methodology Group, Nomura International Securities.

risk factor not marked-to-market at least 24 times a year, historical VaR and ES numbers are multiplied before their final values are set. Combined with our findings, NMRF entails that risk over illiquid or over-the-counter traded (OTC) assets is double-counted under current international regulations.

As in the first chapter, despite no material changes to a bank’s positions, its risk charges are artificially revised depending on the bank’s ability to observe the market. The consequences of double-counting risk under NMRF are momentous. Indeed, the “*Basel*” framework globally offers two options for risk calibration, the Standardised Approach (SA) (BoE, 2022a), and the Internal Models Approach (IMA) (BCBS, 2022b; BCBS, 2022a). Regulated institutions opting for SA use largely predetermined risk charges for most products, shared across the industry. IMA provides guidelines for risk modelling to regulated institutions, which then individually develop and calibrate proprietary models. While the latter was more popular under the previous “*Basel 2.5*”, SA has become the default option for large financial institutions under “*Basel 4.0*”.⁴ We argue that the drop in IMA use is largely motivated by NMRF’s double counting of risk across the aforementioned class of assets. The shift from IMA to SA initiated by a revision in book value calibration is a substantial outcome in itself. Moreover, under SA, all financial institutions follow the same risk models and thus carry the same risk limits. The Gilts crisis (Pinter, 2023) highlighted the core importance of preventing correlated large-scale exposures in capital markets. Recent fire sales triggered by insolvencies (Bouveret and Haferkorn, 2023) targeted the very institutions regulated by “*Basel 4.0*”. The newly widespread and shared SA risk limits thus increase systemic risks, originated from a non-material book valuation methodology change.

MACHINE LEARNING FOR SOVEREIGN CREDIT GRADE PREDICTION

The final chapter of the thesis presents a novel model for prediction of Sovereign Credit Grades. Leveraging the third party data available at Nomura, and collaborating with colleagues from Nomura’s Credit Risk Division, the predictors developed significantly improve over the accuracy of previous internal implementations. More importantly, our models outperform all known published sovereign grades’ predictors.

As banks intend on keeping exposure to sovereign debts over long horizons, it is mandated that their default and downgrading (migration) risks models encompass long future periods (BCBS, 2009). “*Basel*” SA regulation directly ties sovereign bonds’ risk to their sovereign credit grade (BIS, 2013) (*see fig. 3.1*). The better modelling of sovereign credit grades’ evolution is thus crucial for banks. Yet, sovereign credit grades have no direct link to bonds’ market valuations nor to their default probabilities. Credit

⁴As of 2024, to our knowledge, only two major investment banks plan on implementing IMA for “*Basel 4.0*”. ISDA surveys reported IMA usage of 97.5% under “*Basel 2.5*” and 52.5% under “*Basel 4.0*” (ISDA and for Financial Markets in Europe, 2024).

grades' sole repercussion is on the book value of the bonds, through their risk charge. The improvements in credit grades' predictions do not impact banks' portfolios and PnL, and their exposures to sovereign bonds remain unchanged.

However, by changing the grades applied to sovereign bonds in risk models, the calibration of credit portfolios' risk has improved. In turn, the uncertainty attached to sovereign credit positions is reduced. Thus, bond holders can reduce the conservative margins and haircuts they apply to sovereign exposures. Such improvements at a systemic level can participate to the adjustment of pivotal sovereign borrowing rates.⁵

This thesis shows in three distinct settings that executives, risk managers and more generally decision makers' actions can be motivated by the optimisation of features usually considered as non-material indicators. That is, while they consider classical metrics such as capital allocation, investors' utility and PnL, their actions are also driven by features including earnings-per-share ratio, share price, market capitalisation, market risk and credit risk. While many market completeness hypotheses imply the irrelevance of these aforementioned "non-material" indicators, we present empirical, statistical and analytical frameworks in which their consideration is key.

⁵In practice, nations usually ultimately borrow from markets, though banks still act as market-makers. Yet, sovereign credit can be issued by banks directly ([Werner, 2015](#)). Bank thus have not only a central role in the secondary market for bonds, but also as primary lenders.

“I hate stock buybacks. [...] I think they’re one of the most self-serving things that corporate America does. [...] They don’t do a thing to make their company better.”

U.S. Senator of New York, and Democratic
Senate Majority Leader, Chuck Schumer

1

BUYBACKS, DIVIDENDS, EQUIVALENCE AND DIFFERENCES

It is empirically agreed that various capital redistribution methods by publicly listed firms are not equivalent. Analytical research has however generally focused on studying optimal capital allocation through dividends. With the recent rise in public interest on the importance of firm share buybacks, it is crucial to understand where and how they differ from the more classical dividend payments. By understanding the analytical conditions under which dividends and share buybacks are equivalent, we can set boundaries to the extension of classical optimal dividend allocation results. More particularly, it might explain why firms’ practical habits do not universally agree with the behaviours entailed by analytical studies. We also confirm findings from empirical works: share repurchases do permit the manipulation and stabilization of the company’s stock price as well as management of the firm’s all-important earnings per share ratio is also enabled by buybacks. Finally, taking into account managerial stock options and different kinds of investors, the utility equivalence between the two setting is broken.

In parallel, we show that if the results from Modigliani & Miller stand in certain settings, they do not systematically apply, depending on the indicators considered and on the model’s specific setup. We finally find that without some market completeness hypothesis, the two mechanisms output very contrasting equilibria.

INTRODUCTION

The Dutch East India Company (VOC) is widely known to be the first ever company with publicly traded shares. Established through an anachronistic Initial Public Offering (IPO), by the evening of Saturday the thirty-first of August 1602, 1143 shareholders started being able to publicly trade the company's shares on the simultaneously created Amsterdam Stock Exchange (Petram, 2014). The world's oldest share certificate still available to this day thus dates back to August 27th, 1604 (*see fig. 1.1*). 48 years after its creation, in 1650, the VOC had paid its investors back thirteen times their initial investment in dividends, at an average yearly coupon rate of 27% (Koot, 1995).

Isaac Le Maire was one of the original 1143 investors in the VOC. He invested 85,000 out of the total 3,674,945 guilders raised by the VOC at its inception (Petram, 2014), and served on its Board of Directors until his dismissal over verified claims of embezzlement on February 22nd, 1605 concerning the journey of Wybrand van Warwijck in 1602. After multiple attempts at establishing a rival to the VOC, in 1609, Isaac le Maire spearheaded the first ever "Bear Run". The Groote Compagnie that he cofounded sold VOC shares for future delivery. The premium at which VOC shares traded was nearly halved, from 212% in 1607, to 127% in 1609 (Dillen and Poitras, 2006).

At the core of Isaac Le Maire's claims was one issue: by 1609, the VOC had not distributed any dividends back to its investors. Despite increasing revenues, the company's free cash-flow was not allocated to shareholders. This "war chest" would arguably become the central reason for the VOC's success over its competitors (Koot, 1995).

While unsuccessful, Le Maire's actions forced the VOC to start distributing dividends to its shareholders in 1610: a 75% dividend in silver but also a 50% dividend in pepper. The share of VOC's profit set aside for dividends would remain the crucial issue for shareholders over its history (Jongh, 2011).

Seven years after the creation of the first publicly traded share, the distribution of dividends had already become the most important issue the shares' holders.

Stock buybacks, on the other hand, are a more recent apparition in shareholder finance. While SEC Rule 10-18b, providing a "safe harbor" for the execution of share repurchases (Lazonick, 2015), is often pointed at as the origin for the modern use of share buybacks, the prevalence of share repurchases in publicly traded companies however predates this ruling (Holt and Morris, 1934; Klug, 1993). Indeed, in 1964 and 1966, Richard West, Harold Bierman and Eugene Brigham wrote in the first two analytical studies on stock buybacks that "In the past decade, major US corporations have increasingly repurchased significant amount of their common shares" (Bierman Jr and West, 1966), and "The dollar value of repurchases increased from approximately \$600 million in 1960 to \$1.7 billion in 1964" (Brigham, 1964), demonstrating the importance



Figure 1.1: Right: Dutch East India Company foundational 1602 charter establishing its monopoly over trade with the “East Indies”.

Left: Earliest known share certificate receipts, belonging to Hans Thijs, referring to share ownerships in the Dutch East India Company by Leonardt Rans (Top) and Jan Brouwer (Bottom) on August 27th 1604 and August 28th 1604 respectively, probably referring to the same share. The share registers “Enkhuizen” themselves are lost.

Source: National Archives, The Hague. Taken from Leo Akveld and Els M. Jacobs, eds., *The Colourful World of the VOC* (Amsterdam: 2002), p. 12 (Right). Archives at Leiden University Libraries, Arend Pietersma (Left).

buybacks already had reached.

The analytical study of shareholder corporations more generally, and in turn of dividends allocation can arguably be traced back to 1932 in “The Modern Corporation and Private Property”, by Adolf A. Berle and Gardiner C. Means (Berle and Means, 1932). Dividends allocation was already an active field of research in the inter-war period (HBR, 1926, 1927b). The first dividends allocation modelling attempts relevant to our analysis refer to John Lintner’s eponymous model, published in 1965 (Lintner, 1962, 1964), inspired by a previous empirical study of his (Lintner, 1956) and the Capital Asset Pricing Model by William Sharpe in 1964 (Sharpe, 1964), which resulted in a Nobel prize in 1990.

Both models were inspired by the Modigliani-Miller capital allocation results (Modigliani and Miller, 1961, 1958, 1963). The Nobel-prize winning 1958 Modigliani-Miller theorem states that in a perfect market, the way a company is financed does not influence its valuation, only the expected prospects for its growth do. A perfect market is character-

ized by the absence of transaction costs, no asymmetry of information and no personal nor corporate taxes. In particular, the Modigliani-Miller theorem implies that the company’s redistribution policy is independent from its valuation: “the division of the stream between cash dividends and retained earnings in any period is a mere detail” (Modigliani and Miller, 1958). This result provides the framework in which the vast majority of analytical studies are included in. As outlined below in this introduction, the first part of our analysis confirms that this result indeed holds implementing share repurchases as well as dividends in our framework.

Since these foundational developments in the analytical description of share repurchases and dividend distributions, many alternative models were advanced, in parallel to the still active study of these original models (Andres et al., 2015; Bauer and Bhattacharyya, 2006). An introduction to payout policy modelling can be found in (Allen and Michaely, 2003). Section 3.1 of our analysis is reminiscent of the model for dividends modelling in the context of asymmetric information by Merton Miller and Kevin Rock (Miller and Rock, 1985). Going as far back as Adam Smith (Smith, 1776), the agency problem, that is the diverging interests of publicly traded firms’ Directors and shareholders has also been an early area of study (Crum, 1935; HBR, 1927a). Michael Jensen and William Meckling’s dedicated a publication in 1976 to the modelling of the agency problem (Jensen and Meckling, 1976). Also encompassing asymmetric information, the Bhattacharyya’s 1979 dividend signaling model argues that dividends act as free cash-flow signaling tool (Bhattacharyya, 1979). The signaling and free-cash flow hypotheses are discussed further below. Comparing stock repurchases, dividends and equity emissions, other general equilibrium models have been developed (John and Williams, 1985; Løkka and Zervos, 2008; Tse, 2018), building upon previous models (Dionne and Ouederni, 2011; Howe et al., 1992) and focusing on the agency problem (Kumar, 1988). While not incorporating bankruptcies, our analysis (*see Section 4.2*) takes inspiration from Bart Lambrecht and Stewart Myers’ more recent model (Lambrecht and Myers, 2017). Specifically, we introduce a firm and risk-averse shareholders who make static optimal investment and redistribution decisions. Both maximise investors’ expected and risk-adjusted utility from the firm’s uncertain future returns, reminiscent of the static optimal liquidation strategies outlined by Robert Almgren and Neil Chriss (Almgren and Chriss, 1997). Other models outlining the importance of bankruptcies regarding the optimal capital debt structure of a firm were put forward (Gomes and Schmid, 2019; Leland, 1994; Leland and Toft, 1996).

There are multiple ways of executing share buybacks: open tender offers, private negotiations, Dutch auctions or put executions. Open market repurchases are the most popular nowadays, representing more than 95% of the volume of buybacks executed by publicly listed firms. Our analysis thus focuses on open market repurchases, though older studies that we reference use open tender offers instead, which used to be popular

before the SEC Rule 10-18b ruling (Comment and Jarrell, 1991).

A multitude of factors arise to explain the differences between the two main redistribution methods, dividends distribution and share repurchase plans, as well as the level of redistribution operated by public firms. Multiple reviews attempt an exhaustive overview of these factors (Chen and Obizhaeva, 2022; Weston and Siu, 2003). Depending on the methodologies used in different studies, results can differ, and analyses often obtain diverging results (Banyi et al., 2008). Extensive surveys of Board Executives were carried out to determine which factors are considered in practice (Brav et al., 2005; Cudd et al., 1996; Dittmar, 2000; Grullon and Ikenberry, 2000; Wansley et al., 1989). We focus on the medium to long-term effects of redistribution. For an introduction to the dynamics of stock prices around redistribution announcements, *see* (Comment and Jarrell, 1991; Masulis, 1980), and after dividend announcements specifically, *see* (Petit, 1972). In parallel, our analysis and the works we reference typically consider firms in a standalone fashion. However, this assumption is not always correct, as corporate governance is not isolated from the average managerial behaviour (Williamson, 1984), nor are the investors' preferences that managers target rational (Shefrin and Statman, 1984).

We now provide a list of the main drivers outlined in previous academic literature that are important in understanding the motivations behind redistribution either through share repurchases or dividend distribution.

1. Unequal Tax Treatment of Redistribution Methods:

Tax treatment in various countries might be the most important factor driving redistribution method decisions. Shareholders' taxation rates vary across periods and countries (Hammer, 1975). Generally, dividends are treated as income, and thus carry a higher marginal tax rate than proceeds from repurchases, which are treated as capital gains.

In the U.S., dividends received are classified as income and thus are taxed upwards of 40% (all income above \$539,901 is taxed at a rate of 37% as of 2023). Investors' gains from buybacks, classified as capital gains, are taxed at a rate capped at 20% (depending on their marital status, capital gains tax brackets vary for individuals, gains entering the 20% tax bracket from \$258,600 to \$488,500 per year). Recent legislation in the U.S. attempted to shrink this gap while maintaining global capital gains and income tax brackets (IRS, 2022a; IRS, 2022b). *See Report* (KPMG, 2024) for a review of their effect. In the U.K., the same qualitative difference in taxation is present. France increased its tax level on dividends by 70% in 2013 (Boissel and Matray, 2019; Lefebvre et al., 2021).

Academic concern over the levels of taxation of capital gains and dividends is

old, dating back to at least 1922 (May, 1922), and the classification of dividends as income too (HBR, 1926). Mentioned earlier, Ronald Masulis also argues that deferral of earnings to optimise the effective tax rates is the main driver in redistribution decisions (Masulis and Trueman, 1988). The original Modigliani-Miller setup was also extended to include distinct tax treatments of dividends and repurchases (DeAngelo and Masulis, 1980; Farrar and Selwyn, 1967). Other studies further included the heterogeneity of shareholders' tax rates stemming from their various levels of wealth (Brennan, 1970; Elton and Gruber, 1970). Merton Miller and Myron Scholes also empirically exhibited that the tax level on dividends is practically taken into account in U.S. firm valuations (Miller and Scholes, 1982). Yet, other replicating studies have been inconclusive (Litzenberger and Ramaswamy, 1982). The bracket structure of modern taxation systems also implies that very small company dividends are little taxed, if taxed at all. Thus, depending on the amount redistributed, the competitiveness of repurchases and dividends varies (Brennan and Thakor, 1990). Malcolm Baker and Jeffrey Wurgler show that managers take the tax situations of their shareholders for redistribution into account (Baker and Wurgler, 2004). In particular, managers decide the type of redistribution based on the type of shareholders (Allen et al., 2000; Bajaj and Vijh, 1990). Indeed, the favorable tax treatment of buybacks only applies to individuals: a high ratio of institutional shareholders yields a higher use of dividends. Our analysis in Section 4.1 confirms that different shareholder structures affect the optimal redistribution method. Finally, reminiscent of the 2022 Nobel Prize methodology (*see* (Hull et al., 2022)), Raj Chetty and Emmanuel Saez showed that dividends tax cuts lead to increases in dividends payments (Chetty and Saez, 2005).

2. Stock Buybacks as an EPS and debt-to-equity ratios management tools:

Buybacks reduce the stock of available shares. Over the short term, repurchases thus generally increase the Earnings-Per-Share ratio (EPS). Failure to meet periodic EPS ratio targets carries negative consequences for a firm's executives, as well as for its market capitalisation. It has consequently been shown empirically that firms execute buybacks around target dates to manipulate their EPS ratios (Gao and Kronlund, 2020; Hribar et al., 2006). However other studies have shown the opposite (Dann et al., 1991). Devaluations following share buybacks announcements have happened in the past too, the announcements being interpreted as a signal of a firm's lack of good investment opportunities (Décamps and Villeneuve, 2007). Our analysis also concludes that not all repurchases, even optimal, increase share price. While the specific use of repurchases for EPS management is discussed, it is agreed upon that cash-flow and earnings management

by executives is common (Burgstahler and Dichev, 1997), and that executives' compensation is affected by EPS levels (Baber et al., 1999). Our analysis confirms that the use of share repurchases allow for the management of the EPS ratio.

While it is true that buybacks mechanically raise the EPS ratio over the short term, the firm's liquid assets spent on redistribution cannot then be invested towards future growth. As such, repurchase plans are nowadays most often financed by debt (Elgouacem and Zago, 2020). That is, buybacks are used as a capital restructuring mechanism to move companies towards optimal equity-to-debt ratios (Hovakimian et al., 2002; Lei and Zhang, 2016; Shoven and Simon, 1987). Other works have studied the interplay between dividends distribution, cost of debt and the agency problem (Rozeff, 1982).

3. The Agency Problem and Managerial Interest:

As mentioned, the study of separation between corporate ownership and control is ancient, going back to Adam Smith (Smith, 1776). It has been an uninterrupted active field of research until today. An extensive review of the corporate agency problem, that is, the diverging interests of managers and shareholders, is outside the scope of our study. We focus on the agency problem and its relationship to cash flow redistribution¹.

Some aforementioned models already exhibit the analytical implications of diverging interests between shareholders and managers (Jensen and Meckling, 1976; Lambrecht and Myers, 2017). Other influential models establish the importance of the managers-investors relationship towards the definition of redistribution levels (Fama and Jensen, 1983). Bart Lambrecht and Stewart Myers' model in particular outlines that the managers-shareholders relationship can be summarized as managers choosing to pay out the minimal amount of redistribution acceptable to investors (Lambrecht and Myers, 2017). This argument was verified empirically: in settings where managers' opinions are more decisive, firms retain their earnings from being redistributed twice as much (Dittmar et al., 2003). Analogously, there is a significant correlation between strong investor boards and higher dividends (Chang and Dutta, 2012; Ginglinger and L'her, 2006). Under legislatures that do not defend shareholders rights, firms' cash is valued less, taking into account the low probability that executives will distribute it back to investors (Pinkowitz et al., 2006).

¹Many studies can be outlined towards a general understanding of the agency problem. Eugene Fama's work is central (Fama, 1980; Fama and Jensen, 1983). Mark Mizruchi's analysis shows how powerful firms' executives can get, and thus outlines the magnitude of the ramifications of the agency problem (Mizruchi, 1983a,b). We refer to its impact on market valuations but its impacts are far ranging, e.g. on employee average salaries (Cronqvist et al., 2009).

Managers sometimes have information unavailable to investors which can lead to better investment on their behalf (Myers and Majluf, 1984). However, typically, redistribution of the firm’s earnings as dividends can align the managers’ interests with the shareholders, increasing the executives’ risk appetite. Indeed, it hinders management’s reckless investment of available cash (Easterbrook, 1984; Rozeff, 1982). Thus, good corporate governance is tied to higher payout ratios (Hwang et al., 2013).

Finally, to align managers’ utility to the investors’, executives are often endowed stock options. While modern volumes of managerial stock options are high, their implementation within executives’ compensations packages is not new (Crum, 1935; HBR, 1927a). Empirically, it has been shown that the endowment of stock options decreases dividends (Lambert et al., 1989), motivates buybacks in the hope of raising the stock price (Weisbenner, 2000), and exacerbates the attention given to the EPS ratio (Gao and Kronlund, 2020). Section 4.2 of our analysis considers the impact of endowing managers with stock options on our model’s equilibrium (Carpenter, 1998). We thus opine that rising amounts of managerial stock options are substantial driver of the rise in popularity of share repurchases. We also show that emission of dividends generally reduce the value of executives’ stock options.

4. Shareholder Concentration and Takeover Defence

We already established that shareholders’ tax exposure heterogeneity influences the redistribution decisions. The shareholding structure of a firm also impacts the redistribution strategy of the firm. Firstly, the presence of shareholders with controlling voting rights but a lower share of earnings’ rights, often resulting from multiple equity classes, affects the redistribution. These investors reduce the amount given back to investors (Lins, 2003), especially when they represent individual and not institutional investors (Thanatawee, 2013b). Actually, the presence of controlling shareholders with a minority stake over the firm’s cash flow modifies the valuation of firms directly: a 6% to 25% reduction in market capitalisation (Cronqvist and Nilsson, 2003), and much worse performance in stressed market conditions (Baek et al., 2004). As such, markets globally also price a significant premium for corporate control over earnings’ rights (Atanasov, 2005; Zingales, 1994).

Conversely, large shareholders controlling a large proportion of the earnings’ right of a public company significantly increase the amount the firm redistributes (Gugler and Yurtoglu, 2003; Harada and Nguyen, 2011). Simon Johnson & al. described this behaviour as “tunneling” in their eponymous 2000 study (Johnson et al., 2000). Even one large investor is sufficient for tunneling to happen (Yun

and Hai-ming, 2011).

These actions often come at the cost of the minority shareholders' rights (Claessens et al., 2000). Legislation strengthening minority stakeholders' rights thus impacts the capital structure of firms and their level of redistribution (Franks and Mayer, 2002; Guney et al., 2007; Houston et al., 2018; Shleifer et al., 2000). Rules as simple as ensuring the representation of minority shareholders during company board meetings have an effect on dividend policy (Lv et al., 2023). For a review of the factors driving corporate governance, one can refer to (Shleifer and Vishny, 1997). Our analysis in Section 4.1 indeed confirms that the presence of controlling interests among the firm's shareholders impacts its optimal redistribution strategy.

Buybacks are also used to actively alter the shareholder structure of the firm. Indeed, implementing repurchases, the most pessimistic investors are bought out first. The remaining optimistic investors provide defence against takeovers, as they are more expensive to buy-out. Buybacks have thus been shown to be used as a corporate defence tool by management (Denis, 1990; Sinha, 1991), not always to the benefit of the remaining shareholders (Bradley and Wakeman, 1983). As the remaining shareholders are mechanically more optimistic, the firm itself becomes less attractive for takeovers. Also, its liquid assets have been spent on buying back shares, or the debt burden has increased in order to finance the transactions. On top of these two effects, implementing a repurchase plan can have a price stabilization effect (Thanatawee, 2013a; Wang et al., 2021), which helps towards rendering the company less accessible to hostile takeover plans.

5. Signaling and Free-Cash Flow Hypotheses

Since the establishment of the Modigliani-Miller framework, the main factor motivating the study of dividends, buybacks and redistribution globally are the "signaling" and free cash-flow hypotheses. Both relate to the same underlying question: under Modigliani-Miller, since a firm's capital structure is irrelevant towards its valuation, why do firms still distribute dividends, considering their effective tax rate is generally twice superior to buybacks' capital gains taxation? And thus, are dividends, but also repurchases, used as a method by a firm's managers to communicate about the firm's future earnings' prospects: are dividends, repurchases and redistribution correlated to the company's earnings? The free cash-flow hypothesis states that firms tend to redistribute retained earnings to investors when past profits have been high. Meanwhile, the signaling hypothesis states that managers distribute dividends when they expect the firm's future earnings to be good, and that they repurchase shares when they consider that the company's current stock price undervalues it. That is, managers are using

redistribution as a signaling tool to convey the optimistic asymmetric information they have about the firm’s future. Despite numerous studies, no consensus exists over the signaling hypothesis and, to a lesser extent, over the free cash-flow hypothesis. First of all, many analytical models have attempted to take the unequal taxation into account to conclude that using dividends for signaling is indeed optimal for firms’ executives. The Modigliani-Miller original framework was extended in this direction (Ross, 1977). Sudipto Bhattacharyya’s influential “Bird-in-hand” model is largely designed to model asymmetric information between a company’s executives and investors (Bhattacharyya, 1979; Dionne and Ouederni, 2011). Other models conclude that dividends must be used for signaling as their use for redistributing retained earnings is sub-optimal (John and Williams, 1985). Many other promising models have been proposed (Ambarish et al., 1987; Vermaelen, 1984). Some focus on specific behaviour: Aharon Ofer and Anjan Thakor exhibit analytically that some repurchase announcements can entail a price drop (Ofer and Thakor, 1987), which our model confirms. Sasson Bar-Yosef and Itzhak Venezia show analytically that firms willing to lower dividends cannot do so without sending bad outlook signals to their investors (Bar-Yosef and Venezia, 1991). Indeed, in practice, regular dividends cuts are very frequent for firms experiencing revenue drops, while they are rare otherwise (Balachandran et al., 2012; DeAngelo et al., 1992).

Literature agrees that the free cash-flow hypothesis is a major driver of redistribution (Howe et al., 1992; Vafeas and Joy, 1995). In particular, redistribution is correlated to earnings in the last available period (Benartzi et al., 1997a; Denis et al., 1994). Whilst the previous studies solely state that the free cash-flow hypothesis is true, others also contradict the signaling hypothesis: statistically, redistribution is correlated to past earnings, and not to future ones (Benartzi et al., 1997b; Gonedes, 1978; Rozeff, 1982; Watts, 1973). Eugene Fama and Harvey Babiak show that dividends in practice are linearly distributed according to the original Lintner’s model predictions, which do not incorporate asymmetric information (Fama and Babiak, 1968). Other studies have showed that dividends’ correlation is highest to capital expenditures, not to earnings (Yoon and Starks, 2015), and that earnings’ movements best explain investment figures (Vogt, 1994). Some general equilibrium models have argued that past earnings and tax rates define the level of redistribution (Goldstein et al., 2001).

Some empirical studies have shown that independent of the signaling hypothesis being true or not, markets do react to dividends and repurchases announcements as if signaling was intended by management (Asquith and Mullins Jr, 1983, 1986), and markets even price redistribution announcements before they happen (Banker et al., 1993). Some work has also attempted to show that positive

changes in dividends contain information about the future stability of cash flows (Dyl and Weigand, 1998), while other studies refute it (Carroll, 1995). While dividends signaling is disputed, repurchases have been shown to be correlated to future earnings (Vermaelen, 1981). Also, while not all repurchases signal under-valuation, some instances are initiated by management to signal their belief that the company’s stock is undervalued (Ikenberry et al., 1995; Stephens and Weisbach, 1998). Publications have outlined that while dividends overall do not seem to display signaling, carefully separating dividends in different categories (e.g. exceptional, large, small) allows one to conclude that only some types of dividends carry signaling (C. Mitchell Conover and Simpson, 2016). Similarly, signaling is present once one controls for the level of earnings’ growth observed in the period during which dividends are announced (Aharony and Swary, 1980). Finally, few studies find that signaling exists over all forms of dividends (Dedman et al., 2014; Nissim and Ziv, 2002).²

The list of effects presented above is not intended to be exhaustive³. It outlines that repurchases and dividends have been shown to affect public firms structurally in non-trivial ways, and that their study has been a non-interrupted field of academic research since Modigliani and Miller’s seminal results. Depending on the era (DeAngelo et al., 2004; Julio and Ikenberry, 2004; Klug, 1993), sector (Dempsey et al., 1993) and region (Kim et al., 2024), different effects will dominate over others, resulting both in varying optimal behaviour but also distinct impacts from similar redistributions plans. Indeed, empirical research over redistribution has been carried in many countries and regions, Australia (Harris and Ramsay, 1995), Egypt (Asquith and Mullins Jr, 1983), Europe (Gupta and Wagner, 2018), France (Ginglinger and L’her, 2006; Mellouli and Khiari, 2015), Germany (Klug, 1993), India (Kandpal and Kavidayal, 2015; Mishra, 2005; Thirumagal and Vasantha, 2018; Thirumalvalavan and Sunitha, 2006), Iran (Avaz-zadeh, 2015), Japan (Wada, 2005), Malaysia (Alquhaif et al., 2020), Pakistan (Afza and Mirza, 2011), South Korea (Lee et al., 2005), Thailand (Vithessonthi, 2008) and

²Signaling has been studied and shown to exist within specific settings and other firm financing decisions. The use of convertible bonds over standard bonds indicates optimistic prospects for the company (Constantinides and Grundy, 1989). Stock splits (McNichols and Dravid, 1990), and stock issuance (Myers and Majluf, 1984), especially in large amounts (Elgers and Murray, 1985), are positive signals for the firm’s future earnings. Dividends in newly public companies are a strong indicator of its future earnings (Lipson et al., 1998), and cross-listed firms with dividends outperform others (Ghadhab, 2023).

³Among others, we did not mention the Wealth Transfer Hypothesis, shifting current wealth from creditors to shareholders, against future earnings (Nishikawa et al., 2011), nor the impact of creditors’ confidence on payout policies (Byrne and O’Connor, 2017). The flexibility hypothesis argues that buybacks are more easily decided upon since dividends distributed are empirically shown to repeat in the future, and thus require a long-term commitment from the firm (Guay and Harford, 2000; Jagannathan et al., 2000). Modigliani and Miller assumed a perfectly liquid market. Studies have studied the optimal behaviour of firms in the presence of transaction costs (Hunting and Paulsen, 2013).

Vietnam (Wang et al., 2021) among others.

The main goal of this Chapter is to distinguish analytically the behaviours of modern firm value models implementing repurchases or dividends, as well as establishing equivalences between the two redistribution methods. From the analytical grounding of our framework, we can then analytically decipher the settings and situations in which each of the aforementioned factors will dominate the redistribution decisions. Indeed, there has been little analytical modelling of share repurchase mechanisms since the SEC’s 10-18b ruling (Baird et al., 2020).

However, buybacks have recently attracted widespread political attention (Henault, 2023; Sorkin et al., 2022). We already mentioned that high volumes of common stock repurchases are not new (Brigham, 1964), yet, their level in recent years is eye-watering, “Analysts at Goldman Sachs project that total S&P 500 repurchases will reach \$925 billion this year and \$1.075 trillion in 2025” (Grant, 2024). Their volume has been on the rise for the last two decades: more than half of the total S&P 500’s earnings from 2002 to 2013 were allocated to share repurchases (Lazonick, 2014). Apple alone will buy back \$110 billion of stock in 2024 despite falling revenue (Nellis et al., 2024). Their total yearly value in the US has now reached more than a percent of the total amount of shares outstanding (Aramonte, 2020), and their volume has already surpassed dividends’ twice over⁴. For the most active firms in this field, it represents over 10% of their stocks being bought back every financial year. This is equivalent to 4% of the U.S.’s GDP spent on share repurchases each year (Graeme, 2020).

Previously mentioned, the current U.S. administration has targeted share repurchases specifically. French President Emmanuel Macron has argued against the use of buybacks⁵. Some regulators also have expressed concern over the magnitude of repurchases implemented in the last decade (Jackson Jr., 2018). Institutions (Asness et al., 2017) and leading academics (DeAngelo, 2023) have pointed out that these critics often stem from a misunderstanding of how share repurchases work, and why companies use them.

Thus, enhancing our understanding of the analytical behaviour of repurchases is particularly relevant today. Academic research has shown that across eras, regions, and legislative frameworks, the uses and effects of redistribution methods varied substantially. Establishing further grounding of the consequences of redistribution methods, and in particular of the factors distinguishing dividends and buybacks, is needed to best assess the potential repercussions of proposed legislation changes.

As described, optimal dividends allocation is a very active field of research, while repurchases mathematical analysis is sparser. Inspired by the widespread Almgren-Chriss liquidation framework, our general equilibrium setup implements both share repur-

⁴Source: S&P Dow Jones Indices (September 2022)

⁵Source: Emmanuel Macron, 22/03/2023, LCI.

chases and dividends at once (Almgren and Chriss, 1997). Also deviating from the Almgren-Chriss setup, the market impact of trades is obtained endogenously through offer & demand market clearing conditions instead of being measured through a preset equation. Indeed, dividend distribution and share repurchases were considered empirically for nearly a century as two analogous types of capital redistribution. Our study aims at putting forward a modern common framework for their analysis. Our analysis can in turn be used to establish mathematical grounding to the various effects driving redistribution decisions. In particular, we confirm analytically some of the factors listed above, and reject others.

Our setting, presented in Section 1, models in two periods the actions of a firm's shareholders and management. As in the Almgren-Chriss liquidation framework (Almgren and Chriss, 1997), while the firm's returns are stochastic, defined by drift and volatility terms, the investors and managers' optimisation is static and determined at time 0 for the two periods. In short, risk-averse investors are endowed a position in a publicly traded firm which can be traded at a price determined through market clearing. The firm's assets can either be redistributed to investors proportionally to their share holdings, or invested by the firm. Funds invested in the firm's risky technology grow at a stochastic rate Π between the two periods, at the end of which, all the firm's assets are distributed to the remaining investors. The investors' cash will grow deterministically at a risk-free bank account rate. The investors choose their level of investment in the firm in order to maximise their utility. In turn, the firm decides on a level of asset redistribution to investors, either through dividends, or through buybacks. Dividends impact the firm's valuation as they alter the investment in the firm's risky technology. The firm's repurchases intentions change the stock's demand: buybacks modify both the market clearing condition and the scope of the investment in the risky technology. We compare the equilibria of the model depending on which redistribution methods is implemented.

Section 1 formally presents the baseline model, with a representative investor and exogenous risk-free rate. We derive the complete optimal equilibrium solution of the model with dividends strategies. With share repurchases implemented instead, we first prove that a single optimal repurchases strategy exists. Using the proved equivalence between the two redistribution modes, described in Section 2, we then exhibit the unique repurchases optimal strategy.

As expected, redistribution increase with risk aversion and volatility. In accordance with the Modigliani-Miller theorem, the structure of redistribution is irrelevant to investors in terms of maximising their utility (Modigliani and Miller, 1961, 1958, 1963). Section 2 contains the formal proof of dividends and buybacks' economic equivalence in this setting. The proof is also constructive, allowing us to find the analytical behaviour

of the share repurchases at equilibrium from the dividends' one.

Yet, this equivalence between the two modes of redistribution only stands when solely taking investor utilities into account. Other model quantities are not equivalent under the two setups at equilibrium. Even in the most straightforward setting, with one representative investor, the resulting firm valuation differs.

Using share buybacks, the firm's price remains stable for varying investor risk-aversion. Indeed, the redistributed volume increases with the risk-aversion as investors put a heightened premium over the risk carried by the firm's stock. Redistribution being executed through repurchases, the higher risk-aversion induces larger demand for the firm's equity. Thus, the firm's price is maintained thanks to the use of share repurchases. Practically, in this model, the reduction in the risky firm's value when increasing risk-aversion exactly equals the price augmentation prompted by bigger equity repurchases volumes: the optimal firm's price remains constant for varying risk-aversion levels.

Whereas, using dividends, the firm's price decreases rapidly with increasing risk-aversion from investors, despite the firm spending the same amount on redistribution in both setups. Similarly, choosing to execute repurchases instead of dividends accrues the earnings-per-share ratio of the company.

Section 3 extends the baseline model in two directions in which the repurchases-dividends equivalence is maintained.

First, we show the economic equivalence of dividends and buybacks is robust to heterogeneous investors (*see Section 3.1*). We introduce heterogeneous investors: they disagree on the firm's prospects for growth, some being optimistic and others pessimistic. As such, one category of investors would ideally see more of the firm's assets be redistributed back to investors than the other. While the optimists naturally have a higher expected utility than pessimists, their respective expected utilities are still independent of the redistribution method used. The ratio of cash spent by the firm on redistribution is also independent of the redistribution mode used, and the same directional observations can be made as for the baseline model.

The second extension to the baseline model endogenizes the risk-free rate and allows for multi-period consumption of income by the investors (*see Section 3.2*). We show that buybacks and dividends also remain equivalent in this setting. This equivalence extends to the resulting risk-free rate: the risk-free rate obtained is identical using the two methods. The optimal market capitalisation of the firm still favors repurchases in all typical settings. In the risk-free endogenization method presented, the firm takes in consideration the impact of its policy of the risk-free rate. We show that our results are robust to the risk-free endogenization process used, whether the firm takes into account its impact on the risk-free rate or not (*see Section 2*).

Finally, Section 4 outlines two model extensions that break the equivalence between share repurchases and dividends. The previous sections, while maintaining the dividends-buybacks equivalence, yielding identical investor utility, show that buybacks, but not dividends, allow for the management of the firm’s share price and earnings-per-share ratio. Contrasting with these conclusions, Section 4.1 exhibits an extension which favors dividends over share repurchases.

Many equity positions are not intended to be managed as liquid position. Various types of investors will not sell their stake in a firm, independently of their expected utilities. Some, such as pension funds, insurers and re-insurers have prohibitive legal requirements to manage when rebalancing portfolios. ETF providers are legally bounded to hold certain stocks. Some other investors can only trade during limited blackout periods. Also, investors whose share holdings carry corporate control of the firm will not take part in repurchase plans as it would but their majority voting rights in jeopardy.

Introducing an investor whose stake in the firm is locked confirms that dividends are beneficial to this type of investor. Buybacks do not allow these investors to get any cash redistributed at time 0. In this setting, dividends are beneficial to the “locked” investor but to other regular investors too. Indeed, the firm maximises the aggregated investors’ utility. Since the “locked” investor is unable to sell his shares directly, the only way for the firm to reduce the risk exposure of the “locked” investor is to alter its behaviour regarding the regular investor too. In all typical settings, dividends strategies strictly dominate repurchases ones.

In more extreme scenarios, the uniqueness of the optimal buybacks strategy is not guaranteed. We numerically exhibit the two equilibria and discuss their impact.

These results confirm that when multiple type of investors hold a firm’s equity, dividends are beneficial utility-wise to all investors, and not only to investors who will not sell.

Finally, Section 4.2 investigates the endowment of stock options to the firm’s management. The legal obligation of management to act in the investors’ best interest is maintained. In this scenario, we prove that deciding to use share repurchases always benefits the executives, given that the optimal redistribution strategies do not result in the firm emitting new shares or liquidating the totality of its assets.

1 SETUP OF THE BASELINE MODEL

We work with a discrete-time model with time points indexed by $t \in \{-1, 0, 1\}$. At $t = -1$, a representative investor is endowed with some cash and shares of a firm. At $t = 0$, the firm distributes some of its equity back to the investor in form of dividend payout

and/or shares buyback. Immediately after the distribution, the investor rebalances their portfolio across the risky holdings in the firm and the cash position in a risk-free bank account. Meanwhile, all residual capital of the firm post-distribution is invested in a risky technology which grows stochastically between $t = 0$ and $t = 1$. Finally, the firm is liquidated at $t = 1$ and its entire outstanding capital is paid to the investor in form of a lump sum cash dividend.

1.1 THE FIRM

At $t = 0$ just before dividend payout and shares buyback, the firm has an initial capital of $K_0 > 0$ and the number of outstanding shares is assumed to be unity. The market price of each share of the firm is denoted by P_0 .

DISTRIBUTION VIA DIVIDEND AND SHARES BUYBACK

A distribution strategy is defined as a pair $(\theta_D, \theta_{BB}) \in \mathbb{R}^2$. θ_D denotes the dividend payout ratio, where a cash dividend amount of $\theta_D K_0$ per unit share is paid to the shareholders. θ_{BB} represents the number of shares to be repurchased by the firm. We assume the shares buyback is performed at the prevailing market price P_0 , and that this operation is entirely financed by the firm's capital. Since the price at which repurchases are executed is set through market clearing, all homogeneous investors are willing to participate in the buybacks plan, selling θ_{BB} of their optimal investment target in the firm. Rebalancing and repurchases occur simultaneously on the same exchange. In turn, the firm's level of repurchases influences the market price P_0 as it impacts the global offer and demand balance for the firm's equity. Recalling that the total number of outstanding shares is normalized to one, the total equity value distributed back to the representative investor is then $\theta_D K_0 + \theta_{BB} P_0$. After the distribution, the residual capital at the firm becomes $K_0 - \theta_D K_0 - \theta_{BB} P_0$, while the number of outstanding shares is reduced to $1 - \theta_{BB}$.

Throughout the paper, we will refer to a distribution strategy in form of $(\theta_D, \theta_{BB} = 0)$ as a (pure) dividend strategy. Likewise, $(\theta_D = 0, \theta_{BB})$ is referred to as a (pure) buyback strategy.

THE FIRM'S RISKY TECHNOLOGY AND LIQUIDATION PAYOUT

The firm has access to a risky technology such that each unit of available capital at $t = 0$ will grow stochastically to Π at $t = 1$. In our baseline model, we assume that $\Pi \sim N(\mu, \sigma^2)$ with some given $\mu \in \mathbb{R}$, $\sigma > 0$. In an extension presented in Section 3.1, we will consider a case featuring investors with heterogeneous beliefs over the mean return of the risky technology.

Since the starting capital of the firm immediately after distribution is $K_0 - \theta_D K_0 -$

$\theta_{BB}P_0$, the terminal net worth of the firm at $t = 1$ is hence $(K_0 - \theta_D - \theta_{BB}P_0)\Pi$. In turn, the liquidation dividend per unit share at $t = 1$ is

$$D_1 := \frac{(K_0 - \theta_D - \theta_{BB}P_0)\Pi}{1 - \theta_{BB}}.$$

Remark 1.1 (Firm liquidation at time 0). *The firm is endowed with assets K_0 at time $t = -1$. By setting θ_D to 1, the firm redistributes all its assets to investors, and is thus liquidated. When the risk-adjusted growth rate of the firm in excess of the risk-free rate is low enough, owning any stake in the company is not beneficial and thus the optimal behaviour is for the whole firm to be liquidated.*

When liquidated through dividends, investors naturally receive K_0 from the firm's liquidation.

When the firm buys back the entirety of the existing share supply, we ensure that the resulting spending equals precisely K_0 (see Definition 1.6). That is, the market price for the liquidated company equals the value of its assets at the time of liquidation.

1.2 THE REPRESENTATIVE INVESTOR

Each shareholder of the firm is initially endowed with some cash and shares at $t = -1$. Without loss of generality in our friction-less setup with exponential utility, the population is summarized by a single representative investor (see Remark 1.4) with initial cash endowment of W_{-1} and their share holding is equal to the number of the outstanding shares, which we have normalized to one.

INVESTOR'S PORTFOLIO DYNAMICS

When the firm distributes equity at $t = 0$, the representative investor receives a total cash dividend of $\theta_D K_0$. In addition, we assume all investors fully participate in the buyback program such that the representative investor sells θ_{BB} units of shares to the firm at price level P_0 . Consequently, the investor's cash wealth immediately after the firm's distribution is $W_{-1} + \theta_D K_0 + \theta_{BB} P_0$ and they are left with $1 - \theta_{BB}$ units of shares. The investor is allowed to rebalance their positions at the same share price level P_0 . The share repurchases and rebalancing happen simultaneously at market price P_0 . Suppose they choose θ_0 as the number of shares to be held after the rebalancing. The total demand for shares then is $(\theta_0 + \theta_{BB})$, i.e. the investor opts to sell $1 - \theta_{BB} - \theta_0$ shares. At the end of $t = 0$, the investor's portfolio consists of cash wealth of

$$W_0 := W_{-1} + \theta_D K_0 + \theta_{BB} P_0 + (1 - \theta_{BB} - \theta_0) P_0 = W_{-1} + \theta_D K_0 + (1 - \theta_0) P_0$$

and θ_0 units of risky share. The cash wealth W_0 grows at a risk-free rate of $R_f > 0$. We assume R_f to be exogenously fixed in the baseline model, but we will consider an extension in Section 3.2 where R_f will eventually be endogenized.

At $t = 1$, the firm pays a liquidation dividend of D_1 per unit share. Hence the terminal portfolio value of the investor is given by

$$\begin{aligned} W_1 &:= W_0(1 + R_f) + \theta_0 D_1, \\ &= (W_{-1} + \theta_D K_0 + (1 - \theta_0)P_0)(1 + R_f) + \theta_0 \frac{(K_0 - \theta_D K_0 - \theta_{BB} P_0)\Pi}{1 - \theta_{BB}}. \end{aligned}$$

INVESTOR'S UTILITY FUNCTION AND OPTIMIZATION PROBLEM

Suppose the investor has an exponential utility function given by

$$U(x) := \frac{1 - \exp(-\gamma x)}{\gamma},$$

where $\gamma > 0$ is their constant absolute risk aversion coefficient. The certainty equivalent associated with the terminal portfolio value under a given trading strategy θ_0 , the market price of the firm P_0 and the distribution strategy (θ_D, θ_{BB}) is

$$\begin{aligned} J(\theta_0; P_0, \theta_D, \theta_{BB}) &:= U^{-1}(\mathbb{E}[U(W_1)]), \\ &= -\frac{1}{\gamma} \ln \left\{ 1 - \gamma \mathbb{E} \left[\exp \left(-\gamma \left((W_{-1} + \theta_D K_0 + (1 - \theta_0)P_0)(1 + R_f) \right. \right. \right. \right. \\ &\quad \left. \left. \left. + \theta_0 \frac{(K_0 - \theta_D K_0 - \theta_{BB} P_0)\Pi}{1 - \theta_{BB}} \right) \right) \right] \right\}, \\ &= (W_{-1} + \theta_D K_0 + (1 - \theta_0)P_0)(1 + R_f) \\ &\quad + \left\{ \frac{\theta_0(K_0 - \theta_D K_0 - \theta_{BB} P_0)\mu}{1 - \theta_{BB}} - \frac{\sigma^2 \gamma \theta_0^2 (K_0 - \theta_D K_0 - \theta_{BB} P_0)^2}{2(1 - \theta_{BB})^2} \right\}. \end{aligned}$$

1.3 COMPETITIVE EQUILIBRIUM UNDER EXOGENOUSLY GIVEN DISTRIBUTION STRATEGY

We now introduce the concept of equilibrium in the economy under an exogenously given distribution strategy of the firm.

Definition 1.2. *Given the firm's distribution strategy (θ_D, θ_{BB}) , a competitive equilibrium is a pair of the investor's trading strategy and market price of the firm $(\hat{\theta}_0, \hat{P}_0)$ such that:*

1. *The representative investor maximizes the certainty equivalent of their terminal*

portfolio value, i.e.

$$J(\hat{\theta}_0; \hat{P}_0, \theta_D, \theta_{BB}) = \sup_{\theta_0} J(\theta_0; \hat{P}_0, \theta_D, \theta_{BB}).$$

2. The market clears, i.e. the total demand for the shares is equal to the number of outstanding shares post-buyback, or equivalently

$$1 = \hat{\theta}_0 + \theta_{BB}.$$

We refer to $\hat{J}(\theta_D, \theta_{BB}) := J(\hat{\theta}_0; \hat{P}_0, \theta_D, \theta_{BB})$ as an equilibrium value function under the distribution strategy (θ_D, θ_{BB}) . We further define $\hat{J}_D(\theta_D) := \hat{J}(\theta_D, \theta_{BB} = 0)$ and $\hat{J}_{BB}(\theta_{BB}) := \hat{J}(\theta_D = 0, \theta_{BB})$ as equilibrium value functions under a (pure) dividend strategy and a (pure) buyback strategy respectively.

In this chapter, we are interested in the two polar cases of the distribution strategy: dividend strategy $(\theta_D, \theta_{BB} = 0)$ and buyback strategy $(\theta_D = 0, \theta_{BB})$. The two propositions below characterize the forms of the competitive equilibrium under these two cases.

Proposition 1.3 (Equilibrium under exogenously given dividend strategy). *If the firm adopts a dividend distribution strategy in form of $(\theta_D, \theta_{BB} = 0)$, the competitive equilibrium is uniquely given by*

$$\hat{\theta}_0 = 1, \quad \hat{P}_0 = \frac{K_0(1 - \theta_D)}{1 + R_f} (\mu - \sigma^2 \gamma K_0(1 - \theta_D)). \quad (1.1)$$

Proof. We work in a pure dividends strategy scenario $(\theta_{BB} = 0)$.

Thus, the investor's expected utility $J(\theta_0; P_0, \theta_D, \theta_{BB})$ becomes:

$$\begin{aligned} J(\theta_0; P_0, \theta_D, 0) = & (W_{-1} + \theta_D K_0 + (1 - \theta_0) P_0)(1 + R_f) \\ & + \left\{ \theta_0 K_0(1 - \theta_D) \mu - \frac{\sigma^2 \gamma}{2} \theta_0^2 (K_0(1 - \theta_D))^2 \right\}. \end{aligned}$$

$J(\theta_0; P_0, \theta_D, \theta_{BB})$ is a concave quadratic function of θ_0 , it is thus enough to consider the first-order condition $(\partial_{\theta_0} J(\theta_0; P_0, \theta_D, \theta_{BB}) = 0)$ to characterize the optimal θ_0 .

We first determine the optimal investor ownership after redistribution $\hat{\theta}_0$, given the company price P_0 and dividends spending $(\theta_D \cdot K_0)$. The optimal ownership ratio maximises the investor's total expected utility, from both his holding in the firm and his bank account's growth. Thus it satisfies:

$$\frac{\partial J(\hat{\theta}_0; P_0, \theta_D, 0)}{\partial \theta_0} = 0.$$

That is,

$$0 = -P_0(1 + R_f) + K_0(1 - \theta_D)\mu - \hat{\theta}_0\gamma\sigma^2(K_0(1 - \theta_D))^2.$$

Thus,

$$\begin{aligned}\hat{\theta}_0 &= \frac{K_0(1 - \theta_D)\mu - P_0(1 + R_f)}{\gamma\sigma^2(K_0(1 - \theta_D))^2}, \\ &= \frac{1}{\gamma\sigma^2 K_0(1 - \theta_D)} \left(\mu - \frac{P_0(1 + R_f)}{K_0(1 - \theta_D)} \right).\end{aligned}$$

Note that we used the $(1 - \theta_D) \neq 0$ constraint (i.e. the firm is allowed to use its own cash to go in private ownership), we have that θ_0 is always well-defined as γ, σ, K_0 are all strictly positive.

We determine $\hat{P}_0$, the company's equilibrium price at $t = 0$, by market clearing. That is, the price is set so that the total aggregated demand for company equity from investors equals the available supply of shares. Since we have one representative investor, the total demand for shares is $\hat{\theta}_0$. Without loss of generality, the total number of shares is normalized to 1.

Consequently, the price is set to satisfy $\hat{\theta}_0 = 1$, giving

$$\hat{P}_0 = \frac{K_0(1 - \theta_D)}{1 + R_f} \left(\mu - \sigma^2 \gamma K_0(1 - \theta_D) \right).$$

□

Remark 1.4. *The demand in company shares incurred by N indistinguishable investors is identical to the demand for shares initiated by the representative investor:*

$$\hat{\theta}_0 = \sum_{n=1}^N \hat{\theta}_n.$$

Where $\hat{\theta}_n$ is defined as the demand for company stock incurred by individual investor n , with risk aversion γ_n .

From this, the resulting company price at equilibrium does not depend on whether individual investors are aggregated as an representative one or not. From this, the optimal equilibrium is identical using a the representative investor or the original indistinguishable ones.

Proof. We work in the dividends setting, the buybacks one is analogous.

As in the previous proof, we get:

$$\begin{aligned} \frac{\partial J_n(\hat{\theta}_n, P_0, \theta_D, \theta_{BB})}{\theta_n} &= 0, \\ \iff \hat{\theta}_n &= \frac{1}{\gamma_n \sigma^2 K_0 (1 - \theta_D)} \left(\mu - \frac{P_0(1 + R_f)}{K_0(1 - \theta_D)} \right). \end{aligned}$$

Thus, using $\forall n \in \llbracket N \rrbracket$, $\gamma_n = \frac{1}{N} \gamma$:

$$\sum_{n=1}^N \hat{\theta}_n = N \hat{\theta}_N = \hat{\theta}_0.$$

□

Proposition 1.5 (Equilibrium under exogenously given buyback strategy).

Define

$$\Delta = \Delta(\theta_{BB}) := [\mu \theta_{BB} + (1 + R_f)(1 - \theta_{BB})]^2 - 4\sigma^2 \gamma (1 + R_f) K_0 \theta_{BB} (1 - \theta_{BB}). \quad (1.2)$$

If the firm adopts a buyback strategy in form of $(\theta_D = 0, \theta_{BB})$ with $\theta_{BB} \neq 0$. Then:

1. *If $\Delta \geq 0$, there exists at most two competitive equilibria. Moreover, for any competitive equilibrium $(\hat{\theta}_0, \hat{P}_0)$, we must have $\hat{\theta}_0 = 1 - \theta_{BB}$ and*

$$\hat{P}_0 \in \left\{ \frac{2\sigma^2 \gamma \theta_{BB} K_0 - \mu \theta_{BB} - (1 + R_f)(1 - \theta_{BB}) \pm \sqrt{\Delta(\theta_{BB})}}{2\sigma^2 \gamma \theta_{BB}^2} \right\}. \quad (1.3)$$

2. *If $\Delta < 0$ instead, there does not exist any competitive equilibrium.*

Proof. We work in a pure buybacks strategy scenario ($\theta_D = 0$). The investor's total expected utility becomes:

$$\begin{aligned} J(\theta_0; P_0, 0, \theta_{BB}) &= (W_{-1} + (1 - \theta_0)P_0)(1 + R_f) \\ &\quad + \left\{ \mu \theta_0 \frac{K_0 - \theta_{BB} P_0}{1 - \theta_{BB}} - \frac{\sigma^2 \gamma}{2} \theta_0^2 \left(\frac{K_0 - \theta_{BB} P_0}{1 - \theta_{BB}} \right)^2 \right\}. \end{aligned}$$

To obtain $\hat{\theta}_0$, as in the pure dividends setting, we use:

$$\frac{\partial J(\hat{\theta}_0; P_0, 0, \theta_{BB})}{\partial \theta_0} = 0.$$

Thus,

$$0 = -P_0(1 + R_f) + \mu \frac{K_0 - \theta_{BB}P_0}{1 - \theta_{BB}} - \sigma^2 \gamma \hat{\theta}_0 \left(\frac{K_0 - \theta_{BB}P_0}{1 - \theta_{BB}} \right)^2.$$

We obtain

$$\hat{\theta}_0 = \frac{1 - \theta_{BB}}{\gamma \sigma^2 (K_0 - \theta_{BB}P_0)} \left(\mu - \frac{P_0(1 + R_f)(1 - \theta_{BB})}{K_0 - \theta_{BB}P_0} \right).$$

As in the pure dividends setting, the equilibrium price as a function of θ_{BB} is obtained through market clearing. That is, $\hat{P}_0$ is set such that the demand for company shares equals the share supply. Contrasting with the pure dividends setting however, the total demand for shares at time $t = 0$ is driven by both the investor through θ_0 , but also by the repurchases initiated by the company, of volume θ_{BB} . Since the available supply of shares is normalized to 1, the market clearing condition for $\hat{P}_0$ becomes:

$$\hat{\theta}_0 + \theta_{BB} = 1.$$

Equivalently,

$$\hat{\theta}_0 = 1 - \theta_{BB}.$$

That is:

$$1 - \theta_{BB} = \frac{1 - \theta_{BB}}{\gamma \sigma^2 (K_0 - \theta_{BB}\hat{P}_0)} \left(\mu - \frac{\hat{P}_0(1 + R_f)(1 - \theta_{BB})}{K_0 - \theta_{BB}\hat{P}_0} \right).$$

Which gives

$$0 = \mu(K_0 - \theta_{BB}\hat{P}_0) - \hat{P}_0(1 + R_f)(1 - \theta_{BB}) - \gamma \sigma^2 (K_0 - \theta_{BB}\hat{P}_0)^2.$$

$\hat{P}_0$ as a function of $(\theta_{BB}, R_f, \gamma, \mu, \sigma, K_0)$ solves a quadratic equation of the following form:

$$a \cdot \hat{P}_0^2 + b \cdot \hat{P}_0 + c = 0,$$

With coefficients,

$$\begin{aligned} a &= \theta_{BB}^2 \gamma \sigma^2, \\ b &= \mu \theta_{BB} + (1 - \theta_{BB})(1 + R_f) - 2 \theta_{BB} K_0 \gamma \sigma^2, \\ c &= \gamma \sigma^2 K_0^2 - \mu K_0. \end{aligned}$$

This polynomial's determinant is given by:

$$\Delta = (\mu \theta_{BB} + (1 - \theta_{BB})(1 + R_f))^2 - 4 \cdot \theta_{BB} K_0 \gamma \sigma^2 (1 - \theta_{BB})(1 + R_f)$$

So, when Δ is non-negative, we obtain the following for $\hat{P}_0$:

$$\hat{P}_0 \in \left\{ \frac{2 \theta_{BB} K_0 \gamma \sigma^2 - \mu \theta_{BB} - (1 - \theta_{BB})(1 + R_f) \pm \sqrt{\Delta}}{2 \gamma \sigma^2 \theta_{BB}^2} \right\}.$$

□

In view of Proposition 1.5, the existence of an equilibrium under buyback depends crucially on the quantity Δ which in turn depends on the given distribution strategy of the firm as well as other model parameters. Moreover, there are typically two valid equilibria. It is useful to investigate the behavior of the equilibrium prices to determine which of the two equilibria is more economically sensible.

Definition 1.6. *A competitive equilibrium $(\hat{\theta}_0(\theta_{BB}), \hat{P}_0(\theta_{BB}))$ under a pure buyback strategy $(\theta_D = 0, \theta_{BB})$ is said to be “consistent” if:*

$$\lim_{\theta_{BB} \rightarrow 0} \theta_{BB} \hat{P}_0(\theta_{BB}) = 0. \quad (1.4)$$

$$\lim_{\theta_{BB} \rightarrow 1} \theta_{BB} \hat{P}_0(\theta_{BB}) = K_0. \quad (1.5)$$

In the following, we guarantee that the equilibrium we study are consistent to avoid incoherent behaviour as the ratios bought back approach 0 or 1.

Indeed, (1.5) ensures that, when the company implements share repurchases in order to liquidate the company, then as θ_{BB} nears retiring the whole share pool, the amount spent also tends towards the entirety of the firm's assets K_0 (see Remark 1.1). The original shareholders receive all the firm's assets through redistribution, no assets are left without ownership, owned by a firm with no shareholders.

Analogously, (1.4) entails that as the ratio of shares bought back decreases, the amount spent on repurchases tends to 0. That is, the amount spent to buy marginally few shares is also marginal. It also implies that as θ_{BB} tends towards 0, the representative investor's expected utility also converges towards its original value with no redistribution implemented (see Remark 1.9).

Finally, this definition of consistency implies that given a repurchases ratio, the price at which repurchases are executed relative to the amount of shares repurchased is always finite. Removing (1.4), the value of the infinitesimal amount of shares actually bought back would be non-zero, yielding an infinite amount per share bought back. Conversely, without (1.5), the negligible amount of shares left in the firm after repurchases, as θ_{BB}

reaches 1, would hold a non-negligible value. The resulting value per unit of share would thus become infinite.

While the amount spent as well as the relative value paid per share is always finite, the book value of the whole firm as θ_{BB} reaches 0 or 1 is not. Indeed, as θ_{BB} tends to 0, $\hat{P}_0(\theta_{BB})$ gets arbitrarily high. This assigns an infinite book value to the entire company. However, as no trades happen other than repurchases, of finite volume, this book value is not realised by investors and is not benefited from.

To ensure existence of equilibria (equivalently $\Delta \geq 0$ for Δ defined in (1.2)), we further impose the following standing assumption on the model parameters which we enforce throughout the rest of this paper.

Standing Assumption 1.7. *The model parameters are assumed to satisfy*

$$\mu \geq K_0 \gamma \sigma^2.$$

The assumption implies that per dollar invested in the firm, its expected growth rate μ outweighs the risk this investment carried, $K_0 \gamma \sigma^2$.

Corollary 1.8. *A unique consistent competitive equilibrium exists under any fixed buyback strategy θ_{BB} and is given by*

$$\hat{\theta}_0 = 1 - \theta_{BB}, \quad \hat{P}_0 = \frac{2\sigma^2 \gamma \theta_{BB} K_0 - \mu \theta_{BB} - (1 + R_f)(1 - \theta_{BB}) + \sqrt{\Delta(\theta_{BB})}}{2\sigma^2 \gamma \theta_{BB}^2}. \quad (1.6)$$

We outlined two equilibria (see Proposition 1.5) for a given pure buybacks strategy. We focus on the consistent equilibrium for the rest of this study.

Proof. Using Proposition 1.5, and under Assumption 1.7, we know there are at most two potential competitive equilibria in the pure repurchases setting, given by:

$$\begin{aligned} 1. \quad & \hat{\theta}_0 = 1 - \theta_{BB}. \\ 2. \quad & \hat{P}_0 = \begin{cases} \frac{2\sigma^2 \gamma \theta_{BB} K_0 - \mu \theta_{BB} - (1 + R_f)(1 - \theta_{BB}) + \sqrt{\Delta(\theta_{BB})}}{2\sigma^2 \gamma \theta_{BB}^2}, \\ \frac{2\sigma^2 \gamma \theta_{BB} K_0 - \mu \theta_{BB} - (1 + R_f)(1 - \theta_{BB}) - \sqrt{\Delta(\theta_{BB})}}{2\sigma^2 \gamma \theta_{BB}^2}. \end{cases} \end{aligned}$$

Let $\hat{P}_0^+$ and $\hat{P}_0^-$ respectively be the two expressions for $\hat{P}_0$, and let us check the consistency conditions.

For $\hat{P}_0^+$, with the help of L'Hôpital's rule, as intended we get first that $\hat{P}_0^+$ satisfies

(1.4):

$$\lim_{\theta_{BB} \rightarrow 0} \theta_{BB} \hat{P}_0^+(\theta_{BB}) = K_0 + \lim_{\theta_{BB} \rightarrow 0} \frac{(1 + R_f) - \mu + \frac{\partial_{\theta_{BB}} \Delta(\theta_{BB})}{2\sqrt{\Delta(\theta_{BB})}}}{2\sigma^2\gamma\theta_{BB}}.$$

From Δ 's definition,

$$\Delta(\theta_{BB}) := [\mu\theta_{BB} + (1 + R_f)(1 - \theta_{BB})]^2 - 4\sigma^2\gamma(1 + R_f)K_0\theta_{BB}(1 - \theta_{BB}),$$

we have,

$$\lim_{\theta_{BB} \rightarrow 0} \Delta(\theta_{BB}) = (1 + R_f)^2,$$

and

$$\partial_{\theta_{BB}} \Delta(\theta_{BB}) = 2[\mu\theta_{BB} + (1 + R_f)(1 - \theta_{BB})](\mu - (1 + R_f)) - 4\sigma^2\gamma(1 + R_f)K_0(1 - 2\theta_{BB}).$$

So,

$$\begin{aligned} \lim_{\theta_{BB} \rightarrow 0} \theta_{BB} \hat{P}_0^+(\theta_{BB}) &= K_0 + \frac{(1 + R_f) - \mu + \frac{2(\mu - (1 + R_f))(1 + R_f) - 4\sigma^2\gamma(1 + R_f)K_0}{2(1 + R_f)}}{2\sigma^2\gamma}, \\ &= K_0 - \frac{2\sigma^2\gamma K_0}{2\sigma^2\gamma} = 0. \end{aligned}$$

We obtain (1.5) for $\hat{P}_0^+$ directly:

$$\lim_{\theta_{BB} \rightarrow 1} \theta_{BB} \hat{P}_0^+(\theta_{BB}) = K_0 + \frac{-\mu + \sqrt{\mu^2}}{2\sigma^2\gamma} = K_0.$$

$\hat{P}_0^-$ is not consistent however, as (1.4) is not satisfied:

$$\begin{aligned} \lim_{\theta_{BB} \rightarrow 0} \theta_{BB} \hat{P}_0^-(\theta_{BB}) &= K_0 + \lim_{\theta_{BB} \rightarrow 0} \frac{(1 + R_f) - \mu - \frac{\partial_{\theta_{BB}} \Delta(\theta_{BB})}{2\sqrt{\Delta(\theta_{BB})}}}{2\sigma^2\gamma\theta_{BB}}, \\ &= K_0 + \frac{(1 + R_f) - \mu - \frac{2(\mu - (1 + R_f))(1 + R_f) - 4\sigma^2\gamma(1 + R_f)K_0}{2(1 + R_f)}}{2\sigma^2\gamma}, \\ &= 2K_0 - \frac{\mu - (1 + R_f)}{\sigma^2\gamma}. \end{aligned}$$

Condition (1.5) is not satisfied either:

$$\lim_{\theta_{BB} \rightarrow 1} \theta_{BB} \hat{P}_0^-(\theta_{BB}) = K_0 + \frac{-\mu - \sqrt{\mu^2}}{2\sigma^2\gamma} = K_0 - \frac{\mu}{\sigma^2\gamma}.$$

We obtain that the competitive equilibrium $(\hat{\theta}_0 = 1 - \theta_{BB}, \hat{P}_0^+)$ is consistent while the

only other competitive equilibrium $(\hat{\theta}_0 = 1 - \theta_{BB}, \hat{P}_0^-)$ is not consistent.

□

Remark 1.9 (Model Behaviour as $\theta_{BB} \rightarrow 0$).

When $\theta_{BB} = 0$, the firm does not redistribute any of its assets to investors. It invests all its endowed assets in the risky technology. Since there is a single representative investor, no share is traded at time $t = 0$. Without redistribution, the resulting equilibrium expected utility, optimal ownership and market capitalisation are:

$$\hat{\theta}_0 = 1, \hat{P}_0 = K_0 \frac{\mu - \sigma^2 \gamma K_0}{1 + R_f}.$$

And,

$$J(\hat{\theta}_0; \hat{P}_0, 0, 0) = \hat{J}(0, 0) = (1 + R_f)W_{-1} + K_0\mu - \frac{\sigma^2 \gamma}{2} K_0^2.$$

As the amount of shares bought back, θ_{BB} , tends to 0, the equilibrium investor expected utility converges to $\hat{J}(0, 0)$:

$$\lim_{\theta_{BB} \rightarrow 0} \hat{J}(\theta_{BB}, 0) = \hat{J}(0, 0).$$

Remark 1.9 ensures that the model at equilibrium behaves smoothly around $\theta_{BB} = 0$. Specifically, that the investor's expected utility converge to $J(0, 0)$ as $\theta_{BB} \rightarrow 0$ is only ensured since $(\hat{\theta}_0, \hat{P}_0)$ is a consistent equilibrium.

Proof of Remark 1.9.

$$J(\theta_0; P_0, 0, 0) = (W_{-1} + (\theta_{-1} - \theta_0)P_0)(1 + R_f) + \theta_0 K_0 \mu - \frac{\sigma^2 \gamma}{2} \theta_0^2 K_0^2.$$

We obtain the expressions for $\hat{\theta}_0$ and $\hat{P}_0$ by replacing θ_{BB} by 0 in their characterization in the previous Corollary 1.8. We then obtain

$$J(\hat{\theta}_0; \hat{P}_0, 0, 0) = \hat{J}(0, 0) = (1 + R_f)W_{-1} + K_0\mu - \frac{\sigma^2 \gamma}{2} K_0^2.$$

Since $(\hat{\theta}_0, \hat{P}_0)$ is a consistent equilibrium, we have

$$\hat{\theta}_0 = 1 - \theta_{BB},$$

and

$$\lim_{\theta_{BB} \rightarrow 0} \theta_{BB} \hat{P}_0(\theta_{BB}) = 0.$$

Thus, we get the desired limit for $\hat{J}(\theta_{BB}, 0)$ as θ_{BB} tends to 0:

$$\begin{aligned}
\lim_{\theta_{BB} \rightarrow 0} \hat{J}(\theta_{BB}, 0) &= \lim_{\theta_{BB} \rightarrow 0} \left[(1+R_f)(W_{-1} + (1 - \hat{\theta}_0)\hat{P}_0(\theta_{BB})) + \mu \hat{\theta}_0 \frac{K_0 - \theta_{BB}\hat{P}_0(\theta_{BB})}{1 - \theta_{BB}} \right. \\
&\quad \left. - \frac{\gamma\sigma^2}{2} \hat{\theta}_0^2 \left(\frac{K_0 - \theta_{BB}\hat{P}_0(\theta_{BB})}{1 - \theta_{BB}} \right)^2 \right], \\
&= \lim_{\theta_{BB} \rightarrow 0} \left[(1+R_f)(W_{-1} + \theta_{BB}\hat{P}_0(\theta_{BB})) \right] + \mu K_0 - \frac{\gamma\sigma^2}{2} K_0^2, \\
&= W_{-1}(1 + R_f) + \mu K_0 - \frac{\gamma\sigma^2}{2} K_0^2, \\
&= \hat{J}(0, 0).
\end{aligned}$$

□

Remark 1.10 (Economic Characterization of $\hat{P}_0$).

Market clearing prices $\hat{P}_0$, consistent or not, are characterized through the following equation,

$$1 - \theta_{BB} = \frac{1 - \theta_{BB}}{\gamma\sigma^2(K_0 - \theta_{BB}\hat{P}_0)} \left(\mu - \frac{\hat{P}_0(1 + R_f)(1 - \theta_{BB})}{K_0 - \theta_{BB}\hat{P}_0} \right),$$

Rewriting this condition, $\hat{P}_0$ solves,

$$(1 + R_f)(\hat{P}_0(1 - \theta_{BB})) = \mu(K_0 - \theta_{BB}\hat{P}_0) - \gamma\sigma^2(K_0 - \theta_{BB}\hat{P}_0)^2.$$

The left hand side corresponds to the market value, expected at time $t = 1$, of the company after repurchases. Indeed, after repurchases are executed, $(1 - \theta_{BB})$ shares are left in the company, at current last executed price $\hat{P}_0$. As the bank account grows at risk-free rate R_f , $(1 + R_f)(\hat{P}_0(1 - \theta_{BB}))$ represents the market value of the firm after repurchases are implemented, at time 1.

The right-hand side represents the risk-adjusted expected value of holding all the firm's shares after redistribution.

Rewriting the equation characterising $\hat{P}_0$ again,

$$\hat{P}_0(1 - \theta_{BB}) = \frac{1}{1 + R_f} \left(\mu(K_0 - \theta_{BB}\hat{P}_0) - \gamma\sigma^2(K_0 - \theta_{BB}\hat{P}_0)^2 \right).$$

That is, $\hat{P}_0$ is set to satisfy that the market capitalisation of the firm after the shares bought back are retired equals the discounted risk-adjusted expected value of the firm at the time of its liquidation, $t = 1$.

Definition 1.11. From the previous Remark 1.10, define MC and EV as functions of

$\hat{P}_0$ and θ_{BB} ,

$$MC(\hat{P}_0, \theta_{BB}) = \hat{P}_0(1 - \theta_{BB}),$$

$$EV(\hat{P}_0, \theta_{BB}) = \frac{1}{1 + R_f} \left(\mu(K_0 - \theta_{BB}\hat{P}_0) - \gamma\sigma^2(K_0 - \theta_{BB}\hat{P}_0)^2 \right).$$

For $\hat{P}_0$ to be a market clearing price given θ_{BB} is equivalent to it satisfying,

$$MC(\hat{P}_0, \theta_{BB}) = EV(\hat{P}_0, \theta_{BB}).$$

Remark 1.12 (Negative Δ Economic Interpretation). We provide an economic characterization of the settings where $\hat{P}_0$ does not exist, that is when $\Delta < 0$.

Assumption 1.7 is equivalent to ensuring $EV(0, \theta_{BB})$ is non-negative. Let us assume that $\theta_{BB} \in (0, 1)$: the firm does not issue shares and the investors' aggregated position is not short.

Given Assumption 1.7 and $\theta_{BB} \in (0, 1)$, we have:

$$\begin{aligned} MC(0, \theta_{BB}) &= 0 \\ MC\left(\frac{K_0}{\theta_{BB}}, \theta_{BB}\right) &\geq 0 \\ EV(0, \theta_{BB}) &\geq 0, \text{ relying on Assumption 1.7,} \\ EV\left(\frac{K_0}{\theta_{BB}}, \theta_{BB}\right) &= 0 \end{aligned}$$

Since EV and MC both are continuous functions of $\hat{P}_0$, by the intermediate value theorem, for any $\theta_{BB} \in (0, 1)$, there exists a positive market clearing price $\hat{P}_0$ such that,

$$MC(\hat{P}_0, \theta_{BB}) = EV(\hat{P}_0, \theta_{BB}).$$

When Assumption 1.7 is not satisfied, we have $EV(0, \theta_{BB}) < 0$. Essentially, the market capitalisation MC of the firm is independent of $(R_f, \mu, \sigma, \gamma, K_0)$ and always rises as $\hat{P}_0$ increases: $\partial_{\hat{P}_0} MC = (1 - \theta_{BB})$.⁶ EV however depends on $(R_f, \mu, \sigma, \gamma, K_0)$. For fixed $(\theta_{BB}, \hat{P}_0)$, as the expected growth of the risky investment μ decreases, $EV(\theta_{BB}, \hat{P}_0)$ is reduced, while $MC(\theta_{BB}, \hat{P}_0)$ is not. In turn, when μ reaches a low enough value ($\mu < \gamma\sigma^2 K_0$), for some values of θ_{BB} , EV is smaller than MC for all values of $\hat{P}_0$.

Thus, $\Delta < 0$ can be interpreted economically as: the firm's intrinsic value as an investment never attains the market value implied by the repurchases, and thus a clearing price cannot be found.

⁶As for EV , we have $\text{sgn}(\partial_{\hat{P}_0} EV) = \text{sgn}(\hat{P}_0 - P_0^*)$, where $P_0^* = \frac{K_0}{\theta_{BB}} - \frac{\mu}{2\gamma\sigma^2\theta_{BB}}$.

1.4 OPTIMALLY ENDOGENIZED DIVIDEND PAYOUT AND BUYBACK STRATEGY

A particular choice of dividend strategy θ_D or buyback strategy θ_{BB} induces a unique pair of equilibrium share price $\hat{P}_0$ and investor's share holding $\hat{\theta}_0$, which in turn uniquely determine the welfare level of the investor measured by their certainty equivalent $\hat{J}_D(\theta_D)$ or $\hat{J}_{BB}(\theta_{BB})$ as defined in Definition 1.2. The executives' redistribution decisions are set in order to maximise shareholder value, measured by their certainty equivalent of the representative investor. That is, we assume that the firm maximizes the investor's welfare by optimally endogenizing the dividend or buyback strategy via solving the maximization problem

$$V_k := \sup_{\theta_k} \hat{J}_k(\theta_k), \quad k \in \{D, BB\}.$$

The corresponding optimizer θ_k^* is said to be the optimal dividend strategy (for $k = D$) or the optimal buyback strategy (for $k = BB$).

A welfare-maximizing dividend strategy and the corresponding equilibrium share price are indeed available in close-form.

Proposition 1.13. *The unique optimal dividend strategy is given by*

$$\theta_D^* = 1 - \frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0}$$

under which the associated equilibrium share price is

$$\hat{P}_0 = \frac{\mu - (1 + R_f)}{\sigma^2 \gamma}.$$

The optimal dividends distribution is inversely proportional to the expected growth of the firm in excess of the risk-free rate ($\mu - (1 + R_f)$), normalized by the per-dollar risk charge of the firm $\sigma^2 \gamma K_0$. In turn, the equilibrium price is proportional to $(1 - \theta_D^*) \cdot K_0$, i.e. to $(\mu - (1 + R_f))$ normalized by the firm's risk adjusted volatility $\sigma^2 \gamma$.

Proof of Proposition 1.13. From Proposition 1.3, we have that the sole competitive equilibrium in the pure dividends setting has price:

$$\hat{P}_0 = \frac{K_0(1 - \theta_D)}{1 + R_f} (\mu - \sigma^2 \gamma K_0(1 - \theta_D)).$$

$$\text{And } \hat{\theta}_0 = 1.$$

Under this competitive equilibrium, the investor's expected utility becomes:

$$\hat{J}_D(\theta_D) = (W_{-1} + \theta_D K_0)(1 + R_f) + \left(K_0(1 - \theta_D)\mu - \frac{\sigma^2 \gamma}{2} (K_0(1 - \theta_D))^2 \right).$$

The level of dividends θ_D is set by the company's executives so that it maximises the investor's expected utility $\hat{J}_D(\theta_D)$.

The expression $\hat{J}_D(\theta_D)$ is a concave quadratic function of θ_D . To characterize the optimal redistribution level θ_D^* , it is thus enough to consider the first order condition

$$\frac{\partial \hat{J}_D(\theta_D^*)}{\partial \theta_D} = 0.$$

That is,

$$0 = \frac{\partial \hat{J}_D(\theta_D^*)}{\partial \theta_D} = K_0(1 + R_f) - K_0\mu + \sigma^2 \gamma K_0(1 - \theta_D^*).$$

Finally,

$$\theta_D^* = 1 - \frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0}.$$

$\hat{P}_0$ then becomes:

$$\begin{aligned} \hat{P}_0 &= \frac{K_0}{1 + R_f} \left(\frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0} \right) \left(\mu - \sigma^2 \gamma K_0 \left(\frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0} \right) \right), \\ &= \frac{\mu - (1 + R_f)}{\sigma^2 \gamma}. \end{aligned}$$

As θ_D^* is the only solution to $\frac{\partial \hat{J}_D(\theta_D^*)}{\partial \theta_D} = 0$, we get that $(\theta_D^*, \hat{P}_0)$ is the only optimal pure dividends strategy. □

Meanwhile, the optimal buyback strategy is not available in close-form but it can still be conveniently characterized as a solution to an explicit equation.

Proposition 1.14. *An optimal buyback ratio θ_{BB}^* always exists and is unique.*

It is given by the unique solution to the scalar equation

$$\theta_{BB}^* \hat{P}_0(\theta_{BB}^*) = K_0 - \frac{\mu - (1 + R_f)}{\sigma^2 \gamma} \quad (1.7)$$

where $\hat{P}_0(\theta_{BB})$ is defined in (1.6).

Furthermore, the unique such θ_{BB}^* equals θ_D^* (see Proposition 2.3).

We first prove the equivalence of the two redistribution methods, and then provide the proof to this statement (*see* 2.3).

2 EQUIVALENCE OF DIVIDEND AND BUYBACK IN THE BASELINE MODEL

We now state the first central result of this paper. The subsequent sections present further extensions to this result, and also provide limitations to its applications.

Theorem 2.1 (Equivalence of dividend and buyback). *Dividend and buyback are economically equivalent in terms of investor's welfare and equity value distributed. Specifically,*

$$V_D = V_{BB}$$

and the value of equity distributed to the investor are identical across dividend and buyback strategy, i.e.

$$K_0 \theta_D^* = \theta_{BB}^* \hat{P}_0(\theta_{BB}^*).$$

We first prove that pure dividends strategies dominate pure share repurchases ones.

$$\begin{aligned} V_D &\geq V_{BB}, \\ \sup_{\theta_D} \hat{J}_D(\theta_D) &\geq \sup_{\theta_{BB}} \hat{J}_{BB}(\theta_{BB}). \end{aligned}$$

The second subsection proves the analogous statement:

$$\begin{aligned} V_D &\leq V_{BB}, \\ \sup_{\theta_D} \hat{J}_D(\theta_D) &\leq \sup_{\theta_{BB}} \hat{J}_{BB}(\theta_{BB}). \end{aligned}$$

We actually show that optimal pure dividends and optimal pure repurchases strategies, while equivalent, also dominate any mixed strategy ($\theta_{BB} > 0, \theta_D > 0$). Indeed, we first prove that pure dividends strategies dominate all mixed strategies. The second part of the shows that pure repurchases strategies dominate pure dividends ones. In turn, both optimal pure dividends and buybacks strategies yield a higher expected utility to investors than mixed strategies combining dividends' distribution and share repurchases.

PURE DIVIDENDS STRATEGIES DOMINATE MIXED DIVIDENDS AND REPURCHASES ONES

For any strategy with both dividends and share repurchases, let us show that one can build a strategy with dividends only that yields the same utility for investors. We call the later a pure dividends strategy and the former a mixed strategy. Using a mixed

strategy, the firm is allowed to distribute dividends and to repurchase its own shares in the same strategy:

$$J(\theta_{BB}, \theta_D) = (1 + R_f)(W_{-1} + \theta_D \theta_{-1} K_0 + (\theta_{-1} - \theta_0) P_0) \\ + \mu \frac{K_0(1 - \theta_D) - P_0 \theta_{BB}}{1 - \theta_{BB}} \theta_0 - \frac{\sigma^2 \gamma}{2} \frac{(K_0(1 - \theta_D) - P_0 \theta_{BB})^2}{(1 - \theta_{BB})^2} \theta_0^2.$$

Assume that the couple $(\theta_{BB}^*, \theta_D^*)$ maximizes $J(\theta_{BB}, \theta_D)$. That is, $V_{BB} = J(\theta_{BB}^*, \theta_D^*)$. In this strategy, the investor has stake θ_0^* in the firm from time $t = 0$ to time $t = 1$. The resulting company price P_0^* .

Without loss of generality, in this setting was set $\theta_{-1} = \frac{\theta_0^*}{1 - \theta_{BB}^*}$ so that the investor does not actively rebalance their portfolio and only accept the repurchase offer from the company. As the expressions for θ_0^* and P_0^* are independent of θ_{-1} , this assumption does not affect the model's equilibrium.

The investor's optimal utility then becomes:

$$J(\theta_{BB}^*, \theta_D^*) = (1 + R_f)(W_{-1} + \theta_{-1} \theta_D^* K_0) \\ + \mu \frac{K_0(1 - \theta_D^*) - P_0^* \theta_{BB}^*}{1 - \theta_{BB}^*} \theta_0^* - \frac{\sigma^2 \gamma}{2} \frac{(K_0(1 - \theta_D^*) - P_0^* \theta_{BB}^*)^2}{(1 - \theta_{BB}^*)^2} \theta_0^{*2}.$$

Let's exhibit a pure dividends strategy θ_D^u which yields $J(0, \theta_D^u) \geq J(\theta_{BB}^*, \theta_D^*)$.

$$\text{Choose } \theta_D^u \text{ such that } K_0 \theta_D^u = K_0 \theta_D^* + P_0^* \theta_{BB}^*.$$

$$\text{Implying } K_0(1 - \theta_D^u) = K_0 - K_0 \theta_D^* - P_0^* \theta_{BB}^*.$$

This implies that the firm spends the same amount on redistribution in both strategies, if one does not require the firm to go into debt, nor will the other. Typically, this ensures $\theta_D^u \in [0, 1]$.

Analogously, suppose the investor chooses his rebalanced equity position $\theta_0^u = \frac{\theta_0^*}{1 - \theta_{BB}^*}$.

This is always possible as, when the investor cannot short the firm, $\theta_0^* \in [0, 1 - \theta_{BB}^*]$ and thus $\theta_0^u \in [0, 1]$. If shorting the stock is allowed, then shorting in the pure dividends

strategy is equivalent to shorting in the mixed dividends & share buybacks strategy.

$$\begin{aligned}
J(0, \theta_D^u) &= (1 + R_f)(W_{-1} + \theta_{-1}\theta_D^u K_0 + (\theta_{-1} - \theta_0^u)P_0) + \mu K_0(1 - \theta_D^u) \theta_0^u \\
&\quad - \frac{\sigma^2 \gamma}{2} (K_0(1 - \theta_D^u))^2 \theta_0^{u2}, \\
&= (1 + R_f) \left(W_{-1} + \theta_{-1}\theta_D^* K_0 + \theta_{-1}P_0^* \theta_{BB}^* + \left(\theta_{-1} - \frac{\theta_0^*}{1 - \theta_{BB}^*} \right) P_0 \right) \\
&\quad + \mu \frac{K_0 - K_0\theta_D^* - P_0^* \theta_{BB}^*}{1 - \theta_{BB}^*} \theta_0^* - \frac{\sigma^2 \gamma}{2} \frac{(K_0 - K_0\theta_D^* - P_0^* \theta_{BB}^*)^2}{(1 - \theta_{BB}^*)^2} \theta_0^{*2}, \\
&= (1 + R_f)(W_{-1} + \theta_{-1}\theta_D^* K_0) - \frac{1}{\gamma} \ln(\mathbb{E}[\exp(-\gamma \theta_0^* \cdot I_1(\theta_D^*, \theta_{BB}^*))]), \\
J(0, \theta_D^u) &= J(\theta_{BB}^*, \theta_D^*).
\end{aligned}$$

This pure dividends strategy thus replicates the investor's expected utility obtained with the optimal $(\theta_D^*, \theta_{BB}^*)$ mixed strategy.

As we did not pick θ_0^u nor θ_D^u optimally, $\sup_{\theta_D} J(0, \theta_D) \geq \sup_{\theta_D, \theta_{BB}} J(\theta_{BB}^*, \theta_D^*)$. As any pure share repurchases strategies is also a mixed strategy with $(\theta_D = 0)$, we get the desired $V_D \geq V_{BB}$.

SHARE REPURCHASE STRATEGIES DOMINATE PURE DIVIDENDS ONES

We show that $V_D \leq V_{BB}$ under Assumption 1.7.⁷ We first state our results and then provide their proofs.

Proposition 2.2. *For any $(\gamma, R_f, K_0, \mu, \sigma)$ satisfying Assumption 1.7, there exists a unique $\theta_{BB}^u \in \mathbb{R}$ such that:*

$$K_0\theta_D^* = \theta_{BB}^u \hat{P}_0(\theta_{BB}^u). \quad (1.8)$$

We also have

$$\hat{J}(0, \theta_D^*) = \hat{J}(\theta_{BB}^u, 0),$$

And in turn, by generality of θ_{BB}^u ,

$$V_{BB} \geq V_D.$$

Proposition 2.3. *The unique solution to (1.8) is*

$$\theta_{BB}^u = \theta_D^*.$$

⁷The general proof with Assumption 1.7 lifted is presented in the Appendix 1.

Remark 2.4. Given any $\theta_D \in (0, 1)$,

$$\exists! \theta_{BB} \in (0, 1) \text{ s.t. } K_0 \theta_D = \theta_{BB} \hat{P}_0(\theta_{BB}).$$

Proof of Proposition 2.2. Define $f : \mathbb{R}^* \mapsto \mathbb{R}$ as

$$\forall \theta_{BB} \in \mathbb{R}^*, f(\theta_{BB}) := \theta_{BB} \hat{P}_0(\theta_{BB}),$$

and define $f(0) = 0$ by continuity, since we have that $(\hat{\theta}_0, \hat{P}_0)$ is a consistent equilibrium.

With Assumption 1.7, f is a continuous function of θ_{BB} over $\mathbb{R}$. First, as $|\theta_{BB}| \rightarrow \infty$, we have:

$$\begin{aligned} \lim_{\theta_{BB} \rightarrow +\infty} f(\theta_{BB}) &= K_0 - \frac{\mu - 1 - R_f}{2\sigma^2\gamma} + \frac{\sqrt{(\mu - 1 - R_f)^2 + 4K_0\sigma^2\gamma(1 + R_f)}}{2\sigma^2\gamma} \\ &> K_0 - \frac{\mu - 1 - R_f}{2\sigma^2\gamma} + \frac{|\mu - 1 - R_f|}{2\sigma^2\gamma} \\ &\geq K_0 - \frac{\mu - 1 - R_f}{\sigma^2\gamma} \\ &= K_0 \theta_D^*, \end{aligned}$$

and

$$\begin{aligned} \lim_{\theta_{BB} \rightarrow -\infty} f(\theta_{BB}) &= K_0 - \frac{\mu - 1 - R_f}{2\sigma^2\gamma} - \frac{\sqrt{(\mu - 1 - R_f)^2 + 4K_0\sigma^2\gamma(1 + R_f)}}{2\sigma^2\gamma} \\ &< K_0 - \frac{\mu - 1 - R_f}{2\sigma^2\gamma} - \frac{|\mu - 1 - R_f|}{2\sigma^2\gamma} \\ &\leq K_0 - \frac{\mu - 1 - R_f}{\sigma^2\gamma} \\ &= K_0 \theta_D^*. \end{aligned}$$

We now investigate the monotonicity of $f(\theta_{BB})$. f is differentiable over $\mathbb{R}^*$, and direct differentiation gives

$$f'(\theta_{BB}) = \frac{1 + R_f - \Delta(\theta_{BB})^{-\frac{1}{2}}(\Delta(\theta_{BB}) - \frac{\theta_{BB}}{2} \Delta'(\theta_{BB}))}{2\sigma^2\gamma\theta_{BB}^2}.$$

We recall that

$$\begin{aligned} \Delta(\theta_{BB}) &:= (\mu\theta_{BB} + (1 + R_f)(1 - \theta_{BB}))^2 - 4\sigma^2\gamma(1 + R_f)K_0\theta_{BB}(1 - \theta_{BB}), \\ &= a\theta_{BB}^2 + b\theta_{BB} + c, \end{aligned}$$

where

$$\begin{aligned} a &:= (\mu - 1 - R_f)^2 + 4K_0\gamma\sigma^2(1 + R_f) > 0, \\ b &:= 2(1 + R_f)(\mu - 1 - R_f) - 4K_0\gamma\sigma^2(1 + R_f), \\ c &:= (1 + R_f)^2. \end{aligned}$$

Thus, $f'(\theta_{BB}) > 0$ is equivalent to

$$\sqrt{\Delta(\theta_{BB})(1+R_f)} - (\Delta(\theta_{BB}) - \frac{\theta_{BB}}{2}\Delta'(\theta_{BB})) > 0.$$

That is:

$$\begin{aligned} f'(\theta_{BB}) > 0 &\iff (1+R_f)\sqrt{\Delta(\theta_{BB})} > \Delta(\theta_{BB}) - \frac{\theta_{BB}}{2}\Delta'(\theta_{BB}), \\ &\iff (1+R_f)^2\Delta(\theta_{BB}) > (\Delta(\theta_{BB}) - \frac{\theta_{BB}}{2}\Delta'(\theta_{BB}))^2, \\ &\iff c(a\theta_{BB}^2 + b\theta_{BB} + c) > (\frac{b}{2}\theta_{BB} + c)^2, \\ &\iff (ac - \frac{b^2}{4})\theta_{BB}^2 > 0, \\ &\iff (\mu - K_0\gamma\sigma^2)\theta_{BB}^2 > 0. \end{aligned}$$

Which is equivalent to Assumption 1.7.

Therefore, $f'(\theta_{BB}) > 0$ for all $\theta_{BB} \in \mathbb{R}^*$. Hence f must be strictly increasing and continuous on $\mathbb{R}$.

Together with the facts that $\lim_{\theta_{BB} \rightarrow -\infty} f(\theta_{BB}) < K_0\theta_D^*$ and $\lim_{\theta_{BB} \rightarrow \infty} f(\theta_{BB}) > K_0\theta_D^*$, the equation $f(\theta_{BB}) = K_0\theta_D^*$ must yield a unique solution.

Let us now consider the resulting utilities $\hat{J}(0, \theta_D^*)$ and $\hat{J}(\theta_{BB}^u, 0)$. Under the optimal pure dividends strategy θ_D^* , the representative investor's expected utility is:

$$\begin{aligned} \hat{J}(0, \theta_D^*) &= (1 + R_f)(W_{-1} + \theta_{-1}\theta_D^*K_0 + \hat{P}_0(\theta_D^*) \cdot (\theta_{-1} - \hat{\theta}_0(\theta_D^*))) \\ &\quad + \mu(K_0 - K_0\theta_D^*)\hat{\theta}_0(\theta_D^*) - \frac{\sigma^2\gamma}{2}(K_0 - K_0\theta_D^*)^2\hat{\theta}_0(\theta_D^*)^2. \end{aligned}$$

At equilibrium, we have $\theta_{-1} = \hat{\theta}_0(\theta_D^*) = 1$, thus $\hat{J}(0, \theta_D^*)$ becomes:

$$\hat{J}(0, \theta_D^*) = (1 + R_f)(W_{-1} + \theta_D^*K_0) + \mu(K_0 - K_0\theta_D^*) - \frac{\sigma^2\gamma}{2}(K_0 - K_0\theta_D^*)^2.$$

In $\hat{J}(0, \theta_D^*)$, replacing $K_0\theta_D^*$ by the replicating $\theta_{BB}^u \hat{P}_0(\theta_{BB}^u)$, we get:

$$\begin{aligned} \hat{J}(0, \theta_D^*) &= (1 + R_f)(W_{-1} + \theta_{BB}^u \cdot \hat{P}_0(\theta_{BB}^u)) + \mu(K_0 - \theta_{BB}^u \cdot \hat{P}_0(\theta_{BB}^u)) \frac{1 - \theta_{BB}^u}{1 - \theta_{BB}^u} \\ &\quad - \frac{\sigma^2 \gamma}{2} (K_0 - \theta_{BB}^u \cdot \hat{P}_0(\theta_{BB}^u))^2 \left(\frac{1 - \theta_{BB}^u}{1 - \theta_{BB}^u} \right)^2 \end{aligned}$$

Using that $\hat{\theta}_0(\theta_{BB}^u) = 1 - \theta_{BB}^u$:

$$\begin{aligned} \hat{J}(0, \theta_D^*) &= (1 + R_f)(W_{-1} + (1 - \hat{\theta}_0(\theta_{BB}^u)) \cdot \hat{P}_0(\theta_{BB}^u)) + \mu(K_0 - \theta_{BB}^u \cdot \hat{P}_0(\theta_{BB}^u)) \frac{\hat{\theta}_0(\theta_{BB}^u)}{1 - \theta_{BB}^u} \\ &\quad - \frac{\sigma^2 \gamma}{2} (K_0 - \theta_{BB}^u \cdot \hat{P}_0(\theta_{BB}^u))^2 \left(\frac{\hat{\theta}_0(\theta_{BB}^u)}{1 - \theta_{BB}^u} \right)^2. \end{aligned}$$

$$\hat{J}(0, \theta_D^*) = \hat{J}(\theta_{BB}^u, 0).$$

As we did not choose θ_{BB}^u and θ_0^u optimally, $V_{BB} \geq V_D$.

□

Proof of Proposition 2.3. We recall the formulation for the equilibrium price in the pure share buybacks setting, $\hat{P}_0(\theta_{BB})$.

From Proposition 1.5 and given a repurchases ratio θ_{BB} , we have:

$$\Delta(\theta_{BB}) = [\mu\theta_{BB} + (1 + R_f)(1 - \theta_{BB})]^2 - 4\sigma^2\gamma(1 + R_f)K_0\theta_{BB}(1 - \theta_{BB}).$$

Given $\Delta(\theta_{BB}) \geq 0$, the equilibrium price then is:

$$\hat{P}_0(\theta_{BB}) = \frac{2\sigma^2\gamma\theta_{BB}K_0 - \mu\theta_{BB} - (1 + R_f)(1 - \theta_{BB}) + \sqrt{\Delta(\theta_{BB})}}{2\sigma^2\gamma\theta_{BB}^2}.$$

Let us now determine $\Delta(\theta_{BB}^u)$. Thanks to Assumption 1.7, $\Delta(\theta_{BB}^u) \geq 0$.

$$\begin{aligned}
\Delta(\theta_{BB}^u) &= \left[\mu \left(1 - \frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0} \right) + (1 + R_f) \frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0} \right]^2 \\
&\quad - 4\sigma^2 \gamma (1 + R_f) K_0 \left(1 - \frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0} \right) \frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0}, \\
&= \mu^2 \left(1 - \frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0} \right)^2 + 2\mu(1 + R_f) \left(\frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0} - \left(\frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0} \right)^2 \right) \\
&\quad + (1 + R_f)^2 \left(\frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0} \right)^2 - 4(1 + R_f)(\mu - (1 + R_f)) + 4(1 + R_f) \frac{(\mu - (1 + R_f))^2}{\sigma^2 \gamma K_0}, \\
&= \mu^2 - 2\mu^2 \frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0} + (\mu^2 + (1 + R_f)^2 - 2\mu(1 + R_f)) \left(\frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0} \right)^2 \\
&\quad + 2\mu(1 + R_f) \left(\frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0} \right) - 4(1 + R_f)(\mu - (1 + R_f)) + 4(1 + R_f) \frac{(\mu - (1 + R_f))^2}{\sigma^2 \gamma K_0}, \\
&= \frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0} (-2\mu^2 + 2\mu(1 + R_f) + 4(1 + R_f)(\mu - (1 + R_f))) + \\
&\quad \frac{(\mu - (1 + R_f))^4}{(\sigma^2 \gamma K_0)^2} + \mu^2 - 4(1 + R_f)(\mu - (1 + R_f)), \\
&= \frac{(\mu - (1 + R_f))^4}{(\sigma^2 \gamma K_0)^2} + 2 \frac{\mu - (1 + R_f)}{\sigma^2 \gamma K_0} ((1 + R_f)(\mu - (1 + R_f)) - (\mu - (1 + R_f))^2) \\
&\quad + (\mu - (1 + R_f))^2 - 2(1 + R_f)(\mu - (1 + R_f)) + (1 + R_f)^2, \\
&= \left(\frac{(\mu - (1 + R_f))^2}{\sigma^2 \gamma K_0} \right)^2 - 2 \frac{(\mu - (1 + R_f))^2}{\sigma^2 \gamma K_0} (\mu - 2(1 + R_f)) + (\mu - 2(1 + R_f))^2, \\
&= \left(\frac{(\mu - (1 + R_f))^2}{\gamma K_0 \sigma^2} - (\mu - 2(1 + R_f)) \right)^2.
\end{aligned}$$

Since $\Delta(\theta_{BB}^u)$ is expressed as a squared analytical expression, we can write:

$$\sqrt{\Delta(\theta_{BB}^u)} = \left| \frac{(\mu - (1 + R_f))^2}{\gamma K_0 \sigma^2} - (\mu - 2(1 + R_f)) \right|.$$

Using Assumption 1.7, we have:

$$\frac{(\mu - (1 + R_f))^2}{\gamma K_0 \sigma^2} \geq \mu - 2(1 + R_f) + \frac{(1 + R_f)^2}{\mu} \geq (\mu - 2(1 + R_f)).$$

Thus,

$$\sqrt{\Delta(\theta_{BB}^u)} = \frac{(\mu - (1 + R_f))^2}{\gamma K_0 \sigma^2} - (\mu - 2(1 + R_f)).$$

Plugging this expression back in the expression for the company's price, $\hat{P}_0(\theta_{BB}^u)$:

$$\begin{aligned}
\hat{P}_0(\theta_{BB}^u) &= \frac{2\sigma^2\gamma\theta_{BB}^u K_0 - \mu\theta_{BB}^u - (1+R_f)(1-\theta_{BB}^u) + \sqrt{\Delta(\theta_{BB}^u)}}{2\sigma^2\gamma\theta_{BB}^u{}^2}, \\
&= \frac{1}{2\sigma^2\gamma\left(1 - \frac{\mu-(1+R_f)}{\sigma^2\gamma K_0}\right)^2} \left((2\sigma^2\gamma K_0 - \mu) \left(1 - \frac{\mu-(1+R_f)}{\sigma^2\gamma K_0}\right) \right. \\
&\quad \left. - (1+R_f) \frac{\mu-(1+R_f)}{\sigma^2\gamma K_0} + \frac{(\mu-(1+R_f))^2}{\sigma^2\gamma K_0} - (\mu - 2(1+R_f)) \right), \\
&= \frac{1}{2\sigma^2\gamma\left(1 - \frac{\mu-(1+R_f)}{\sigma^2\gamma K_0}\right)^2} \left(2\sigma^2\gamma K_0 - \mu - 2(\mu - (1+R_f)) - (\mu - 2(1+R_f)) \right. \\
&\quad \left. + (\mu - (1+R_f)) \frac{\mu-(1+R_f)}{\sigma^2\gamma K_0} + \frac{(\mu-(1+R_f))^2}{\sigma^2\gamma K_0} \right), \\
&= \frac{2}{2\sigma^2\gamma\left(1 - \frac{\mu-(1+R_f)}{\sigma^2\gamma K_0}\right)^2} \left(\sigma^2\gamma K_0 - 2(\mu - (1+R_f)) + \frac{(\mu-(1+R_f))^2}{\sigma^2\gamma K_0} \right), \\
&= \frac{K_0}{\left(1 - \frac{\mu-(1+R_f)}{\sigma^2\gamma K_0}\right)^2} \left(1 - 2 \frac{\mu-(1+R_f)}{\sigma^2\gamma K_0} + \left(\frac{\mu-(1+R_f)}{\sigma^2\gamma K_0} \right)^2 \right), \\
&= K_0 \cdot \frac{\left(1 - \frac{\mu-(1+R_f)}{\sigma^2\gamma K_0}\right)^2}{\left(1 - \frac{\mu-(1+R_f)}{\sigma^2\gamma K_0}\right)^2} = K_0.
\end{aligned}$$

With the chosen θ_{BB}^u , $\hat{P}_0(\theta_{BB}^u)$ precisely equals K_0 . Thus, with $\theta_{BB}^u = \theta_D^*$, $\theta_{BB}^u \hat{P}_0(\theta_{BB}^u) = \theta_D^* K_0$. □

Proof of Remark 2.4. We know that given θ_{BB} , $(\hat{\theta}_0, \hat{P}_0)$ is a consistent equilibrium strategy, that is:

$$\begin{aligned}
\lim_{\theta_{BB} \rightarrow 0} \theta_{BB} \cdot \hat{P}_0(\theta_{BB}) &= 0, \\
\lim_{\theta_{BB} \rightarrow 1} \theta_{BB} \cdot \hat{P}_0(\theta_{BB}) &= K_0.
\end{aligned}$$

Since $\hat{P}_0(\theta_{BB})$ is a continuous function of θ_{BB} over $\mathbb{R}$, we have that

$$\forall \theta_D \in (0, 1), \exists \theta_{BB} \text{ s.t. } \theta_{BB} \hat{P}_0(\theta_{BB}) = K_0 \theta_D.$$

As in the proof of Proposition 2.2,

$$\forall \theta_{BB} \in \mathbb{R}^*, \partial_{\theta_{BB}} \theta_{BB} \hat{P}_0(\theta_{BB}) > 0.$$

Thus,

$$\forall \theta_D \in (0, 1), \exists! \theta_{BB} \text{ s.t. } \theta_{BB} \hat{P}_0(\theta_{BB}) = K_0 \theta_D.$$

□

For varying risk-aversion factor γ and expected growth rate μ , Figures 1.2 and 1.3 show the optimal behaviour of equilibrium of the pure dividends and repurchases models. Results are shown for three distinct values of K_0 (*see fig. 1.2*) and R_f (*see fig. 1.3*). In particular, we present the optimal ratio spent on dividends θ_D^* and ratio of shares bought back θ_{BB}^* (*top-right*). At these optimal levels of spending, we show the investor's expected utility (*top-left*), the company's price (*bottom-right*) and the amount spent from the firm's assets on redistribution (*bottom-left*).

In the numerical studies, we apply a scaling of the risk aversion factor γ by $(1/K_0)$ so that the risk incurred by owning the whole firm does not vary according to the firm size. That is, this scaling ensures that larger firms are not penalized for their size.

At the optimal levels of cash spent / shares bought back, the investor's utility, cash spent and redistribution ratios are identical for the dividend and buyback models. We get the expected behaviours: the higher the risk aversion, the heftier the risk penalty for investing in the company gets. The company thus redistributes more to alleviate the risk carried by its investors. The higher the risk-free bank account growth rate R_f , the least growth in excess of the risk-free rate the investor can expect from the risky technology. Thus, as R_f grows, the firm redistributes more of its assets back to investors. Inversely, the higher the expected growth rate of the risk technology μ , the least the firm spends on redistribution. Since the risk-aversion scales as the inverse of the firm's total assets, the share of assets given back to investors is independent of K_0 .

However, in the buybacks setup, price is not influenced by the level of risk aversion, whereas the dividends' setup price is (*see fig. 1.3, bottom-right*).

Indeed, the artificial increased demand in shares incurred by the share repurchases keeps the company's price high. In practice, $\hat{P}_0$ is determined through market clearing to ensure the aggregate demand for shares matches the unit supply: $\theta_{BB} + \theta_0 = 1$. Equivalently, repurchases reduce supply, yielding a higher price at equilibrium $\hat{P}_0$. The optimal values of θ_D and θ_{BB} are set solely considering the maximisation of the investor's expected utility. As such, dividends and share repurchases are economically equivalent. Yet, implementing share repurchases allows for a higher company share price, and management of the total amount of outstanding shares.

Many other actors, both within and exterior to the firm, are exposed to the company's stock price directly, or have an exposure tied to the share price. Similarly, reduction of the share supply has significant consequences. A multitude of factors motivate the

management stock price and number of shares outstanding.

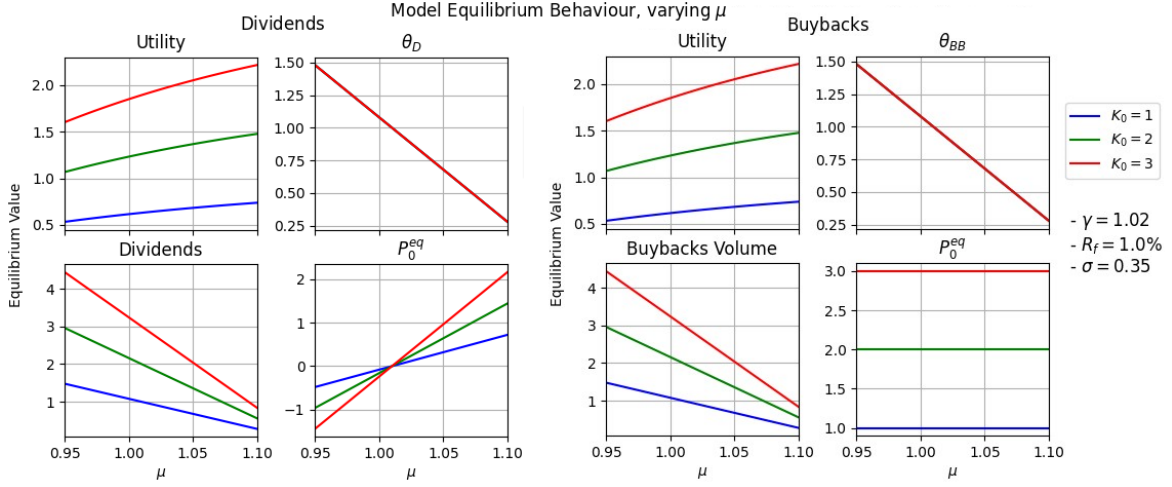


Figure 1.2: For varying expected growth rate μ , are presented the equilibria behaviour of the dividends (Left) and share repurchases (Right) models.

As the K_0 increases (Low: Blue, Average: Green and High: Red), the investor utility, redistribution and price increase proportionally. However the ratio of assets redistributed remains unchanged.

As the firm becomes more attractive, the investor's expected utility increases, the amount redistributed is reduced, and the firm's price in the dividends setting rises, though even for high μ , $\hat{P}_0(\theta_D^*)$ is smaller than $\hat{P}_0(\theta_{BB}^*)$. When $(\mu \leq (1 + R_f))$, $\hat{P}_0(\theta_D^*)$ is negative.

We focus on two such factors in particular.

Firstly, analysts' firm value models are generally ultimately compared to the firm's per share price. Such indicators as Earnings Per Share (EPS) are central in determining a company's outlook (Consler et al., 2011; Machuga et al., 2002). Share repurchases are in practice widely used for the purpose of meeting periodic EPS targets fixed by analysts. Precedent research empirically shows this relationship (Gao and Kronlund, 2020; Hribar et al., 2006). Our model confirms this view. Indeed, equal amounts are redistributed to investors in the dividends and buybacks setups. Since the company's prospects are proportional to its remaining size, the two redistribution methods yield the same expected earnings at time $t = 1$ for the firm. As repurchases reduce the supply of shares, buybacks in turn relatively increase the EPS ratio, without modifying the outlook for the firms earnings at time $t = 1$. In practice, meeting EPS goals is all-important for executives, whose tenure might be put in jeopardy should their company fail to do so.

In parallel, stock options are priced relative to the company's stock price $\hat{P}_0$. Indeed, the payoff of stock options at expiry is immediately determined by the share price at that time. Before expiration, the moneyness of an option is generally its main pricing component. Hence, the sizable derivatives market is largely exposed to movements in the company's stock price. Since the two redistribution methods have a distinct impact on $\hat{P}_0$, the choice of distributing a dividend or repurchasing outstanding shares

has momentous repercussions on the derivatives market. Furthermore, the company's shareholders and more importantly its executives often own stock options. The influence redistribution methodologies have on the stock options' value could be taken into account by the company decision makers. The company's decision makers' exposure to stock options could therefore influence the mode, and crucially the volume of redistribution to execute. Section 4.2 explores the impact of the different redistribution methods on executive stock options.

Other important factors also benefit from the management of the share supply and company price. Non-exhaustively, our results confirm that repurchases entail better performance regarding the following:

- Inclusion in the most popular Exchange Traded Funds is conditional to Market Capitalisation, which in our model share repurchases allow.
- For companies paying regular coupon dividends, retiring a portion of their shares instead of paying an exceptional dividend shrinks the future amount spent coupon payments.
- As mentioned in the introduction, repurchases can be executed in various forms. Both open market and tender offers allow for the investors whose perspective for the company is the bleakest to sell their shares at a higher price. Tender offers specifically, involving block sales at an inflated price permit large shareholders to exit their positions without impacting the market. Section 3.1 presents an extension of this section's model with multiple heterogeneous investors.

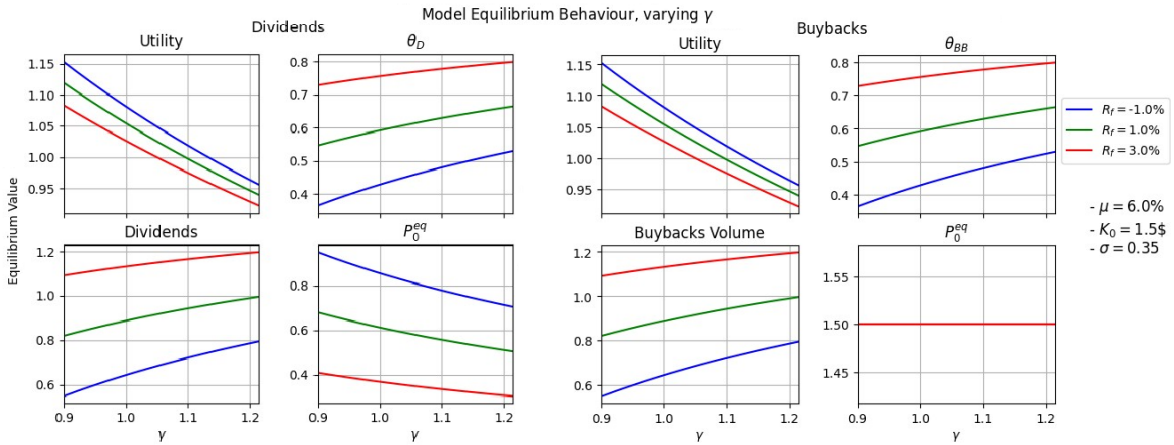


Figure 1.3: For varying investor risk-aversion γ , are presented the equilibria behaviour of the dividends (Left) and share repurchases (Right) models.

As the bank account growth rate rises (Negative: Blue, Low: Green and High: Red), we can see that investor utility and redistribution grow proportionally. The company's price in turn decreases in the dividends setting. It is however independent of R_f in the buybacks setting: buybacks allow to shield the share price from inflationary pressures.

As observed in the previous graph, everything but share price are equivalent in the two models.

3 MODEL EXTENSIONS UNDER WHICH DIVIDEND-BUYBACK EQUIVALENCE REMAINS VALID

In this section, we consider a few variations of the baseline model where the equivalence of dividend and buyback as shown in Theorem 2.1 still remains valid.

3.1 INVESTORS WITH HETEROGENEOUS BELIEFS OVER FIRM'S PROFITABILITY

In the baseline model, investors are homogeneous such that the population can be summarized by a single representative investor. Now suppose the investors have different beliefs about the profitability of the firm. To simplify the setup, we consider two groups of infinitesimal investors, an optimistic group and a pessimistic one. In our friction-less setup, with exponential utility, each group is summarized by a single representative investor, an optimist (*resp.* *pessimist*) labeled by $i = 1$ ($i = 2$) believes the firm's capital will grow at a stochastic rate of $\Pi_1 \sim N(\mu_1, \sigma^2)$ ($\Pi_2 \sim N(\mu_2, \sigma^2)$) with $\mu_2 \leq \mu_1$.

The rest of the setup is largely the same as in Section 1. For $i \in \{1, 2\}$, we assume the investor i is initially endowed with cash of W_{-1}^i and θ_{-1}^i units of shares at time $t = -1$. The number of shares held by investor i immediately after the firm's distribution is denoted by θ_0^i such that the cash amount available to investor i is

$$W_0^i := W_{-1}^i + \theta_D K_0 + (\theta_{-1}^i - \theta_0^i) P_0.$$

The portfolio dynamics of each investor is then given by

$$W_1^i := W_0(1 + R_f) + \theta_0^i D_1^i,$$

where D_1^i is the liquidation dividend per share perceived by investor i with expression of

$$D_1^i := \frac{(K_0 - \theta_D K_0 - \theta_{BB} P_0) \Pi_i}{1 - \theta_{BB}}.$$

The optimist and pessimist have the same utility function $U(x) = (1 - \exp(-\gamma x))/\gamma$ as before. The certainty equivalent of investor i under their given trading strategy θ_0^i , the market price of the firm P_0 and the distribution strategy (θ_D, θ_{BB}) is defined as

$$\begin{aligned} J_i(\theta_0^i; P_0, \theta_D, \theta_{BB}) &:= U^{-1}(\mathbb{E}[U(W_1^i)]) \\ &= (W_{-1}^i + \theta_D K_0 + (\theta_{-1}^i - \theta_0^i) P_0) (1 + R_f) + \\ &\quad \frac{\theta_0^i (K_0 - \theta_D K_0 - \theta_{BB} P_0) \mu_i}{1 - \theta_{BB}} - \frac{\sigma^2 \gamma (\theta_0^i)^2 (K_0 - \theta_D K_0 - \theta_{BB} P_0)^2}{2 (1 - \theta_{BB})^2}. \end{aligned}$$

With the above notations, we can define a competitive equilibrium featuring heterogeneous investors similar to what we did in Definition 1.2.

Definition 3.1. *Given the firm's distribution strategy (θ_D, θ_{BB}) , in the case with heterogeneous investors a competitive equilibrium is a tuple of the investors' trading strategies and market price of the firm $(\hat{\theta}_0^1, \hat{\theta}_0^2, \hat{P}_0)$ such that:*

1. *Both representative investors maximize the certainty equivalent of their terminal portfolio value, i.e.*

$$J_i(\hat{\theta}_0^i; \hat{P}_0, \theta_D, \theta_{BB}) = \sup_{\theta_0^i} J_i(\theta_0^i; \hat{P}_0, \theta_D, \theta_{BB}), \quad i \in \{1, 2\}.$$

2. *The market clears, i.e. the total demand for the shares is equal to the number of outstanding shares post-buyback, or equivalently*

$$1 = \theta_{-1}^1 + \theta_{-1}^2 = \hat{\theta}_0^1 + \hat{\theta}_0^2 + \theta_{BB}.$$

For $i \in \{1, 2\}$, we refer to $\hat{J}_i(\theta_D, \theta_{BB}) := J_i(\hat{\theta}_0^i; \hat{P}_0, \theta_D, \theta_{BB})$ as an equilibrium value function of investor i under the distribution strategy (θ_D, θ_{BB}) . Write $\hat{J}_{i,D}(\theta_D) := \hat{J}_i(\theta_D, \theta_{BB} = 0)$ and $\hat{J}_{i,B}(\theta_{BB}) := \hat{J}_i(\theta_D = 0, \theta_{BB})$ as investor i 's equilibrium value functions under a (pure) dividend strategy and a (pure) buyback strategy respectively.

Similar to the baseline model, the competitive equilibrium can be characterized for both the dividend and buyback strategy. The proposition below is analogous to Proposition 1.3 and Corollary 1.8.

Proposition 3.2. *Write $\bar{\mu} := (\mu_1 + \mu_2)/2$ as the average estimate of the firm's growth by the two investors.*

1. *Under a dividend strategy $(\theta_D, \theta_{BB} = 0)$, the unique competitive equilibrium is given by*

$$\hat{P}_0 = \frac{K_0(1 - \theta_D)}{1 + R_f} \left(\bar{\mu} - \frac{\sigma^2 \gamma K_0(1 - \theta_D)}{2} \right),$$

and

$$\hat{\theta}_0^i = \frac{1}{2} + \frac{\mu_i - \mu_{-i}}{2\sigma^2 \gamma K_0(1 - \theta_D)}, \quad i \in \{1, 2\}.$$

2. *Under a buyback strategy $(\theta_D = 0, \theta_{BB})$, the unique consistent competitive equi-*

librium is given by

$$\hat{P}_0 = \frac{\sigma^2 \gamma \theta_{BB} K_0 - \bar{\mu} \theta_{BB} - (1 + R_f)(1 - \theta_{BB}) + \sqrt{\Delta(\theta_{BB})/4}}{\sigma^2 \gamma \theta_{BB}^2}$$

and

$$\hat{\theta}_0^i = \frac{1 - \theta_{BB}}{\sigma^2 \gamma (K_0 - \hat{P}_0 \theta_{BB})} \left[\mu_i - \frac{\hat{P}_0 (1 + R_f)(1 - \theta_{BB})}{K_0 - \hat{P}_0 \theta_{BB}} \right], \quad i \in \{1, 2\},$$

where

$$\Delta(\theta_{BB}) := 4 \left([\bar{\mu} \theta_{BB} + (1 + R_f)(1 - \theta_{BB})]^2 - 2\sigma^2 \gamma (1 + R_f) K_0 \theta_{BB} (1 - \theta_{BB}) \right).$$

Proof. The derivation of $\hat{P}_0$ and $\hat{\theta}_0^i$ follows the outline from the representative investor setting from Section 1.

Pure Dividends Setting:

The optimal ownership θ_0^i for each investor maximizes their expected utility:

$$\frac{\partial \hat{J}_{i,D}(\theta_D)}{\partial \theta_0^i} = 0.$$

Since,

$$\frac{\partial \hat{J}_{i,D}(\theta_D)}{\partial \theta_0^i} = -P_0(1 + R_f) + \mu_i K_0(1 - \theta_D) - \theta_0^i \gamma \sigma^2 K_0^2 (1 - \theta_D)^2,$$

we get

$$\hat{\theta}_0^i = \frac{1}{\sigma^2 \gamma (1 - \theta_D) K_0} \left(\mu_i - \frac{(1 + R_f) P_0}{(1 - \theta_D) K_0} \right).$$

Using the outlined market clearing condition:

$$\begin{aligned} 1 &= \hat{\theta}_0^1 + \hat{\theta}_0^2, \\ &= \frac{1}{\sigma^2 \gamma (1 - \theta_D) K_0} \left(\mu_1 + \mu_2 - \frac{2(1 + R_f) \hat{P}_0}{(1 - \theta_D) K_0} \right), \\ &= -\frac{2(1 + R_f) \hat{P}_0}{\sigma^2 \gamma (1 - \theta_D)^2 K_0^2} + \frac{\mu_1 + \mu_2}{\sigma^2 \gamma (1 - \theta_D) K_0}. \end{aligned}$$

Thus,

$$\hat{P}_0 = \frac{\sigma^2 \gamma (1 - \theta_D)^2 K_0^2}{2(1 + R_f)} \left(\frac{\mu_1 + \mu_2}{\sigma^2 \gamma (1 - \theta_D) K_0} - 1 \right).$$

And thus,

$$\hat{P}_0 = \frac{(1 - \theta_D)K_0}{1 + R_f} \left(\bar{\mu} - \frac{1}{2}\sigma^2\gamma K_0(1 - \theta_D) \right).$$

Plugging back the market clearing price in the expression of θ_0^i , we obtain:

$$\begin{aligned} \hat{\theta}_0^i &= \frac{1}{\sigma^2\gamma(1 - \theta_D)K_0} \left(\mu_i - \left(\bar{\mu} - \frac{1}{2}\sigma^2\gamma K_0(1 - \theta_D) \right) \right), \\ \hat{\theta}_0^i &= \frac{1}{2} + \frac{\mu_i - \mu_{-i}}{2\sigma^2\gamma(1 - \theta_D)K_0}. \end{aligned}$$

Where $-i$ represents the investor opposite i .

Pure Share Repurchases Setting:

Similarly,

$$\frac{\partial \hat{J}_{i,BB}(\theta_{BB})}{\partial \theta_0^i} = -P_0(1 + R_f) + \mu_i \frac{K_0 - \theta_{BB}P_0}{1 - \theta_{BB}} - \theta_0^i \sigma^2\gamma \frac{(K_0 - \theta_{BB}P_0)^2}{(1 - \theta_{BB})^2}.$$

Thus

$$\begin{aligned} 0 &= -P_0(1 + R_f) + \mu_i \frac{K_0 - \theta_{BB}P_0}{1 - \theta_{BB}} - \theta_0^i \sigma^2\gamma \frac{(K_0 - \theta_{BB}P_0)^2}{(1 - \theta_{BB})^2}, \\ \hat{\theta}_0^i &= \frac{(1 - \theta_{BB})^2}{\sigma^2\gamma(K_0 - \theta_{BB}P_0)^2} \left(\mu_i \frac{K_0 - \theta_{BB}P_0}{1 - \theta_{BB}} - P_0(1 + R_f) \right), \\ \hat{\theta}_0^i &= \frac{1 - \theta_{BB}}{\sigma^2\gamma(K_0 - \theta_{BB}P_0)} \left(\mu_i - \frac{P_0(1 - \theta_{BB})(1 + R_f)}{K_0 - \theta_{BB}P_0} \right), \end{aligned}$$

From this, using $1 = \hat{\theta}_0^1 + \hat{\theta}_0^2 + \theta_{BB}$,

$$\begin{aligned} 1 &= \frac{1}{\sigma^2\gamma(K_0 - \theta_{BB}\hat{P}_0)} \left(\mu_1 + \mu_2 - \frac{2\hat{P}_0(1 - \theta_{BB})(1 + R_f)}{K_0 - \theta_{BB}\hat{P}_0} \right), \\ 0 &= 2\bar{\mu}(K_0 - \theta_{BB}\hat{P}_0) - \sigma^2\gamma(K_0 - \theta_{BB}\hat{P}_0)^2 - 2\hat{P}_0(1 - \theta_{BB})(1 + R_f), \\ 0 &= \hat{P}_0^2(-\sigma^2\gamma\theta_{BB}^2) \\ &\quad + \hat{P}_0 2(\sigma^2\gamma K_0\theta_{BB} - \bar{\mu}\theta_{BB} - (1 - \theta_{BB})(1 + R_f)) \\ &\quad + 2\bar{\mu}K_0 - \sigma^2\gamma K_0^2. \end{aligned}$$

So, as in the representative investor setting, $\hat{P}_0$ is the solution to a quadratic polynomial

which has:

$$\Delta = 4 \cdot \left[(\bar{\mu} \theta_{BB} + (1 + R_f)(1 - \theta_{BB}))^2 - 2K_0 \gamma \sigma^2 \theta_{BB}(1 - \theta_{BB})(1 + R_f) \right].$$

And given $\Delta \geq 0$,

$$\hat{P}_0(\theta_{BB}) \in \left\{ \frac{\sigma^2 \gamma \theta_{BB} K_0 - \bar{\mu} \theta_{BB} - (1 + R_f)(1 - \theta_{BB}) \pm \sqrt{\Delta(\theta_{BB})/4}}{\sigma^2 \gamma \theta_{BB}^2} \right\}.$$

$\hat{P}_0^-$ yields a non-consistent equilibrium, using l'Hôpital's rule in the $\hat{P}_0^+$ case:

$$\begin{aligned} \lim_{\theta_{BB} \rightarrow 0} \theta_{BB} \hat{P}_0^-(\theta_{BB}) &= K_0 - \frac{\bar{\mu}}{\sigma^2 \gamma} - 2 \lim_{\theta_{BB} \rightarrow 0} \frac{(1 + R_f)}{\sigma^2 \gamma \theta_{BB}} \neq 0, \\ \lim_{\theta_{BB} \rightarrow 0} \theta_{BB} \hat{P}_0^+(\theta_{BB}) &= \frac{\sigma^2 \gamma K_0 - \bar{\mu} + (1 + R_f) + \frac{2(\bar{\mu} - (1 + R_f))(1 + R_f) - 2K_0 \gamma \sigma^2 (1 + R_f)}{2(1 + R_f)}}{\sigma^2 \gamma}, \\ \lim_{\theta_{BB} \rightarrow 0} \theta_{BB} \hat{P}_0^+(\theta_{BB}) &= 0. \end{aligned}$$

$\hat{P}_0^-$ fails to be consistent for $\theta_{BB} = 0$. $\hat{P}_0^+$ is however consistent:

$$\begin{aligned} \lim_{\theta_{BB} \rightarrow 1} \theta_{BB} \hat{P}_0^+(\theta_{BB}) &= \frac{\sigma^2 \gamma K_0 - \bar{\mu} + \bar{\mu}}{\sigma^2 \gamma}, \\ \lim_{\theta_{BB} \rightarrow 1} \theta_{BB} \hat{P}_0^+(\theta_{BB}) &= K_0. \end{aligned}$$

Thus, given $\Delta \geq 0$, there is a unique consistent equilibrium in the repurchases setting.

Note that Assumption 1.7, with μ replaced by $\bar{\mu}$, is enough to guarantee that $\forall \theta_{BB} \in \mathbb{R}$, $\Delta(\theta_{BB}) \geq 0$.

Actually, we only need a less restrictive condition to ensure that Δ is always non-negative: Assumption 3.3 is sufficient. $\square$

Standing Assumption 3.3. *The model parameters are assumed to satisfy:*

$$\bar{\mu} \geq \frac{K_0 \gamma \sigma^2}{2}.$$

With two representative investors, we can consider the aggregate welfare as the sum of the certain equivalent across the optimist and the pessimist defined via

$$\bar{J}_k(\theta_k) := \hat{J}_{1,k}(\theta_k) + \hat{J}_{2,k}(\theta_k) \quad (1.9)$$

under distribution method $k \in \{D, BB\}$. The following result shows that the equivalence of dividend and buyback still holds under heterogeneous beliefs of the investors.

Theorem 3.4. *Under the case with heterogeneous beliefs, dividend and buyback are economically equivalent in terms of maximizing the total certain equivalent of the two representative investors, i.e.*

$$\sup_{\theta_D} \bar{J}_D(\theta_D) = \sup_{\theta_{BB}} \bar{J}_{BB}(\theta_{BB}).$$

Furthermore, the value of equity distributed to the investors are identical across dividend and buyback strategy, i.e.

$$K_0 \theta_D^* = \theta_{BB}^* \hat{P}_0(\theta_{BB}^*)$$

where $\theta_k^* = \arg \max_{\theta_k} \bar{J}_k(\theta_k)$ for $k \in \{D, BB\}$.

To prove Theorem 3.4, we rely on the following two lemmas, which we prove first.

Lemma 3.5 (Optimal Dividends Equilibrium, Heterogeneous Investors). *The unique optimal pure dividends strategy θ_D^* in the heterogeneous investors setting is*

$$\theta_D^* = 1 - 2 \frac{\bar{\mu} - (1 + R_f)}{\sigma^2 \gamma K_0}.$$

At equilibrium it yields:

$$\hat{P}_0 = 2 \frac{\bar{\mu} - (1 + R_f)}{\gamma \sigma^2}, \text{ and } \hat{\theta}_0^i = \frac{1}{2} + \frac{\mu_i - \mu_{-i}}{4(\bar{\mu} - (1 + R_f))}.$$

The optimal aggregate utility then becomes

$$\bar{J}_D(\theta_D^*) = (W_{-1}^1 + W_{-1}^2 + K_0)(1 + R_f) + \frac{(\bar{\mu} - (1 + R_f))^2}{\sigma^2 \gamma} + \frac{(\mu_1 - \mu_2)^2}{4\sigma^2 \gamma}.$$

Proof of Lemma 3.5.

$$\begin{aligned}
\bar{J}_D(\theta_D) &= (W_{-1}^1 + W_{-1}^2 + \theta_D K_0) (1 + R_f) + K_0(1 - \theta_D) (\mu_1 \hat{\theta}_0^1 + \mu_2 \hat{\theta}_0^2) \\
&\quad - \frac{\sigma^2 \gamma}{2} K_0^2 (1 - \theta_D)^2 ((\hat{\theta}_0^1)^2 + (\hat{\theta}_0^2)^2), \\
&= (W_{-1}^1 + W_{-1}^2 + \theta_D K_0) (1 + R_f) \\
&\quad + K_0(1 - \theta_D) \left(\mu_1 \left(\frac{1}{2} + \frac{\mu_1 - \mu_2}{2\sigma^2 \gamma K_0(1 - \theta_D)} \right) + \mu_2 \left(\frac{1}{2} + \frac{\mu_2 - \mu_1}{2\sigma^2 \gamma K_0(1 - \theta_D)} \right) \right) \\
&\quad - \frac{\sigma^2 \gamma}{2} K_0^2 (1 - \theta_D)^2 \left(\left(\frac{1}{2} + \frac{\mu_1 - \mu_2}{2\sigma^2 \gamma K_0(1 - \theta_D)} \right)^2 + \left(\frac{1}{2} + \frac{\mu_2 - \mu_1}{2\sigma^2 \gamma K_0(1 - \theta_D)} \right)^2 \right), \\
&= (W_{-1}^1 + W_{-1}^2 + \theta_D K_0) (1 + R_f) + K_0(1 - \theta_D) \bar{\mu} + \frac{(\mu_1 - \mu_2)^2}{2\sigma^2 \gamma} \\
&\quad - \frac{\sigma^2 \gamma}{4} K_0^2 (1 - \theta_D)^2 - \frac{(\mu_1 - \mu_2)^2}{4\sigma^2 \gamma}, \\
&= (W_{-1}^1 + W_{-1}^2 + \theta_D K_0) (1 + R_f) + K_0(1 - \theta_D) \bar{\mu} + \frac{(\mu_1 - \mu_2)^2}{4\sigma^2 \gamma} \\
&\quad - \frac{\sigma^2 \gamma}{4} K_0^2 (1 - \theta_D)^2.
\end{aligned}$$

As $\bar{J}_D(\theta_D)$ is a concave quadratic function, the optimal dividends ratio θ_D^* is characterized by,

$$0 = \frac{\partial \bar{J}_D(\theta_D^*)}{\partial \theta_D} = K_0(1 + R_f) - K_0 \bar{\mu} + \frac{\sigma^2 \gamma}{2} K_0^2 (1 - \theta_D^*).$$

Which gives

$$\begin{aligned}
\theta_D^* &= \frac{2}{\sigma^2 \gamma K_0} \left((1 + R_f) - \bar{\mu} + \frac{\sigma^2 \gamma}{2} K_0 \right), \\
&= 1 - 2 \frac{\bar{\mu} - (1 + R_f)}{\sigma^2 \gamma K_0}.
\end{aligned}$$

Plugging θ_D^* 's expression back in $\hat{\theta}_0^i$, $\hat{P}_0(\theta_D^*)$ and $\bar{J}_D(\theta_D^*)$:

$$\hat{\theta}_0^i = \frac{1}{2} + \frac{\mu_i - \mu_{-i}}{4(\bar{\mu} - (1 + R_f))}.$$

And similarly,

$$\begin{aligned}
\hat{P}_0 &= 2 \frac{\bar{\mu} - (1 + R_f)}{\sigma^2 \gamma (1 + R_f)} \left(\bar{\mu} - (\bar{\mu} - (1 + R_f)) \right), \\
&= 2 \frac{\bar{\mu} - (1 + R_f)}{\sigma^2 \gamma}.
\end{aligned}$$

And finally,

$$\begin{aligned}
\bar{J}_D(\theta_D^*) &= \left(W_{-1}^1 + W_{-1}^2 + \left(1 - 2 \frac{\bar{\mu} - (1 + R_f)}{\sigma^2 \gamma K_0} \right) K_0 \right) (1 + R_f) + 2 \bar{\mu} \frac{\bar{\mu} - (1 + R_f)}{\sigma^2 \gamma} \\
&\quad + \frac{(\mu_1 - \mu_2)^2}{4 \sigma^2 \gamma} - \frac{(\bar{\mu} - (1 + R_f))^2}{\sigma^2 \gamma}, \\
&= (W_{-1}^1 + W_{-1}^2 + K_0) (1 + R_f) + \frac{(\bar{\mu} - (1 + R_f))^2}{\sigma^2 \gamma} + \frac{(\mu_1 - \mu_2)^2}{4 \sigma^2 \gamma}.
\end{aligned}$$

Analogously, the individual utility is:

$$\begin{aligned}
J_{D,i}(\theta_D^*) &= \left(W_{-1}^i + \left(1 - 2 \frac{\bar{\mu} - (1 + R_f)}{\sigma^2 \gamma K_0} \right) K_0 \left(\frac{1}{2} + \frac{\mu_i - \mu_{-i}}{4(\bar{\mu} - (1 + R_f))} \right) + \frac{\mu_i - \mu_{-i}}{2 \gamma \sigma^2} \right) (1 + R_f) \\
&\quad + \mu_i \left(1 + \frac{\mu_i - \mu_{-i}}{2(\bar{\mu} - (1 + R_f))} \right) \frac{\bar{\mu} - (1 + R_f)}{\sigma^2 \gamma} - \left(1 + \frac{\mu_i - \mu_{-i}}{2(\bar{\mu} - (1 + R_f))} \right)^2 \frac{(\bar{\mu} - (1 + R_f))^2}{2 \sigma^2 \gamma}, \\
&= \left(W_{-1}^i + \frac{K_0}{2} + \frac{K_0(\mu_i - \mu_{-i})}{4(\bar{\mu} - (1 + R_f))} - \frac{\bar{\mu} - (1 + R_f)}{\sigma^2 \gamma} - \frac{\mu_i - \mu_{-i}}{2 \sigma^2 \gamma} + \frac{\mu_i - \mu_{-i}}{2 \gamma \sigma^2} \right) (1 + R_f) \\
&\quad + \frac{\mu_i(\bar{\mu} - (1 + R_f))}{\sigma^2 \gamma} + \frac{\mu_i^2 - \mu_{-i} \mu_i}{2 \sigma^2 \gamma} - \frac{1}{2 \sigma^2 \gamma} \left(\frac{\mu_i + \mu_{-i}}{2} - (1 + R_f) + \frac{\mu_i - \mu_{-i}}{2} \right)^2, \\
&= \left(W_{-1}^i + \frac{K_0}{2} + \frac{K_0(\mu_i - \mu_{-i})}{4(\bar{\mu} - (1 + R_f))} \right) (1 + R_f) \\
&\quad + \frac{1}{2 \sigma^2 \gamma} \left(2(\mu_i - (1 + R_f))(\bar{\mu} - (1 + R_f)) + \mu_i^2 - \mu_{-i} \mu_i - (\mu_i - (1 + R_f))^2 \right), \\
&= \left(W_{-1}^i + \frac{K_0}{2} + \frac{K_0(\mu_i - \mu_{-i})}{4(\bar{\mu} - (1 + R_f))} \right) (1 + R_f) \\
&\quad + \frac{1}{2 \sigma^2 \gamma} \left((\mu_i - (1 + R_f))^2 + (\mu_i - \mu_{-i})(1 + R_f) \right).
\end{aligned}$$

□

Lemma 3.6. Setting $\theta_{BB}^u = \theta_D^*$,

$$\hat{P}_0(\theta_{BB}^u) = K_0.$$

And

$$\hat{\theta}_0^i(\theta_{BB}^u) = \hat{\theta}_0^i(\theta_D^*)(1 - \theta_D^*).$$

Proof of Lemma 3.6. Firstly we have:

$$\sqrt{\Delta(\theta_{BB}^u)/4} = \frac{2(\bar{\mu} - (1 + R_f))^2}{K_0 \gamma \sigma^2} - (\bar{\mu} - 2(1 + R_f))$$

And thus,

$$\begin{aligned}
\hat{P}_0(\theta_{BB}^u) &= \frac{\sigma^2 \gamma \theta_{BB}^u K_0 - \bar{\mu} \theta_{BB}^u - (1+R_f)(1 - \theta_{BB}^u) + \frac{2(\bar{\mu} - (1+R_f))^2}{K_0 \gamma \sigma^2} - (\bar{\mu} - 2(1+R_f))}{\sigma^2 \gamma (\theta_{BB}^u)^2}, \\
&= \frac{1}{\sigma^2 \gamma \left(1 - 2 \frac{\bar{\mu} - (1+R_f)}{K_0 \gamma \sigma^2}\right)^2} \left(\sigma^2 \gamma K_0 \left(1 - 2 \frac{\bar{\mu} - (1+R_f)}{K_0 \gamma \sigma^2}\right) - \bar{\mu} \left(1 - 2 \frac{\bar{\mu} - (1+R_f)}{K_0 \gamma \sigma^2}\right) \right. \\
&\quad \left. - (1+R_f) \left(2 \frac{\bar{\mu} - (1+R_f)}{K_0 \gamma \sigma^2}\right) + \frac{2(\bar{\mu} - (1+R_f))^2}{K_0 \gamma \sigma^2} - (\bar{\mu} - 2(1+R_f)) \right), \\
&= \frac{1}{\sigma^2 \gamma \left(1 - 2 \frac{\bar{\mu} - (1+R_f)}{K_0 \gamma \sigma^2}\right)^2} \left(\sigma^2 \gamma K_0 - 3(\bar{\mu} - (1+R_f)) - \bar{\mu} + (1+R_f) \right. \\
&\quad \left. + 4 \frac{(\bar{\mu} - (1+R_f))^2}{K_0 \gamma \sigma^2} \right), \\
&= \frac{K_0}{\left(1 - 2 \frac{\bar{\mu} - (1+R_f)}{K_0 \gamma \sigma^2}\right)^2} \left(1 - \frac{4(\bar{\mu} - (1+R_f))}{\sigma^2 \gamma K_0} + 4 \frac{(\bar{\mu} - (1+R_f))^2}{K_0^2 \gamma^2 \sigma^4} \right), \\
&= K_0.
\end{aligned}$$

So,

$$\begin{aligned}
\hat{\theta}_0^i(\theta_{BB}^u) &= \frac{1 - \theta_{BB}}{\sigma^2 \gamma (K_0 - \theta_{BB} P_0)} \left(\mu_i - \frac{P_0(1 - \theta_{BB})(1 + R_f)}{K_0 - \theta_{BB} P_0} \right), \\
\hat{\theta}_0^i(\theta_{BB}^u) &= \frac{\mu_i - (1 + R_f)}{\sigma^2 \gamma K_0},
\end{aligned}$$

In parallel,

$$\begin{aligned}
\hat{\theta}_0^i(\theta_D^*)(1 - \theta_D^*) &= \left(\frac{1}{2} + \frac{\mu_i - \mu_{-i}}{4(\bar{\mu} - (1 + R_f))} \right) 2 \frac{\bar{\mu} - (1 + R_f)}{\sigma^2 \gamma K_0}, \\
&= \frac{\mu_i + \mu_{-i} - 2(1 + R_f)}{2\sigma^2 \gamma K_0} + \frac{\mu_i - \mu_{-i}}{2\sigma^2 \gamma K_0}, \\
\hat{\theta}_0^i(\theta_D^*)(1 - \theta_D^*) &= \frac{\mu_i - (1 + R_f)}{\sigma^2 \gamma K_0}.
\end{aligned}$$

□

Proof of Theorem 3.4. We first show:

$$\sup_{\theta_{BB}} \bar{J}_{BB}(\theta_{BB}) \geq \sup_{\theta_D} \bar{J}_D(\theta_D).$$

Given the unique optimal dividends strategy θ_D^* , we show that θ_{BB}^u outlined is a replicating strategy, using results from the previous Lemma 3.6.

We first prove that implementing the pure repurchases strategy θ_{BB}^u , each investor receives an identical amount of cash at time 0 from rebalancing and redistribution as with θ_D^* executed. That is, we show:

$$\theta_D^* K_0 / 2 + (1/2 - \hat{\theta}_0^{D,i}) \hat{P}_0^D = (1/2 - \hat{\theta}_0^{B,i}) \hat{P}_0^B.$$

Indeed:

$$\begin{aligned} 0 &= \left(1 - 2 \frac{\bar{\mu} - (1 + R_f)}{\sigma^2 \gamma K_0}\right) \frac{K_0}{2} - \frac{\mu_i - \mu_{-i}}{4(\bar{\mu} - (1 + R_f))} \frac{\bar{\mu} - (1 + R_f)}{\gamma \sigma^2} \cdot 2 \\ &\quad - \left(1/2 - \frac{\mu_i - (1 + R_f)}{\sigma^2 \gamma K_0}\right) K_0, \\ &= -\frac{\bar{\mu} - (1 + R_f)}{\sigma^2 \gamma} - \frac{\mu_i - \mu_{-i}}{2\gamma \sigma^2} + \frac{\mu_i - (1 + R_f)}{\sigma^2 \gamma}, \\ &= \frac{-\mu_i - \mu_{-i}}{2\sigma^2 \gamma} + \frac{-\mu_i + \mu_{-i}}{2\sigma^2 \gamma} + \frac{2\mu_i}{2\sigma^2 \gamma} = 0, \end{aligned}$$

$$\begin{aligned} J_{BB,i}(\theta_{BB}^u) &= (W_{-1}^i + \left(\frac{1}{2} - \hat{\theta}_0^i(\theta_{BB}^u)\right) \hat{P}_0(\theta_{BB}^u))(1 + R_f) + \mu_i \hat{\theta}_0^i(\theta_{BB}^u) \frac{K_0 - \hat{P}_0(\theta_{BB}^u) \theta_{BB}^u}{1 - \theta_{BB}^u} \\ &\quad - \frac{\sigma^2 \gamma}{2} \hat{\theta}_0^i(\theta_{BB}^u)^2 \frac{(K_0 - \hat{P}_0(\theta_{BB}^u) \theta_{BB}^u)^2}{(1 - \theta_{BB}^u)^2}, \end{aligned}$$

$$\begin{aligned} J_{BB,i}(\theta_{BB}^u) &= \left(W_{-1}^i + \frac{\theta_D^* K_0}{2} + \hat{P}_0(\theta_D^*) \left(\frac{1}{2} - \hat{\theta}_0^i(\theta_D^*)\right)\right) (1 + R_f) \\ &\quad + \mu_i \hat{\theta}_0^i(\theta_{BB}^u) K_0 - \frac{\sigma^2 \gamma}{2} \hat{\theta}_0^i(\theta_{BB}^u)^2 K_0^2, \end{aligned}$$

$$\begin{aligned} J_{BB,i}(\theta_{BB}^u) &= \left(W_{-1}^i + \frac{\theta_D^* K_0}{2} + \hat{P}_0(\theta_D^*) \left(\frac{1}{2} - \hat{\theta}_0^i(\theta_D^*)\right)\right) (1 + R_f) + \mu_i \hat{\theta}_0^i(\theta_D^*) (1 - \theta_D^*) K_0 \\ &\quad - \frac{\sigma^2 \gamma}{2} \hat{\theta}_0^i(\theta_D^*)^2 (1 - \theta_D^*)^2 K_0^2, \end{aligned}$$

$$J_{BB,i}(\theta_{BB}^u) = J_{D,i}(\theta_D^*)$$

And thus,

$$\bar{J}_{BB}(\theta_{BB}^u) = \bar{J}_D(\theta_D^*).$$

As θ_{BB}^u was not chosen optimally:

$$\sup_{\theta_{BB}} \bar{J}_{BB}(\theta_{BB}) \geq \sup_{\theta_D} \bar{J}_D(\theta_D)$$

We now show the opposite side of the inequality:

$$\sup_{\theta_{BB}} \bar{J}_{BB}(\theta_{BB}) \leq \sup_{\theta_D} \bar{J}_D(\theta_D).$$

That is, given an optimal repurchases strategy θ_{BB}^* , let θ_D^u the replicating dividends strategy be:

$$\theta_D^u = \frac{\theta_{BB}^* \hat{P}_0(\theta_{BB}^*)}{K_0}.$$

That is, implementing θ_D^u , the same amount is redistributed through dividends as in the optimal buybacks implementation.

Define then $\tilde{\theta}_0^i = \frac{\hat{\theta}_0^i(\theta_{BB}^*)}{1 - \theta_{BB}^*}$ as the replicating strategy's ownership ratios. This is a valid ownership repartition as $(\tilde{\theta}_0^1, \tilde{\theta}_0^2)$ satisfies the market clearing condition:

$$\begin{aligned} \hat{\theta}_0^1(\theta_{BB}^*) + \hat{\theta}_0^2(\theta_{BB}^*) + \theta_{BB}^* &= 1, \\ \hat{\theta}_0^1(\theta_{BB}^*) + \hat{\theta}_0^2(\theta_{BB}^*) &= 1 - \theta_{BB}^*, \\ \frac{\hat{\theta}_0^1(\theta_{BB}^*)}{1 - \theta_{BB}^*} + \frac{\hat{\theta}_0^2(\theta_{BB}^*)}{1 - \theta_{BB}^*} &= 1, \\ \tilde{\theta}_0^1 + \tilde{\theta}_0^2 &= 1. \end{aligned}$$

The expected utility $\bar{J}_D(\theta_D^u)$ entailed by this replicating strategy match the one obtained implementing θ_{BB}^* :

$$\begin{aligned} \bar{J}_D(\theta_D^u, (\tilde{\theta}_0^1, \tilde{\theta}_0^2)) &= \left(W_{-1}^1 + W_{-1}^2 + K_0 \theta_D^u \right) (1 + R_f) + (\mu_1 \tilde{\theta}_0^1 + \mu_2 \tilde{\theta}_0^2) K_0 (1 - \theta_D^u) \\ &\quad - \frac{\sigma^2 \gamma}{2} ((\tilde{\theta}_0^1)^2 + (\tilde{\theta}_0^2)^2) K_0^2 (1 - \theta_D^u)^2, \\ \bar{J}_D(\theta_D^u, (\tilde{\theta}_0^1, \tilde{\theta}_0^2)) &= \left(W_{-1}^1 + W_{-1}^2 + \theta_{BB}^* \hat{P}_0(\theta_{BB}^*) \right) (1 + R_f) + (\mu_1 \hat{\theta}_0^1 + \mu_2 \hat{\theta}_0^2) \frac{K_0 - \theta_{BB}^* \hat{P}_0(\theta_{BB}^*)}{1 - \theta_{BB}^*} \\ &\quad - \frac{\sigma^2 \gamma}{2} ((\hat{\theta}_0^1)^2 + (\hat{\theta}_0^2)^2) \left(\frac{K_0 - \theta_{BB}^* \hat{P}_0(\theta_{BB}^*)}{1 - \theta_{BB}^*} \right)^2. \\ \bar{J}_D(\theta_D^u, (\tilde{\theta}_0^1, \tilde{\theta}_0^2)) &= \bar{J}_{BB}(\theta_{BB}^*, (\hat{\theta}_0^1, \hat{\theta}_0^2)) \end{aligned}$$

Since θ_D^u was not chosen optimally:

$$\sup_{\theta_D} \bar{J}_D(\theta_D) \geq \bar{J}_{BB}(\theta_{BB}^*).$$

□

We now present the results from the numerical modelling of our heterogeneous agents' setup. The starting wealths W_{-1} are set to 0. Both investors start with equal holdings

in the firm, $\theta_{-1}^1 = \theta_{-1}^2 = 1/2$.

For varying risk aversion γ (*figure 1.4, 1.6 - Left*) and optimist growth expectation μ_1 (*figure 1.5, 1.6 - Right*), we present the model's results at its optimal consistent equilibrium with either pure dividends (*figure 1.4, 1.5*) and share repurchases (*figure 1.6*) implemented.

Are thus maximised:

$$\bar{J}_D(\theta_D) = J_{D,1}(\theta_D) + J_{D,2}(\theta_D),$$

$$\bar{J}_{BB}(\theta_{BB}) = J_{BB,1}(\theta_{BB}) + J_{BB,2}(\theta_{BB}).$$

Where,

$$J_{D,i}(\theta_D) = (1+R_f)(W_{-1}^i + \theta_{-1}^i \theta_D K_0 + (\theta_{-1}^i - \theta_0^i) P_0) + \mu_i \theta_0^i (1 - \theta_D) K_0 - \frac{\gamma \sigma^2}{2} \theta_0^{i2} (1 - \theta_D)^2 K_0^2.$$

$$J_{BB,i}(\theta_{BB}) = (1+R_f)(W_{-1}^i + (\theta_{-1}^i - \theta_0^i) P_0) + \mu_i \frac{\theta_0^i (K_0 - \theta_{BB} P_0)}{1 - \theta_{BB}} - \frac{\gamma \sigma^2}{2} \left(\frac{\theta_0^i (K_0 - \theta_{BB} P_0)}{1 - \theta_{BB}} \right)^2.$$

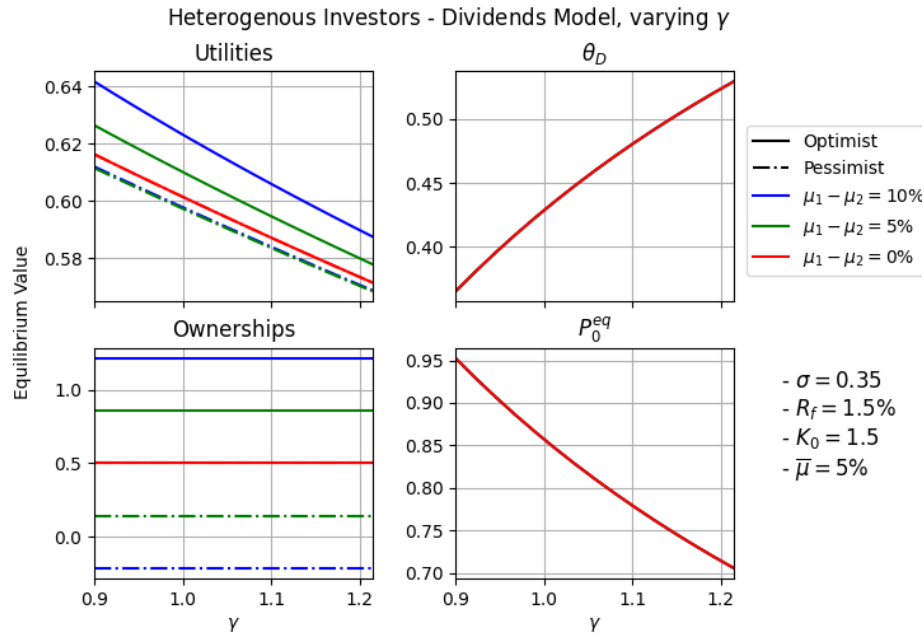


Figure 1.4: For varying risk aversion γ , is presented the equilibrium behaviour of the dividends model with two investors disagreeing over the company's prospects, for various values of $(\mu_1 - \mu_2)$ (High: Blue, Average: Green and None: Red).

We see that the dividends / share repurchases economic equivalence is maintained: both investors are equally rewarded whether the firm uses share repurchases (*see fig. 1.6*) or dividends (*see fig. 1.4 and 1.5*). Except for the company price, always stabilized to equal K_0 in the buybacks setting (*see fig. 1.6, Bottom-Right*), the models' optimal consistent equilibria are identical: same ownership structure and amounts spent on

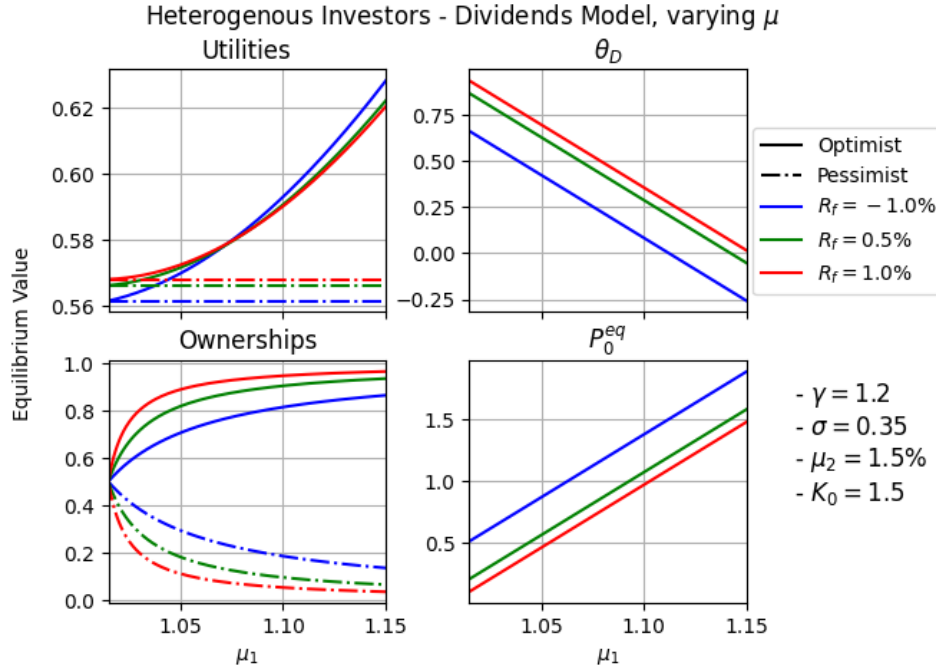


Figure 1.5: For varying optimistic expected growth rate μ_1 , is presented the equilibrium behaviour of the dividends model with two investors disagreeing over the company's prospects. Results are shown for increasing bank account growth rate R_f (Negative: Blue, Average: Green and High: Red).

redistribution. As in the one investor setting, as risk aversion increases, investor utility (see fig. 1.4, Top-Left), and company price (see fig. 1.4, Bottom-Right) decrease proportionally as risk aversion rises. The lower tolerance for risk reduces the company's competitiveness to the investors: its price is reduced and it redistributes more to investors.

In figure 1.4, the average growth prospect $\bar{\mu}$ does not change for varying distinct of $|\mu_1 - \mu_2|$. Thus, the company wide indicators such as price and redistribution ratio (see fig. 1.4, Top-Right) do not vary. Individual metrics depend on both μ_1 and μ_2 , but aggregate ones only depend on $\bar{\mu}$.

Note that it is not practical to compare the holdings of investors in the firm in the two setups as the share repurchases modify the total share supply. Instead, in the figures presented (see fig. 1.4, 1.5 and 1.6, Bottom-Left), we show the optimal ownerships rescaled for the available share supply after the shares bought back are retired.

While the ownership ratios without scaling are different, when taking into account the reduced share supply after repurchases are executed, we see that both investors hold the same effective ratio of the firm in both settings. Naturally, the optimist investor holds a larger stake in the firm than the pessimist investor. This gap in ownership ratios is proportional to the spread in investor expectations $|\mu_1 - \mu_2|$. Moreover, this gap is independent of the risk aversion factor γ .

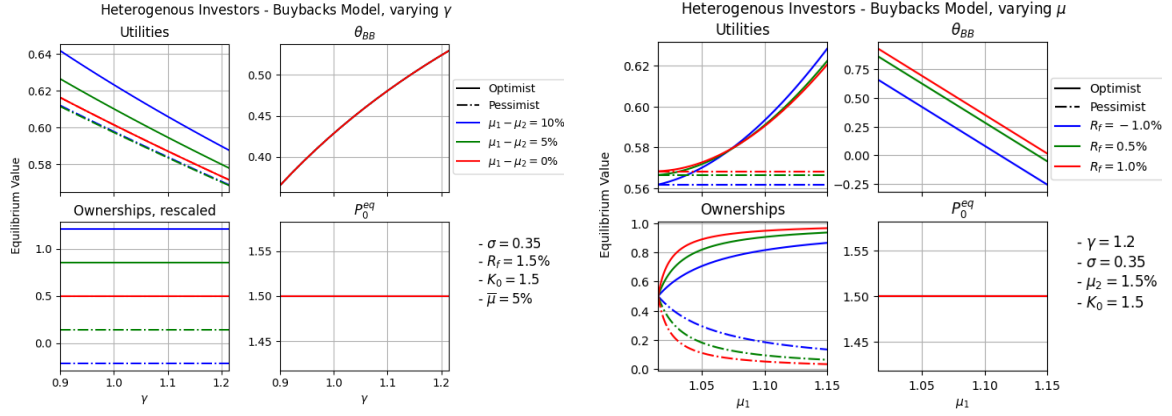


Figure 1.6: For varying risk aversion γ , is presented the equilibrium behaviour of the share repurchases model with two investors disagreeing over the company's prospects (*Left*). For varying optimistic expected growth rate μ_1 , is presented the equilibrium behaviour of the share repurchases model with two investors disagreeing over the company's prospects (*Right*).

Conversely, as μ_2 decreases, the stronger rebalancing allows for the pessimist investor to exit his position, and his utility is not affected significantly. Not obvious in figure 1.4, the pessimist's utility does change with $(\mu_1 - \mu_2)$, but this reduction but this reduction is small. Formally, the only term in the pessimist's utility depending on μ_2 is $(\mu_2 - (1 + R_f))^2$. As μ_2 and $(1 + R_f)$ are close, the dependence on μ_2 around R_f is minimal.

Figure 1.5 shows that each investor's utility is independent of the other's beliefs: as μ_1 varies, $J_{D,2}$ remains constant. Also, as μ_1 increases, the company redistributes less and less since the optimist prefers investing more in the firm. $\hat{\theta}_0^1$ increases too, and while $J_{D,2}$ does not depend on μ_1 , $\hat{\theta}_0^2$ does. The increasingly optimistic investor buys out his counterparty, up to the point where the optimistic investor goes into debt to increase his position in the firm by injecting money through a negative θ_D^* . As the cost of investor debt is controlled by R_f , for high growth expectations, the optimistic investor, and in turn both investors, benefit from a low risk-free rate (*see fig. 1.5, Top-Left*).

To summarize, the equivalence between dividends distribution and share repurchases stands when multiple investors with different beliefs in the firm. Yet, the same distinctions between the two models still hold: share repurchases allow the firm to redistribute cash to investors without impacting its share price. Share repurchases allow for the stabilization of the share price. With typical growth expectation, this entails that buybacks relatively increase the price compared to dividends. Yet, the equilibrium price in the pure dividends setup increases linearly with μ_1 (*see fig. 1.5 - Bottom-Right*). Thus, for large values of μ_1 , the dividends yield a price higher than the respective price in the repurchases setting. We recall:

$$\hat{P}_0(\theta_D^*) = 2 \frac{\bar{\mu} - (1 + R_f)}{\sigma^2 \gamma}.$$

Thus, the dividends equilibrium price is higher than the repurchases' exactly when:

$$\frac{K_0 \sigma^2 \gamma}{2} \leq \bar{\mu} - (1 + R_f).$$

In parallel, both investors expectedly rebalanced their portfolios in this setting. As one investor is more optimistic, he aims at owning more shares than the other pessimistic investor. Using dividends, this rebalancing is done with the investors' funds directly. In practice, this means that the optimistic investor needs to use debt or a substantial share of his own newly acquired wealth to buy the shares he is interested in from the other investor. Meanwhile, the firm buying back the shares itself enables the investors to spend none or a lesser part of their wealth on increasing their stake in the company. As most of the shares repurchased are bought back from the pessimistic investor, the effective optimist's share mechanically increases. By not imposing on the optimistic investor to advance the cost of the rebalancing, the share repurchases might be preferred by a company's board.

3.2 CONSUMPTION AND ENDOGENOUS INTEREST RATE

In this section, we revert to a single investor setup but now we further incorporate investor's consumption decision in order to endogenize the risk-free rate. We first describe how the portfolio dynamics in Section 1 change in presence of consumption. In addition to deciding θ_0 the number of shares to be held at the end of $t = 0$, the representative investor now also consumes a cash amount of C_0 . Their residual cash wealth at $t = 0$ is now given by

$$W_0 := W_{-1} + \theta_D K_0 + \theta_{BB} P_0 + (1 - \theta_{BB} - \theta_0) P_0 - C_0 = W_{-1} + \theta_D K_0 + (1 - \theta_0) P_0 - C_0,$$

which grows at the risk-free rate R_f .

At $t = 1$, the investor receives the terminal dividend and consumes all the available cash wealth. The consumption level at $t = 1$ is hence equal to

$$\begin{aligned} C_1 = W_1 &:= W_0(1 + R_f) + \theta_0 D_1 \\ &= (W_{-1} + \theta_D K_0 + (1 - \theta_0) P_0 - C_0)(1 + R_f) + \theta_0 \frac{(K_0 - \theta_D K_0 - \theta_{BB} P_0) \Pi}{1 - \theta_{BB}}. \end{aligned}$$

The representative investor now derives utility from intertemporal consumption (C_0, C_1) . Their expected utility value under a given investment-consumption strategy (θ_0, C_0) , the market price of firm P_0 , the prevailing risk-free rate R_f and the firm's distribution

strategy (θ_D, θ_{BB}) is now defined as

$$J(\theta_0, C_0; P_0, R_f, \theta_D, \theta_{BB}) := U(C_0) + e^{-\beta} \mathbb{E}[U(C_1)], \quad (1.10)$$

where $\beta > 0$ is the one-period subjective discount rate of the investor and as before we take $U(x) := \frac{1 - \exp(-\gamma x)}{\gamma}$.

Assuming that the money market instrument is in zero net supply, we introduce a modified definition of competitive equilibrium under which the risk-free rate is endogenized as well.

Definition 3.7. *Given the firm's distribution strategy (θ_D, θ_{BB}) , a competitive equilibrium with endogenized risk-free is the collection $(\hat{\theta}_0, \hat{C}_0, \hat{P}_0, \hat{R}_f)$ such that:*

1. *The representative investor maximizes their expected utility, i.e.*

$$J(\hat{\theta}_0, \hat{C}_0; \hat{P}_0, \hat{R}_f, \theta_D, \theta_{BB}) = \sup_{\theta_0, C_0} J(\theta_0, C_0; \hat{P}_0, \hat{R}_f, \theta_D, \theta_{BB}),$$

where J is defined in (1.10).

2. *Stock market clears such that*

$$1 = \hat{\theta}_0 + \theta_{BB}.$$

3. *Money market clears such that*

$$0 = W_{-1} = \theta_D K_0 + (1 - \hat{\theta}_0) \hat{P}_0 - \hat{C}_0.$$

We refer to $\hat{J}(\theta_D, \theta_{BB}) := J(\hat{\theta}_0, \hat{C}_0; \hat{P}_0, \hat{R}_f, \theta_D, \theta_{BB})$ as an equilibrium value function under the distribution strategy (θ_D, θ_{BB}) . $\hat{J}_D(\theta_D)$ and $\hat{J}_{BB}(\theta_{BB})$ are defined similarly as in Definition 1.2.

Proposition 3.8. *Consider an economy with endogenous risk-free rate.*

1. *Under a dividend strategy $(\theta_D, \theta_{BB} = 0)$, the unique competitive equilibrium is given by*

$$\hat{\theta}_0 = 1, \quad \hat{C}_0 = \theta_D K_0,$$

$$\hat{R}_f = \exp \left(\beta + \gamma K_0 \left(\mu(1 - \theta_D) - \theta_D - \frac{\sigma^2 \gamma}{2} K_0 (1 - \theta_D)^2 \right) \right) - 1,$$

and

$$\hat{P}_0 = \frac{K_0(1 - \theta_D)}{1 + \hat{R}_f} (\mu - \sigma^2 \gamma K_0 (1 - \theta_D)).$$

θ_D^* is characterized as the only solution to:

$$\ln(\mu - \gamma K_0 \sigma^2 (1 - \theta_D^*)) = \beta + K_0 \gamma (1 - \theta_D^*) \left(\mu - \frac{\sigma^2}{2} K_0 \gamma (1 - \theta_D^*) \right) - K_0 \gamma \theta_D^*.$$

2. Under a buyback strategy ($\theta_D = 0, \theta_{BB}$), and relying on Assumption 1.7, at least one competitive equilibrium exists and satisfies

$$\hat{\theta}_0 = 1 - \theta_{BB}, \quad \hat{C}_0 = \theta_{BB} \hat{P}_0,$$

$$\hat{P}_0 = \frac{2\theta_{BB}K_0\gamma\sigma^2 - \mu\theta_{BB} - (1 - \theta_{BB})(1 + R_f) + \sqrt{\Delta}}{2\gamma\sigma^2\theta_{BB}^2}.$$

$$\text{With } \Delta = (\mu\theta_{BB} + (1 - \theta_{BB})(1 + R_f))^2 - 4 \cdot \theta_{BB}K_0\gamma\sigma^2(1 - \theta_{BB})(1 + R_f).$$

With an exogenous or endogenous risk-free rate, an unique pure dividends strategy exists. For buybacks however, while an optimal strategy with an endogenized risk-free rate exists, we cannot guarantee its uniqueness. Multiple optimal values of the endogenized risk-free rate could exist. Yet, for each value of $\hat{R}_f$, at most one equilibrium $(\hat{\theta}_0, \hat{C}_0, \hat{P}_0)$ exists.

Appendix 1 provides further conditions on model's parameters to ensure Δ remains non-negative in the absence of Assumption 1.7.

Proof Proposition 3.8. Pure Dividends Setting

We first determine the optimal ownership ratio $\hat{\theta}_0$ as it satisfies $\frac{\partial J_D(\hat{\theta}_0, C_0; P_0, R_f, \theta_D)}{\partial \theta_0} = 0$.

$$\begin{aligned} 0 &= \frac{\partial J_D(\hat{\theta}_0)}{\partial \theta_0}, \\ \iff 0 &= e^{-\beta}(1 + R_f)P_0 + e^{-\beta}\mu K_0(1 - \theta_D) - e^{-\beta}\gamma\sigma^2 K_0^2(1 - \theta_D)^2\theta_0, \\ \text{Thus, } \hat{\theta}_0 &= \frac{1}{\gamma\sigma^2 K_0^2(1 - \theta_D)^2} \left(\mu K_0(1 - \theta_D) - (1 + R_f)P_0 \right), \\ \hat{\theta}_0 &= \frac{1}{K_0\gamma\sigma^2(1 - \theta_D)} \left(\mu - \frac{P_0(1 + R_f)}{K_0(1 - \theta_D)} \right). \end{aligned}$$

Analogously, the optimal consumption level at time 0, $\hat{C}_0$, satisfies $\frac{\partial J_D(\theta_0, \hat{C}_0; P_0, R_f, \theta_D)}{\partial C_0} = 0$.

$$\begin{aligned} \frac{\partial J_D(\hat{C}_0)}{\partial C_0} &= e^{-\gamma C_0} - (1 + R_f) \exp \left[-\beta - \gamma(1 + R_f)(W_{-1} - C_0 + \theta_D\theta_{-1}K_0 + P_0(\theta_{-1} - \theta_0)) \right. \\ &\quad \left. - \gamma\theta_0 K_0\mu(1 - \theta_D) + \frac{\sigma^2}{2}\gamma^2 K_0^2\theta_0^2(1 - \theta_D)^2 \right]. \end{aligned}$$

Thus,

$$\hat{C}_0 = \frac{1}{\gamma(2+R_f)} \left(\beta - \ln(1+R_f) + \gamma \left((1+R_f)(W_{-1} + \theta_D \theta_{-1} K_0 + P_0(\theta_{-1} - \theta_0)) \right. \right. \\ \left. \left. + \theta_0(1-\theta_D)K_0\mu - \frac{\gamma\sigma^2}{2}K_0^2\theta_0^2(1-\theta_D)^2 \right) \right).$$

The market clearing condition is $\hat{\theta}_0 = 1$ as only dividends are executed with one representative investor:

$$1 = \frac{1}{K_0\gamma\sigma^2(1-\theta_D)} \left(\mu - \frac{\hat{P}_0(1+R_f)}{K_0(1-\theta_D)} \right),$$

which gives

$$\hat{P}_0 = \frac{K_0(1-\theta_D)}{1+R_f} \left(\mu - K_0\gamma\sigma^2(1-\theta_D) \right).$$

We now attempt to endogenize the risk-free rate R_f . We set $\theta_{-1} = 1$. For now, we set $W_{-1} = 0$.

To endogenize the risk-free rate, it is chosen so that the investor's optimal behaviour is to spend his entire income at time $t = 0$ immediately: we set $\hat{C}_0 = \theta_D K_0$.

In the expression of $\hat{C}_0$, replacing P_0 and θ_0 by their equilibrium values $\hat{P}_0$ and $\hat{\theta}_0$,

$$\theta_D K_0 = \frac{1}{\gamma(2+\hat{R}_f)} \left(\beta + \gamma \left((1+\hat{R}_f)\theta_D K_0 + (1-\theta_D)K_0\mu - \frac{\gamma}{2}(1-\theta_D)^2 K_0^2 \sigma^2 \right) - \ln(1+\hat{R}_f) \right).$$

The endogenized risk-free rate becomes,

$$(1 + \hat{R}_f) = \exp \left(\beta + \gamma K_0 \left(\mu(1-\theta_D) - \theta_D - \frac{\gamma\sigma^2}{2}K_0(1-\theta_D)^2 \right) \right).$$

Removing the assumption ($W_{-1} = 0$), we cannot explicitly solve the previous equation for R_f :

$$\ln(1 + \hat{R}_f) - \gamma W_{-1}(1 + \hat{R}_f) = \beta + \gamma \left((1-\theta_D)K_0\mu - \theta_D K_0 - \frac{\gamma}{2}(1-\theta_D)^2 K_0^2 \sigma^2 \right).$$

Endogenizing the dividends distribution ratio θ_D by maximising the investor total expected utility at equilibrium, $J_D(\hat{\theta}_0, \hat{C}_0; \hat{P}_0, \hat{R}_f, \theta_D)$:

$$J_D(\theta_D) = \frac{1}{\gamma} - \frac{e^{-\gamma\theta_D K_0}}{\gamma} + \frac{e^{-\beta}}{\gamma} - \frac{e^{-\beta}}{\gamma} \exp \left(-\gamma\theta_0 K_0\mu(1-\theta_D) + \frac{\sigma^2}{2}\gamma^2 K_0^2 \theta_0^2 (1-\theta_D)^2 \right).$$

And its first order derivative then is:

$$\begin{aligned}\frac{\partial J_D(\theta_D)}{\partial \theta_D} &= K_0 \left(e^{-\gamma \theta_D K_0} - (\mu - \gamma K_0 \sigma^2 (1 - \theta_D)) e^{-\gamma (1 - \theta_D) K_0 (\mu - \frac{\gamma}{2} (1 - \theta_D) K_0 \sigma^2) - \beta} \right), \\ &= K_0 e^{-\gamma K_0 \theta_D} \left\{ 1 - [\gamma K_0 \mu - \gamma^2 \sigma^2 K_0^2 (1 - \theta_D)] \cdot \right. \\ &\quad \left. \exp \left(-\beta - \gamma K_0 \mu (1 - \theta_D) + \frac{\gamma^2 \sigma^2}{2} K_0^2 (1 - \theta_D)^2 + \gamma K_0 \theta_D \right) \right\}.\end{aligned}$$

That is, θ_D^* is characterized by

$$\ln(\mu - \gamma K_0 \sigma^2 (1 - \theta_D^*)) = \beta + K_0 \gamma (1 - \theta_D^*) \left(\mu - \frac{\sigma^2}{2} K_0 \gamma (1 - \theta_D^*) \right) - K_0 \gamma \theta_D^*. \quad (1.11)$$

While we cannot exhibit the optimal dividends ratio analytically, we can prove there exists a unique maximizer to the investor's utility, that is, the above equation has a unique solution.

It is clear that $\partial_{\theta_D} J_D(\theta_D) > 0$ if $(\gamma K_0 \mu - \gamma^2 \sigma^2 K_0^2 (1 - \theta_D) \leq 0)$ or equivalently $(\theta_D \leq 1 - \frac{\mu}{\gamma \sigma^2 K_0})$, i.e. J_D must be increasing on $\theta_D \leq 1 - \frac{\mu}{\gamma \sigma^2 K_0}$.

Meanwhile, the exponent inside $\exp \left(-\beta - \gamma K_0 \mu (1 - \theta_D) + \frac{\gamma^2 \sigma^2}{2} K_0^2 (1 - \theta_D)^2 + \gamma K_0 \theta_D \right)$ is a convex quadratic function in θ_D . Hence it is easy to see that this function is increasing on $\theta_D \geq 1 - \frac{\mu+1}{\gamma \sigma^2 K_0}$. Then, for $\theta_D \geq 1 - \frac{\mu}{\gamma \sigma^2 K_0} > 1 - \frac{\mu+1}{\gamma \sigma^2 K_0}$, the expression

$$1 - [\gamma K_0 \mu - \gamma^2 \sigma^2 K_0^2 (1 - \theta_D)] \exp \left(-\beta - \gamma K_0 \mu (1 - \theta_D) + \frac{\gamma^2 \sigma^2}{2} K_0^2 (1 - \theta_D)^2 + \gamma K_0 \theta_D \right)$$

is strictly decreasing in θ_D because its second term is the product of two strictly positive increasing functions. Moreover, this expression tends to $-\infty$ as θ_D tends to ∞ . We therefore conclude that $\partial_{\theta_D} J_D(\theta_D)$ must change sign exactly once from positive to negative at some $\theta_D^* > \frac{\mu}{\gamma \sigma^2 K_0}$. This turning point is the unique maximizer of the function $J_D(\theta_D)$.

Pure Share Repurchases Setting

Similarly, the optimal ownership ratio $\hat{\theta}_0$ in the pure share repurchases setting satisfies $\frac{\partial J_{BB}(\hat{\theta}_0, C_0; P_0, R_f, \theta_{BB})}{\partial \theta_0} = 0$. We obtain:

$$\hat{\theta}_0 = \frac{1 - \theta_{BB}}{\gamma \sigma^2 (K_0 - P_0 \theta_{BB})} \left(\mu - \frac{P_0 (1 + R_f) (1 - \theta_{BB})}{K_0 - P_0 \theta_{BB}} \right).$$

Using the market clearing hypothesis: $\hat{\theta}_0 + \theta_{BB} = 1$, we obtain the company's price, solving a quadratic polynomial and taking the larger root. The other root does not behave consistently when $\theta_{BB} \rightarrow \{0, 1\}$, as shown in previously, in Section 1. Given

Assumption 1.7, Δ is always non-negative and the consistent equilibrium price is thus always real.

$$\hat{P}_0 = \frac{2\theta_{BB}K_0\gamma\sigma^2 - \mu\theta_{BB} - (1 - \theta_{BB})(1 + R_f) + \sqrt{\Delta}}{2\gamma\sigma^2\theta_{BB}^2}.$$

$$\text{With } \Delta = (\mu\theta_{BB} + (1 - \theta_{BB})(1 + R_f))^2 - 4 \cdot \theta_{BB}K_0\gamma\sigma^2(1 - \theta_{BB})(1 + R_f).$$

We obtain C_0 as we obtained $\hat{\theta}_0$, by using that at equilibrium, $\frac{J_{BB}(\hat{\theta}_0, C_0; P_0, R_f, \theta_{BB})}{\partial C_0} = 0$.

$$\frac{\partial J_{BB}(\hat{C}_0)}{\partial C_0} = e^{-\gamma\hat{C}_0} - (1+R_f) \exp\left(-\beta - \gamma(1+R_f)(W_{-1} - \hat{C}_0 - (\theta_0 - \theta_{-1})P_0)\right) \cdot D_1,$$

$$\text{So } \hat{C}_0 = \frac{-1}{\gamma(2+R_f)} \left(\ln(1+R_f) - \beta - \gamma(1+R_f)(W_{-1} - P_0(\theta_0 - \theta_{-1})) - \ln(D_1) \right),$$

$$\begin{aligned} \hat{C}_0 = \frac{-1}{\gamma(2+R_f)} & \left(\ln(1+R_f) - \beta - \gamma(1+R_f)(\theta_{-1} - \theta_0)P_0 - \mu\gamma\theta_0 \frac{K_0 - \theta_{BB}P_0}{1 - \theta_{BB}} \right. \\ & \left. + \frac{\sigma^2\gamma^2\theta_0^2(K_0 - \theta_{BB}P_0)^2}{2(1 - \theta_{BB})^2} \right). \end{aligned}$$

Where we used D_1 as the expected utility term corresponding to the dividend at time 1:

$$D_1 = \mathbb{E} \left[\exp \left(\frac{-\gamma\theta_0}{1 - \theta_{BB}} \Pi \right) \right] = \exp \left(-\mu\gamma\theta_0 \frac{K_0 - \theta_{BB}P_0}{1 - \theta_{BB}} + \frac{\sigma^2\gamma^2\theta_0^2(K_0 - \theta_{BB}P_0)^2}{2(1 - \theta_{BB})^2} \right).$$

From this, replacing first θ_0 by its equilibrium value $\hat{\theta}_0$, we obtain:

$$0 = -\ln(1 + \hat{R}_f) + \beta + \mu\gamma K_0 - (\mu + 1)\gamma\theta_{BB}\hat{P}_0 - \frac{\sigma^2\gamma^2}{2}(K_0 - \theta_{BB}\hat{P}_0)^2,$$

It is not possible to obtain a closed form formula for $\hat{R}_f$ by substituting $\hat{P}_0$ by its closed formula.

Let's define this expression as $\chi(R_f)$:

$$\chi(R_f, \theta_{BB}) := -\ln(1 + R_f) + \beta + \mu\gamma K_0 - (\mu + 1)\gamma\theta_{BB}\hat{P}_0 - \frac{\sigma^2\gamma^2}{2}(K_0 - \theta_{BB}\hat{P}_0)^2.$$

$\hat{R}_f$ is an endogenized value of the risk-free rate for a given θ_{BB} iff $\chi(\hat{R}_f, \theta_{BB}) = 0$.

To finish the proof of the proposition, we need to ensure that for any given θ_{BB} , we can find $\hat{R}_f$ such that $\chi(\hat{R}_f, \theta_{BB}) = 0$. We again rely on the intermediate value theorem. Since we are working with Assumption 1.7, $\hat{P}_0$ is a continuous function over $\mathbb{R}$ of R_f . Thus, χ is a continuous function of R_f over $(-1, \infty)$.

$$\lim_{R_f \rightarrow -1} \Delta = \mu^2 \theta_{BB}^2, \text{ thus } \lim_{R_f \rightarrow -1} \hat{P}_0 = \frac{K_0}{\theta_{BB}}.$$

Consequently,

$$\forall \theta_{BB} \in (0, 1), \lim_{R_f \rightarrow -1} \chi(R_f, \theta_{BB}) = \infty.$$

Setting $(\theta_{BB} \hat{P}_0)$ to 0 for $\theta_{BB} = 0$ in order to continuously extend its value at 0, we can extend the previous statement to $\theta_{BB} \in \mathbb{R}$. Indeed, we showed in Section 1 that $\hat{P}_0$ is a consistent price equilibrium: $\lim_{\theta_{BB} \rightarrow 0} \theta_{BB} \hat{P}_0 = 0$.

As a function of R_f , $\hat{P}_0$ is the sum of a linear and of the square root of another non-negative polynomial. As such, as R_f tends to ∞ , either $\hat{P}_0$ has a well-defined limit, either finite, or infinite. If it is finite, we have:

$$\lim_{R_f \rightarrow \infty} \chi(R_f, \theta_{BB}) = \lim_{R_f \rightarrow \infty} -\ln(1 + R_f) + \text{Constant} = -\infty$$

If $\lim_{R_f \rightarrow \infty} |\hat{P}_0| = \infty$, then:

$$\lim_{R_f \rightarrow \infty} \left(-(\mu + 1)\gamma \theta_{BB} \hat{P}_0 - \frac{\sigma^2 \gamma^2}{2} (K_0 - \theta_{BB} \hat{P}_0)^2 \right) = -\infty.$$

Thus,

$$\lim_{R_f \rightarrow \infty} C(R_f, \theta_{BB}) = -\infty.$$

Since χ is a continuous function of R_f over $(-1, \infty)$, and its range is $\mathbb{R}$ for any given θ_{BB} , there exists $\hat{R}_f > -1$ such that $\chi(\hat{R}_f, \theta_{BB}) = 0$. As such, $\hat{R}_f$ guarantees that $\hat{C}_0 = \theta_{BB} \hat{P}_0$.

While we do not obtain a closed form formula for $\hat{R}_f$, its existence guarantees that at least one optimal θ_{BB}^* exists. Indeed, θ_{BB}^* is a maximizer of the one-dimensional function $\hat{J}_{BB}(\theta_{BB})$. $\square$

The next theorem shows that the equivalence of dividend and buyback still holds even if we endogenize the risk-free rate.

Theorem 3.9. *Under an economy with endogenous risk-free rate, dividend and buyback are economically equivalent such that*

$$\sup_{\theta_D} \hat{J}_D(\theta_D) = \sup_{\theta_{BB}} \hat{J}_{BB}(\theta_{BB}).$$

Furthermore, optimal repurchases strategies redistribute an identical amount of equity to the investors as with the unique optimal dividends strategy, i.e.

$$K_0\theta_D^* = \theta_{BB}^*\hat{P}_0(\theta_{BB}^*).$$

We want to show that for the optimal repurchases and dividends ratios θ_{BB}^* & θ_D^* and corresponding endogenized risk-free rates $\hat{R}_f(\theta_{BB}^*)$ and $\hat{R}_f(\theta_D^*)$, that the representative investor's utility is identical:

$$\sup_{\theta_D} \hat{J}_D(\theta_D) = \sup_{\theta_{BB}} \hat{J}_{BB}(\theta_{BB}).$$

Proof of Theorem 3.9, Dividends Replication of Optimal Buybacks Strategy.

We first show that dividends strategies dominate share repurchases strategies.

To this end, let us first assume that $(\theta_{BB}^*, \hat{R}_f(\theta_{BB}^*))$ maximises the investor's expected utility in the buybacks setting. With this equilibrium, we also have $\hat{C}_0(\theta_{BB}^*, \hat{R}_f(\theta_{BB}^*)) = \theta_{BB}^* \cdot \hat{P}_0$.

Let us construct a replicating pure dividends strategy by first picking θ_D^u as:

$$\theta_D^u = \frac{\theta_{BB}^* \cdot \hat{P}_0}{K_0}.$$

This definition of θ_D^u gets the firm in debt ($\theta_D^u > 1$) or inject cash into the firm ($\theta_D^u < 0$) if and only if the optimal amount redistributed in the buybacks setting, $\theta_{BB}^* \cdot \hat{P}_0$ also does.

The expected utility of the investor under these two setups thus becomes:

$$\begin{aligned} \hat{J}_{BB}(\theta_{BB}^*) &= \frac{1 - e^{-\gamma \theta_{BB}^* \hat{P}_0}}{\gamma} + \frac{e^{-\beta}}{\gamma} - \frac{e^{-\beta}}{\gamma} \exp \left(-\gamma \mu (K_0 - \theta_{BB}^* \hat{P}_0) + \frac{\sigma^2 \gamma^2}{2} (K_0 - \theta_{BB}^* \hat{P}_0)^2 \right). \\ \hat{J}_D(\theta_D^u) &= \frac{1 - e^{-\gamma \hat{C}_0(\theta_D^u, R_f)}}{\gamma} + \frac{e^{-\beta}}{\gamma} - \frac{e^{-\beta}}{\gamma} \exp \left(-\gamma (1 + R_f) (\theta_{BB}^* \hat{P}_0 - \hat{C}_0(\theta_D^u, R_f)) \right. \\ &\quad \left. - \gamma \mu (K_0 - \theta_{BB}^* \hat{P}_0) + \frac{\gamma^2 \sigma^2}{2} (K_0 - \theta_{BB}^* \hat{P}_0)^2 \right). \end{aligned}$$

Where we used that $\theta_D^u K_0 = \theta_{BB}^* \hat{P}_0$, the firm's size after redistribution is $(K_0 - \theta_{BB}^* \hat{P}_0)$ as in the buybacks setting.

We endogenize the risk-free rate in the dividends replicating strategy:

$$\hat{R}_f(\theta_D^u) = \exp \left(\beta + \gamma K_0 \left(\mu (1 - \theta_D^u) - \theta_D^u - \frac{\gamma \sigma^2}{2} K_0 (1 - \theta_D^u)^2 \right) \right) - 1.$$

This ensures that $\hat{C}_0(\theta_D^u, \hat{R}_f(\theta_D^u)) = \theta_D^u K_0 = \theta_{BB}^* \cdot \hat{P}_0$. Thus, $\hat{J}_D(\theta_D^u)$ becomes:

$$\begin{aligned}\hat{J}_D(\theta_D^u) &= \frac{1 - e^{-\gamma \theta_{BB}^* \cdot \hat{P}_0}}{\gamma} + \frac{e^{-\beta}}{\gamma} - \frac{e^{-\beta}}{\gamma} \exp \left(-\gamma \mu (K_0 - \hat{P}_0 \cdot \theta_{BB}^*) + \frac{\gamma^2 \sigma^2}{2} (K_0 - \hat{P}_0 \cdot \theta_{BB}^*)^2 \right). \\ \hat{J}_D(\theta_D^u) &= \hat{J}_{BB}(\theta_{BB}^*).\end{aligned}$$

Since θ_{BB}^* was chosen optimally, we have that:

$$\sup_{\theta_D} \hat{J}_D(\theta_D) \geq \sup_{\theta_{BB}} \hat{J}_{BB}(\theta_{BB}).$$

□

Before showing $\sup_{\theta_{BB}} \hat{J}_{BB}(\theta_{BB}) \geq \sup_{\theta_D} \hat{J}_D(\theta_D)$, we prove some preliminary lemmas.

Lemma 3.10. *Given a fixed risk-free rate R_f , let θ_D^0 be the unique optimal dividends ratio if the investor is not allowed consumption at time $t = 0$, and consequently with β set to 0.*

Let θ_D^ be the optimal dividends ratio with optimal non-zero consumption at time $t = 0$ with the same risk-free rate. Then,*

$$\theta_D^0 = \theta_D^*.$$

The dividends strategy θ_D^ does not endogenize the risk-free rate: cash is not in zero net supply.*

Proof.

$$\begin{aligned}J_D(\theta_D, C_0) &= \frac{1 - e^{-\gamma C_0}}{\gamma} + \frac{e^{-\beta}}{\gamma} - \frac{e^{-\beta - \gamma(1+R_f)(\theta_D K_0 - C_0)}}{\gamma} \cdot \mathbb{E}[U(K_0(1 - \theta_D)\Pi)], \\ J_D(\theta_D, 0) &= \frac{1}{\gamma} - \frac{1}{\gamma} \exp(-\gamma(1 + R_f)\theta_D K_0) \cdot \mathbb{E}[U(K_0(1 - \theta_D)\Pi)].\end{aligned}$$

In this setting, we have:

$$\frac{\partial J_D(\theta_D, 0)}{\partial \theta_D} = 0 \iff \frac{\partial J_D(\theta_D, C_0)}{\partial \theta_D} = 0,$$

Given C_0 fixed, the optimal level of redistribution does not change.

In parallel, we have that $\sup_{\theta_D} (\sup_{C_0} J_D(\theta_D, C_0)) = \sup_{C_0} (\sup_{\theta_D} J_D(\theta_D, C_0))$.

Instead of expressing the optimal consumption C_0^* as a function of θ_D , we can then express the optimal dividends ratio θ_D^* as a function of the consumption C_0 .

For any consumption level C_0 , the optimal level of redistribution is identical to θ_D^0 , the optimal redistribution level with $C_0 = 0$. Thus, for the optimal C_0^* , we still have that $\theta_D^*(C_0^*) = \theta_D^0$.

In short, allowing the investor to consume at time 0, $C_0 \neq 0$, does not change the optimal level of redistribution. $\square$

Lemma 3.11. *For any set of model parameters $(\beta, \gamma, K_0, \mu, \sigma, R_f)$, let θ_D^* be the optimal dividends strategy with consumption at time 0 and exogenous risk-free rate.*

We can find a buybacks ratio $\theta_{BB}^u \in \mathbb{R}$ that yields the same amount of redistribution to investors as θ_D^ . That is:*

$$\forall(\beta, \gamma, K_0, \mu, \sigma, R_f), \exists \theta_{BB}^u \in \mathbb{R} \text{ such that } \theta_{BB}^u \cdot \hat{P}_0(\theta_{BB}^u) = \theta_D^* \cdot K_0.$$

And the defined buybacks strategy θ_{BB}^u is such that:

$$\hat{C}_0(\theta_{BB}^u) = \hat{C}_0(\theta_D^*).$$

And implementing θ_D^ or θ_{BB}^u is economically indifferent:*

$$J(\hat{\theta}_0, \hat{C}_0(\theta_D^*), \hat{P}_0, R_f, \theta_D^*) = J(\hat{\theta}_0, \hat{C}_0(\theta_{BB}^u), \hat{P}_0, R_f, \theta_{BB}^u)$$

Proof. Lemma 3.10 entails that the optimal dividends redistribution level with consumption θ_D^* equals the one without consumption, θ_D^0 .

In the buybacks setting, the company's price as a function of other variables does not change whether consumption at time 0 is implemented or not:

$$\hat{P}_0(\theta_{BB}) = \frac{2\theta_{BB}K_0\gamma\sigma^2 - \mu\theta_{BB} - (1 - \theta_{BB})(1 + R_f) + \sqrt{\Delta}}{2\gamma\sigma^2\theta_{BB}^2}$$

$$\text{With } \Delta = (\mu\theta_{BB} + (1 - \theta_{BB})(1 + R_f))^2 - 4 \cdot \theta_{BB}K_0\gamma\sigma^2(1 - \theta_{BB})(1 + R_f).$$

In the previous setting without consumption at time 0, Section 1, we have that setting the buybacks ratio $\theta_{BB}^u = \theta_D^0$, $\hat{P}_0(\theta_{BB}^u) = K_0$.

Setting the replicating buybacks ratio θ_{BB}^u to θ_D^* , we then automatically get that the replicating buybacks strategy spends the same amount on redistribution to investors as the optimal dividends strategy.

θ_{BB}^u was defined so that the amount redistributed through buybacks equals the optimal amount spent on dividends. Let $S(\theta_D^*) = K_0\theta_D^*$ be the optimal amount redistributed

by the company in the dividends setting and in the replicating buybacks strategy.

$$\begin{aligned}
J(\hat{\theta}_0, C_0, \hat{P}_0, R_f, \theta_D^*) &= \frac{1 - e^{-\gamma C_0}}{\gamma} + \frac{e^{-\beta}}{\gamma} \\
&\quad - \frac{e^{-\beta}}{\gamma} \exp\left(\gamma(1+R_f)(C_0 - S(\theta_D^*))\right) \cdot \mathbb{E}\left[e^{-\gamma\Pi(K_0 - S(\theta_D^*))}\right], \\
J(\hat{\theta}_0, C_0, \hat{P}_0, R_f, \theta_{BB}^u) &= \frac{1 - e^{-\gamma C_0}}{\gamma} + \frac{e^{-\beta}}{\gamma} \\
&\quad - \frac{e^{-\beta}}{\gamma} \exp\left(\gamma(1+R_f)(C_0 - S(\theta_D^*))\right) \cdot \mathbb{E}\left[e^{-\gamma\Pi(K_0 - S(\theta_D^*))}\right].
\end{aligned}$$

Thus,

$$J(\hat{\theta}_0, C_0, \hat{P}_0, R_f, \theta_{BB}^u) \equiv J(\hat{\theta}_0, C_0, \hat{P}_0, R_f, \theta_D^*).$$

From this, the two functions' respective suprema are attained for the same $\hat{C}_0$, and are equal:

$$\begin{aligned}
J(\hat{\theta}_0, \hat{C}_0(\theta_{BB}^u), \hat{P}_0, R_f, \theta_{BB}^u) &= J(\hat{\theta}_0, \hat{C}_0(\theta_D^*), \hat{P}_0, R_f, \theta_D^*), \\
\hat{C}_0(\theta_{BB}^u) &= \hat{C}_0(\theta_D^*).
\end{aligned}$$

□

For any given value of R_f , it is possible to find a repurchases strategy replicating the optimal dividends one. We now extend this result to an endogenized risk-free rate.

*Proof of Theorem 3.9, **Buybacks Replication of Dividends Optimal Strategy.***

To prove that $\sup_{\theta_{BB}} \hat{J}_{BB}(\theta_{BB}) \geq \sup_{\theta_D} \hat{J}_D(\theta_D)$, we work in the “Partial Equilibrium” endogenization setting presented in Appendix 2. In the dividends setting, Appendix 2 proves the equivalence of the two endogenization methodologies. Let $\tilde{J}_D(\theta_D)$ and $\tilde{J}_{BB}(\theta_{BB})$ be the investors' expected utilities in the “partial equilibrium” setting, given θ_D and θ_{BB} . Using Proposition 3.8,

$$\forall \theta_{BB}, \exists \hat{R}_f \text{ s.t. } \hat{C}_0 = \theta_{BB} \hat{P}_0.$$

As $\hat{P}_0$ and $\hat{C}_0$ are defined identically in the “full equilibrium” and “partial equilibrium”, it is also possible to endogenize the risk-free rate for all values of θ_{BB} in the “partial equilibrium”.

Let $(\tilde{R}_f, \tilde{\theta}_{BB})$ be the optimal buybacks strategy in the “partial-equilibrium”. As it endogenizes the risk-free rate and ensures its price satisfies the market clearing condition $\hat{\theta}_0 + \theta_{BB} = 1$, it is also an acceptable strategy in the “full-equilibrium” setting:

$$\begin{aligned} \sup_{\theta_{BB}} J(\hat{\theta}_0, \hat{C}_0, \hat{P}_0, \hat{R}_f, \theta_{BB}) &\geq J(\hat{\theta}_0, \hat{C}_0, \hat{P}_0, \tilde{R}_f, \tilde{\theta}_{BB}), \\ \sup_{\theta_{BB}} \hat{J}_{BB}(\theta_{BB}) &\geq \sup_{\theta_{BB}} \tilde{J}_{BB}(\theta_{BB}). \end{aligned}$$

Appendix 2 shows $(\sup_{\theta_D} \tilde{J}_D(\theta_D) = \sup_{\theta_D} \hat{J}_D(\theta_D))$. As such, given θ_D^* the maximizer of $\tilde{J}_D(\theta_D)$, exhibiting θ_{BB}^u such that $\tilde{J}_D(\theta_D^*) = \tilde{J}_{BB}(\theta_{BB}^u)$ is enough to obtain the desired:

$$\sup_{\theta_{BB}} \hat{J}_{BB}(\theta_{BB}) \geq \sup_{\theta_D} \hat{J}_D(\theta_D).$$

Let $(\hat{R}_f, \theta_D^*, \hat{C}_0(\theta_D^*))$ be an optimal dividends strategy:

$$\tilde{J}_D(R_f) = J(\hat{\theta}_0, \hat{C}_0(\theta_D^*), \hat{P}_0, \hat{R}_f, \theta_D^*).$$

The risk-free rate ensures that cash is in zero net-supply: $\hat{C}_0(\theta_D^*(\hat{R}_f)) = S(\theta_D^*)$. We recall that we defined $S(\theta_D^*) := K_0 \theta_D^*$.

From Lemma 3.11, let $\theta_{BB}^u = \theta_D^*$ be the replicating strategy of the optimal dividends strategy with risk-free rate $\hat{R}_f$. This replicating strategy has consumption level $\hat{C}_0(\theta_{BB}^u)$ equal to $S(\theta_D^*)$.

Lemma 3.11 also guarantees:

$$J(\hat{\theta}_0, S(\theta_D^*), \hat{P}_0, \hat{R}_f, \theta_D^*) = J(\hat{\theta}_0, S(\theta_D^*), \hat{P}_0, \hat{R}_f, \theta_{BB}^u).$$

Finally, Lemma 3.11 has $\hat{P}_0(\theta_{BB}^u) = K_0$. Thus, $\hat{C}_0(\theta_{BB}^u) = \theta_{BB}^u \hat{P}_0$. The investor’s income from the repurchases a time 0 is precisely the optimal amount consumed. So, $(\hat{R}_f, \theta_{BB}^u)$ is a valid share repurchases strategy endogenizing the risk-free rate.

$\hat{R}_f$ was not chosen optimally: $\sup_{R_f} \tilde{J}_{BB}(R_f) \geq J(\hat{\theta}_0, \hat{C}_0(\theta_{BB}^u), \hat{P}_0, \hat{R}_f, \theta_{BB}^u)$.

$$\sup_{R_f} \tilde{J}_{BB}(R_f) \geq \sup_{R_f} \tilde{J}_D(R_f).$$

□

In the following figures, we exhibit that, as for an exogenous R_f , at equilibrium with an endogenized risk-free rate, the company spends the same amount on optimal dividends as it does on optimal buybacks, with $\theta_D^* = \theta_{BB}^*$. The optimal investor utility does not depend on the mode of redistribution used, nor does the amount spent by the company on redistributing assets to investors.

While in theory multiple values of $\hat{R}_f$ and $\hat{\theta}_{BB}$ can exist in the endogenized risk-free rate setting, there is strong numerical evidence that the optimal couple $(\hat{R}_f, \hat{\theta}_{BB})$ is unique. In all instances studied, we confirmed that a single optimal risk-free rate exists. In turn, the optimal share repurchases ratio is also unique, as there exists at most one per value of $\hat{R}_f$. Under the unique optimal found, our results with an exogenous R_f remain true for the endogenized risk-free rate models. The two redistribution methods only differ regarding the firm's price. Buybacks still maintain the firm's price high despite high risk aversions or low expected growth. However, endogenizing the risk-free rate, tying the risk-free asset's returns to the risky ones, reverses the price's behaviour for the dividends setup. Increasing risk aversion also increases the firm's price as it remains competitive: the risk-free asset's rate R_f decreases with increasing γ (see fig. 1.7, *Bottom-Right*).

In the graphs, the risk aversion γ was scaled by K_0^{-1} such that larger companies are not penalized for their size. In the graphs presented, parameters are chosen so that Assumption 1.7 holds. In the buybacks setting, we check numerically that only one value of R_f satisfies $\hat{C}_0 = \theta_{BB}\hat{P}_0$ relation. There also is only one buybacks ratio $\hat{\theta}_{BB}$ maximising the investor's utility .

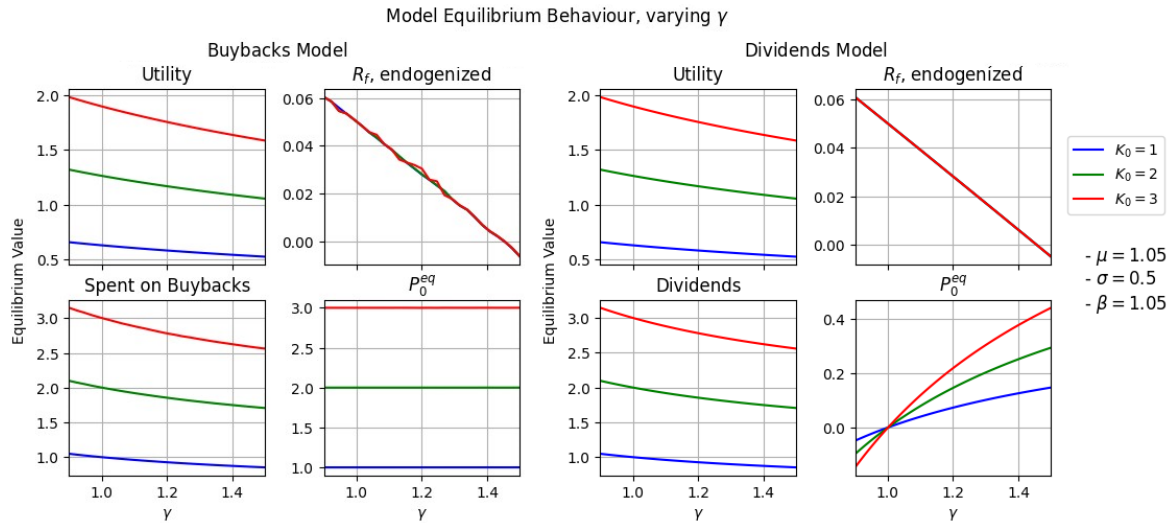


Figure 1.7: Behaviour of the dividends (*Right*) and buybacks (*Left*) models at equilibrium for varying risk aversion ratio γ .

As in the previous non-endogenized cases, cash spent (*Bottom-Left*) and investor utility (*Upper-Left*) are identical in the two settings. $\hat{R}_f$ (*Upper-Right*) is also identical. The firm's valuation (*Bottom-Right*) is typically increased in the buybacks setting compared to the dividends model.

With no starting wealth ($W_{-1} = 0$), for any given set of parameters values (μ, σ, β) , the equilibrium risk-free rate is independent of the size K_0 of the company.

In the dividends model, we also find that for any given set of model parameters, if $\hat{P}_0(K_0) = 0$ for some value of K_0 , then everything else equal, any other value of K_0 would also yield a null price. For increasing γ , the risky asset becomes less attractive

to the investor: the resulting risk free asset R_f also decreases (*see fig. 1.7, Top-Right*). This decrease in the risk-free rate allows for the firm's price to typically increase when γ increases, instead of decreasing sharply for a given exogenous risk-free rate.

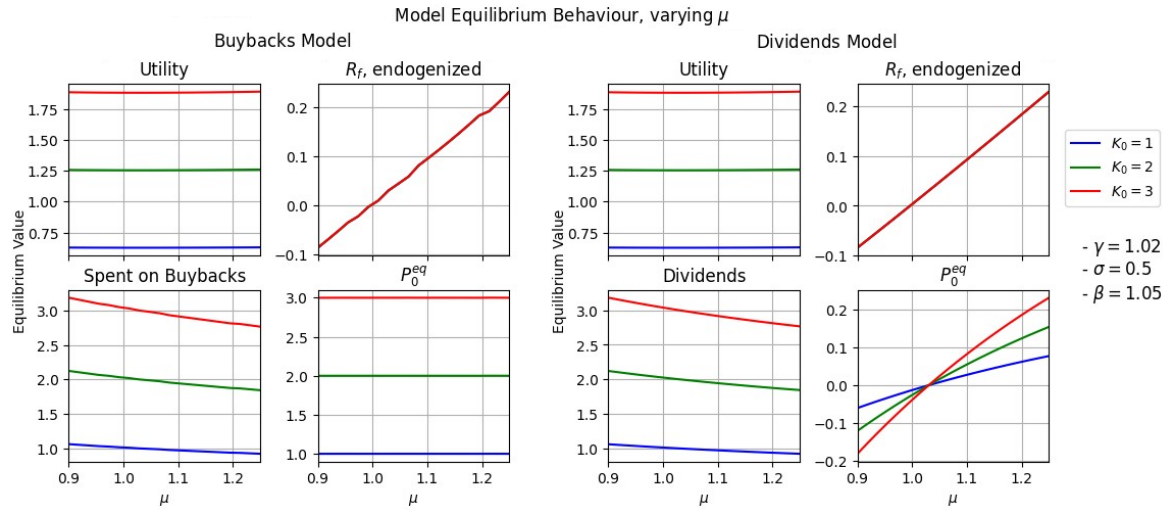


Figure 1.8: Buybacks (*Left*) and dividends (*Right*) equilibrium behaviours for increasing firm growth μ . As expected, the behaviours are analogous to the previous graph. Since the firm's value increases with expected growth μ , the risk-free rate increases with μ here (*Upper-Right*).

The investor's utility and equilibrium company price, as expected, decrease with increasing discount factor β (*see fig. 1.9*). As the value of future earnings continues decreasing as β increases, the price of the company turns negative. Once this reversal happens, the optimal utility of the investor increases with β as they are able to short future earnings of the firm through the negative company price.

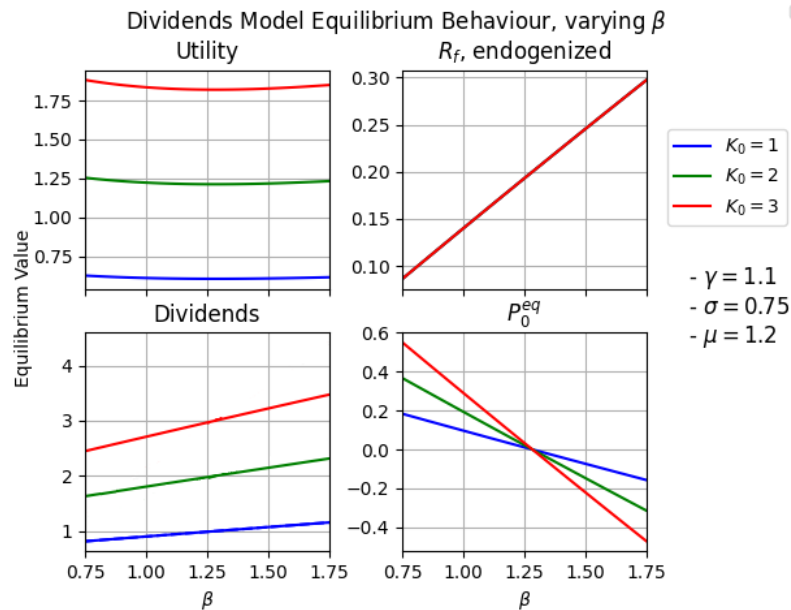


Figure 1.9: Typical dividends' equilibrium behaviour for varying discount factor β .

4 MODEL EXTENSIONS UNDER WHICH DIVIDEND-BUYBACK EQUIVALENCE BREAKS DOWN

4.1 INVESTORS WITH TRADING CONSTRAINT

In this section, we assume some of the investors are facing institutional trading constraints who cannot or are not willing to rebalance their shares holding throughout the investment horizon. Some investors are legally bound to hold positions, such as asset managers who can only trade with their clients approval. For instance, passive investment institutions, such as Exchange Traded Fund (ETF) providers also cannot rebalance their stake, as ETFs' compositions are fixed in advance. Other investors will not reduce their stake in a firm as it carries corporate control over the firm. In our setup, one investor cannot rebalance his ownership in the firm. It replicates the former category of asset managers and passive investment providers. Setting the ownership of this investor not able to rebalance his holdings to 50% of the outstanding shares, our framework thus mirrors an investor prioritizing his control over the firm's voting rights.

Assume there are two representative investors, labeled by $i \in \{1, 2\}$. Investor 1 is required to hold some fixed number of shares at all time points, while investor 2 is unconstrained. We follow the notations introduced in Section 3.1 except we specialize to $\mu = \mu_1 = \mu_2$, i.e. the constrained and unconstrained investors agree on the firm's profitability. For the ease of exposition, we assume investor 1's share holding is fixed at $\theta_{-1}^1 = \theta_0^1 = 1/2$.

Definition 4.1. *Given the firm's distribution strategy (θ_D, θ_{BB}) , in the case with a constrained investor a competitive equilibrium is a pair of the unconstrained investor's trading strategy and the market price of the firm $(\hat{\theta}_0^2, \hat{P}_0)$ such that:*

1. *Investor 2 (the unconstrained investor) maximizes the certainty equivalent of their terminal portfolio value, i.e.*

$$J_2(\hat{\theta}_0^2; \hat{P}_0, \theta_D, \theta_{BB}) = \sup_{\theta_0^2} J_2(\theta_0^2; \hat{P}_0, \theta_D, \theta_{BB}).$$

2. *The market clears, i.e. the total demand for the shares is equal to the number of outstanding shares post-buyback, or equivalently*

$$1 = \frac{1}{2} + \theta_{-1}^2 = \frac{1}{2} + \hat{\theta}_0^2 + \theta_{BB}.$$

Investor i 's equilibrium value function $\hat{J}_i(\theta_D, \theta_{BB})$, $\hat{J}_{i,D}(\theta_D)$ and $\hat{J}_{i,BB}(\theta_{BB})$ are defined as in Definition 3.1.

As in Section 3.1, the firm can determine the optimal dividend or buyback strategy to maximize $\hat{J}_D(\theta_D)$ or $\hat{J}_{BB}(\theta_{BB})$ the sum of the certainty equivalent of the constrained and unconstrained investors in form of (1.9).

Standing Assumption 4.2. *Given model parameters $(\gamma, R_f, K_0, \mu, \sigma)$,*

$$\mu \geq \frac{K_0\gamma\sigma^2}{2} - (1+R_f).$$

Proposition 4.3. *In the setup with a constrained investor, a unique (and consistent) competitive equilibrium exists under any exogenous given pure dividend and buyback strategy.*

1. *Under a dividend strategy $(\theta_D, \theta_{BB} = 0)$, the unique competitive equilibrium is given by*

$$\hat{\theta}_0^2 = \frac{1}{2}, \quad \hat{P}_0 = \frac{K_0(1 - \theta_D)}{1 + R_f} \left(\mu - \frac{\sigma^2\gamma}{2} K_0(1 - \theta_D) \right).$$

Moreover, the optimal dividend level which maximizes the sum of certainty equivalence of the two investors is given by

$$\theta_D^* = 1 - 2 \frac{\mu - (1 + R_f)}{K_0\sigma^2\gamma}$$

such that $\bar{J}_D(\theta_D^) = \sup_{\theta_D} \hat{J}_D(\theta_D)$.*

2. *Under a buyback strategy $(\theta_D = 0, \theta_{BB})$ and Assumption 4.2, the unique consistent competitive equilibrium is given by*

$$\hat{\theta}_0^2 = \frac{1}{2} - \theta_{BB},$$

$$\hat{P}_0 = \frac{K_0}{\theta_{BB}} + \frac{(1 - \theta_{BB})}{\theta_{BB}^2(1 - 2\theta_{BB})\gamma\sigma^2} \left(-\theta_{BB}\mu - (1 - \theta_{BB})(1 + R_f) + \sqrt{\Delta'} \right),$$

where $\Delta'(\theta_{BB})$ is defined as

$$\Delta'(\theta_{BB}) = \left(\theta_{BB}\mu + (1 + R_f)(1 - \theta_{BB}) \right)^2 - 2K_0\gamma\sigma^2\theta_{BB}(1 + R_f)(1 - 2\theta_{BB}).$$

Proof. The dividends optimal behaviour is identical to the one outlined in Section 3.1 with $\mu_1 = \mu_2$. Indeed, the market clearing condition in the dividends setting was,

$$1 = \hat{\theta}_0^1 + \theta_0^2,$$

given $\mu_1 = \mu_2$, we have $\hat{\theta}_0^1 = \hat{\theta}_0^2$. The market clearing condition becomes

$$1 = 2\hat{\theta}_0^2.$$

In this setting, we have $\hat{\theta}_0^1 = 1/2$ and $\hat{\theta}_0^2$ is unchanged from Section 3.1. The market clearing condition thus is

$$1 = 1/2 + \hat{\theta}_0^2 \iff 1 = 2\hat{\theta}_0^2.$$

The market clearing condition is thus identical to Section 3.1's with $\mu_1 = \mu_2$. $\hat{P}_0$ is then also identical, and both ensure $\hat{\theta}_0^1 = \hat{\theta}_0^2 = 1/2$. As such, the expression of $\bar{J}(\hat{\theta}_0^1, \hat{\theta}_0^2, \hat{P}_0, \theta_D)$ is identical in both settings and $\sup_{\theta_D} \bar{J}_D(\theta_D)$ is the same and attained for the same value θ_D^* .

In the pure repurchases setting however, $\hat{\theta}_0^2$ is unchanged from Section 3.1:

$$\hat{\theta}_0^2 = \frac{1 - \theta_{BB}}{\gamma\sigma^2(K_0 - P_0\theta_{BB})} \left(\mu - \frac{P_0(1+R_f)(1 - \theta_{BB})}{K_0 - P_0\theta_{BB}} \right).$$

The market clearing condition however is not the same and becomes:

$$1 = \theta_{BB} + \hat{\theta}_0^1 + \hat{\theta}_0^2,$$

$$1/2 - \theta_{BB} = \frac{1 - \theta_{BB}}{\gamma\sigma^2(K_0 - P_0\theta_{BB})} \left(\mu - \frac{\hat{P}_0(1+R_f)(1 - \theta_{BB})}{K_0 - \hat{P}_0\theta_{BB}} \right),$$

We obtain two potential values for $\hat{P}_0$:

$$\hat{P}_0^+ = \frac{K_0}{\theta_{BB}} + \frac{(1 - \theta_{BB})}{\theta_{BB}^2(1 - 2\theta_{BB})\gamma\sigma^2} (-\theta_{BB}\mu - (1 - \theta_{BB})(1+R_f) + \sqrt{\Delta'}),$$

$$\hat{P}_0^- = \frac{K_0}{\theta_{BB}} + \frac{(1 - \theta_{BB})}{\theta_{BB}^2(1 - 2\theta_{BB})\gamma\sigma^2} (-\theta_{BB}\mu - (1 - \theta_{BB})(1+R_f) - \sqrt{\Delta'}).$$

$$\Delta' = (\theta_{BB}\mu + (1+R_f)(1 - \theta_{BB}))^2 - 2K_0\gamma\sigma^2\theta_{BB}(1+R_f)(1 - 2\theta_{BB}).$$

This definition of consistency in the baseline setting does not apply as $\theta_{BB} \in (0, 1/2)$ instead of $(0, 1)$. In this setting, we call $\hat{P}_0$ "consistent" if $\theta_{BB}\hat{P}_0(\theta_{BB})$ has a finite limits around $1/2$ and tends to 0 as $\theta_{BB} \rightarrow 0$.

$$\lim_{\theta_{BB} \rightarrow 1/2} \theta_{BB}\hat{P}_0^-(\theta_{BB}) = -\infty.$$

Using l'Hôpital's rule, one can check:

$$\begin{aligned}
\lim_{\theta_{BB} \rightarrow 1/2} \theta_{BB} \hat{P}_0^+(\theta_{BB}) &= K_0 + \lim_{\theta_{BB} \rightarrow 1/2} \frac{-2\theta_{BB}\mu}{(1-4\theta_{BB})\gamma\sigma^2} - K_0 \frac{(1-\theta_{BB})(1+R_f)}{\mu\theta_{BB} + (1-\theta_{BB})(1+R_f)}, \\
\lim_{\theta_{BB} \rightarrow 1/2} \theta_{BB} \hat{P}_0^+(\theta_{BB}) &= \frac{\mu}{2\gamma\sigma^2} + \frac{K_0\mu}{\mu + (1+R_f)}, \\
\lim_{\theta_{BB} \rightarrow 0} \theta_{BB} \hat{P}_0^+(\theta_{BB}) &= K_0 + \lim_{\theta_{BB} \rightarrow 0} \frac{\mu\theta_{BB}(\mu - (1+R_f)) - K_0\gamma\sigma^2(1+R_f)(1-4\theta_{BB})}{\gamma\sigma^2(1-\theta_{BB})(1-4\theta_{BB})(1+R_f)}, \\
\lim_{\theta_{BB} \rightarrow 0} \theta_{BB} \hat{P}_0^+(\theta_{BB}) &= 0.
\end{aligned}$$

Thus $(\hat{\theta}_0^2, \hat{P}_0) = (1/2 - \theta_{BB}, \hat{P}_0)$ is the sole consistent market clearing equilibrium.

This equilibrium requires Δ' to exist within the reals. Its minimum is then is

$$\inf_{\theta_{BB} \in \mathbb{R}} \Delta'(\theta_{BB}) = K_0\gamma\sigma^2(1+R_f)^2 \frac{2(1+R_f+\mu) - K_0\gamma\sigma^2}{(1+R_f+\mu)^2 + 4K_0\gamma\sigma^2(1+R_f)}.$$

Thus, Assumption 4.2 guarantees that $\forall \theta_{BB} \in \mathbb{R}, \Delta'(\theta_{BB}) \geq 0$. □

We find that the economic equivalence between repurchases and dividends is broken (*see fig. 1.12*). As the company's relative prospects and risk vary, only the “free” investor can rebalance his exposure. The company acts aiming to maximise the combined utilities of the two investors. As such, when the company carries too much risk, dividends allow for deleveraging equally between the two investors. In the repurchases setting however, deleveraging is not straightforward. While buying back shares from the “free” investor reduces the total risk carried by the risky technology, it also increases the relative share of the “locked” investor's holdings in the company.

As θ_{BB} increases, $\hat{P}_0$ typically grows at first, reaches a maximum value above the face value of the firm's assets and then decreases until it reaches $\hat{P}_0(\theta_{BB} = 1) = K_0$. As θ_{BB} further climbs, $\hat{P}_0$ continues shrinking.

When the risky technology's relative competitiveness is low, i.e. when σ , γ or R_f are high, and μ low, both investors aim at reducing their relative position in the firm. The “free” investor's utility is thus maximised for a high θ_{BB} , nearing 1/2 or even larger, yielding a small or even short position in the firm. The “locked” investor's utility increases as his effective share ownership is reduced. Equivalently, his utility is higher the lower $\left(\frac{K_0 - \theta_{BB}\hat{P}_0}{1 - \theta_{BB}}\right)$ is. $\left(\frac{K_0 - \theta_{BB}\hat{P}_0}{1 - \theta_{BB}}\right)$ is minimised for θ_{BB} resulting in the highest $\hat{P}_0$.

The θ_{BB} maximising $\hat{P}_0$ is typically smaller than the optimal θ_{BB} for the “free” investor. Under typical realistic parameters, there is a unique optimal θ_{BB}^* , which is closer to the buybacks ratio optimal for the “free” investor. However, as the firm becomes less

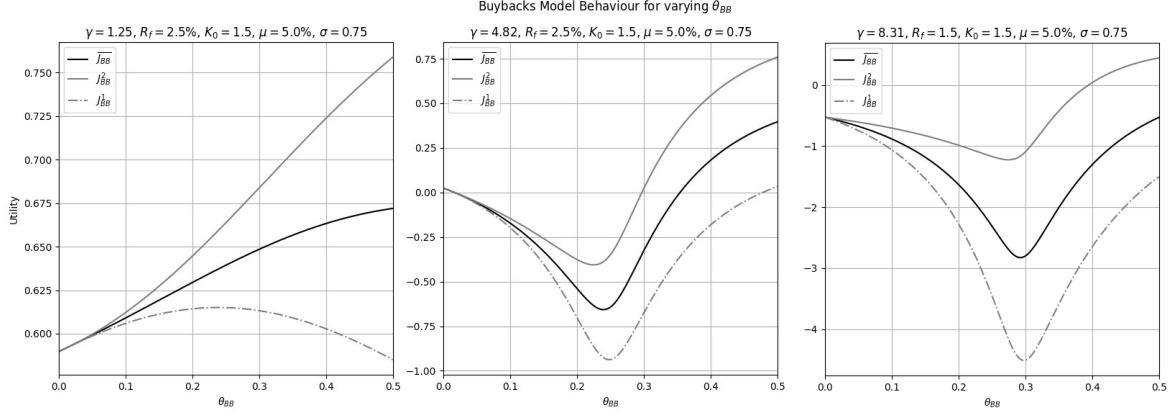


Figure 1.10: Typical behaviour of the buybacks model for varying θ_{BB} .

Left: Firm is competitive - *Center:* Firm is competitive, investors are strongly risk averse - *Right:* Firm is not competitive, investors are strongly risk-averse.

The Right graph exhibits the two distinct model optimum. The low θ_{BB} one favours the “locked” investor, and the high θ_{BB} one favours the “free” investor. This duplication appears for high values of γ and R_f .

All the results presented are under Assumption 4.2: $\gamma < \frac{2(\mu + (1 + R_f))}{K_0 \sigma^2}$.

and less competitive, the individual optimum for the two investors separate. Two aggregate optima then appear (*see fig. 1.10, Right. 1.11*), one with very low θ_{BB}^* , and one with $\theta_{BB} = 1/2$. The former maximises the “locked” investor’s utility, while the latter maximises the “free” investor’s utility.

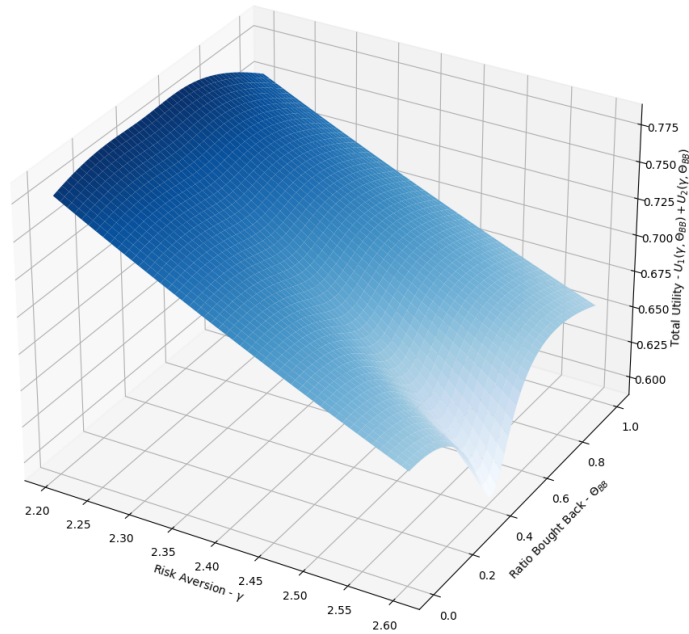


Figure 1.11: Average investor utility for varying repurchases ratio and risk-aversion. As γ increases, the aggregate utility for median values of θ_{BB} lowers and two optimum appear. $\theta_{BB} \approx 0.3$ yields the lowest aggregate value.

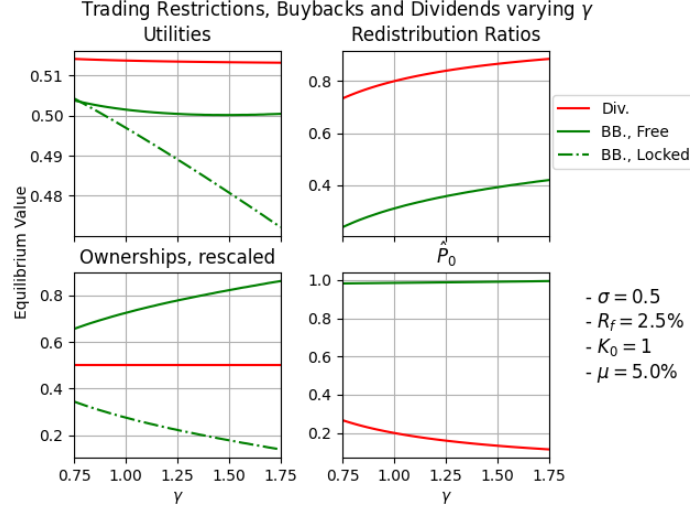


Figure 1.12: Optimal dividends (Red) & buybacks (Green) model behaviours. The economic equivalence is broken (*Top-Left*). Additional demand for shares incurred by repurchases maintains $\hat{P}_0$ high (*Bottom-Right*). As the “locked” investor cannot exit his position, $\theta_D^* > \theta_{BB}^*$ (*Top-Right*).

4.2 MANAGERS WITH STOCK OPTIONS

Managerial stock options have become popular as their introduction within executives’ compensation aligns investors’ interest with the board’s. With no stock options endowment, managers’ entire compensation package is fully comprised of salary and company benefits. The main driver of their utility is thus their tenure length. Since executives’ tenure can only typically end when the company goes bankrupt or when they are fired by investors, they are incentivized to be strongly risk-averse. In practice, their interest can even be opposite to investors’ (Lambrecht and Myers, 2017). Management invests and redistributes the smallest amount of the firm’s assets that will avoid their replacement by investors. With the introduction of stock options within executives’ compensation package, the firm’s management is directly exposed to the company’s equity performance, increasing their risk appetite. The tax treatment of stock options also made them attractive to include in managerial compensation packages: options only carry capital tax, incurred at the options’ execution time. While modern volumes of managerial stock options are high, their implementation within executives’ compensations packages is not new (Crum, 1935; HBR, 1927a). Stock options thus tie the managers’ utility to the investors’ as both actors are then exposed to the company’s future prospects and share price. Empirical studies have outlined plural impacts from the introduction of stock options in managers’ compensation packages: the endowment of stock options decreases dividends (Lambert et al., 1989), motivates buybacks in the hope of raising the stock price (Weisbenner, 2000), and exacerbates the attention given to the EPS ratio (Gao and Kronlund, 2020). As stock options are often assumed to be a major driver of management’s decisions to redistribute firm assets with share repurchases instead of dividends, we implement managerial stock options in the scope of our

model. We also generally assume that executives' stock options are contractually not allowed to be hedged. In other words, stock options endowed to management cannot be used to bet against the firm's prospects.

In this extension, we still assume that the optimal dividend payout ratio θ_D^* or the optimal buyback ratio θ_{BB}^* are determined by the firm in order to maximize the representative investor's welfare, i.e. these quantities are determined by the expressions in Proposition 1.13 or 1.14. However, when the firm is deciding between paying out a dividend ratio of θ_D^* or repurchasing back θ_{BB}^* fraction of the outstanding shares, this decision is guided by the utility value of the managers who own stock options written on the firm underlying shares. That is, since managers are legally required to maximise the investors' utility, the level of redistribution is set in order to maximise investors' expected utility, as in the previous sections. However, when multiple optimal strategies are equivalent to investors, $(0, \theta_D^*)$ and $(\theta_{BB}^*, 0)$ in particular, the firm's executives opt for the decision that maximises their utility.

We assume that at $t = 0$, before repurchases or dividends distributions are announced, the manager is endowed with $\theta_m > 0$ units of "at-the-money" call option which strike price is set at the equilibrium share price level *as if there was no equity distribution*, i.e. the strike price can be expressed as $\kappa := \hat{P}_0(\theta_D = 0, \theta_{BB} = 0)$ where $\hat{P}_0(\theta_D = 0, \theta_{BB} = 0)$ is defined in Definition 1.2 under $(\theta_D = 0, \theta_{BB} = 0)$, i.e.

$$\kappa := K_0 \frac{\mu - K_0 \sigma^2 \gamma}{1 + R_f}.$$

The manager also receives some fixed salary R_m at $t = 0$ and $t = 1$. Under pure optimal redistribution strategies, the payoff to the manager is hence the deterministic quantity

$$J_D^m := (1 + R_f)(R_m + \theta_m(\hat{P}_0(\theta_D^*) - \kappa)^+) + R_m$$

under the choice of dividend strategy, or

$$J_{BB}^m := (1 + R_f)(R_m + \theta_m(\hat{P}_0(\theta_{BB}^*) - \kappa)^+) + R_m$$

under the choice of buyback strategy. In the above, $\hat{P}_0(\theta_D^*)$ and $\hat{P}_0(\theta_{BB}^*)$ are the equilibrium share price under the optimal dividend and buyback strategy respectively as prescribed in the baseline model.⁸ Given that redistribution methods are economically

⁸As the options' payoff function suggests, the options' are European options with expiry at time $t = 0$, after the date at which redistribution is announced to investors. Alternatively, the options' maturity could be set at $t = 1$, the time of liquidation. We showed that in this setting, the optimal buybacks and dividends strategies redistribute the same amounts to investors at time 0. Thus, the total value of the company at time 1 is identical in both settings. However, the share supply is reduced in the repurchases setup. So, given $\theta_{BB}^* \in (0, 1)$, each stock the options allow to buy represents a larger

equivalent, i.e. yield the same expected utility to investors, the managers will adopt the distribution method which yields the higher managerial utility.

Proposition 4.4. *Given model parameters $(\gamma, R_f, K_0, \mu, \sigma)$, if $\theta_{BB}^* \in (0, 1)$ or $\theta_D^* \in (0, 1)$, managerial stock options with θ_{BB}^* implemented are always in-the-money at expiration and*

$$J_{BB}^m > J_D^m.$$

We showed that the condition in Proposition 4.4 that $\theta_D^* \in (0, 1)$ or $\theta_{BB}^* \in (0, 1)$ satisfy:

$$\theta_{BB}^* \in (0, 1) \iff \theta_D^* \in (0, 1) \iff \mu < K_0\gamma\sigma^2 + (1+R_f).$$

Corollary 4.5. *Conversely to Proposition 4.4, if the optimal dividends strategy (resp. buybacks strategy θ_{BB}^*) has the firm raise capital (issue new shares), with $\theta_D^* < 0$ or create an aggregate short position against the company, with $\theta_D^* > 1$, then:*

If $(1+R_f) > K_0\gamma\sigma^2$, only θ_D^ yields in-the-money managerial options and*

$$J_{BB}^m < J_D^m.$$

This case thus has $(1+R_f) > K_0\gamma\sigma^2$ and $\mu < K_0\gamma\sigma^2 + (1+R_f)$

Else, no managerial options are in the money and

$$J_{BB}^m = J_D^m.$$

Proof of Proposition 4.4. Proposition 2.3 shows that in the representative investor setting, the unique optimal pure dividends θ_D^* and pure buybacks θ_{BB}^* strategies have:

$$\begin{aligned}\hat{P}_0(\theta_{BB}^*) &= K_0, \\ \hat{P}_0(\theta_D^*) &= \frac{\mu - (1 + R_f)}{\gamma\sigma^2}.\end{aligned}$$

The options' strike price is set to:

$$\kappa = K_0 \frac{\mu - K_0\sigma^2\gamma}{1 + R_f}.$$

When repurchases are implemented, the managers' stock options are then in-the-money when

$$1 + R_f > \mu - K_0\gamma\sigma^2.$$

share of the company, and is thus more often in-the-money at expiration and is also more valuable when so. So, the manager's expected utility (now stochastic) is strictly higher in the repurchases setting, given $\theta_{BB}^* \in (0, 1)$.

When dividends are distributed, the managerial options are instead in-the-money when

$$0 < \frac{\mu - (1+R_f)}{\gamma\sigma^2} - K_0 \frac{\mu - K_0\sigma^2\gamma}{1+R_f},$$

$$0 < \mu(1+R_f) - (1+R_f)^2 - (\mu K_0\gamma\sigma^2 - (K_0\gamma\sigma^2)^2).$$

In general, we assumed that $\mu \geq (1+R_f) \geq 0$. Assumption 1.7 ensures that $\mu > K_0\gamma\sigma^2$. Thus, the managers' stock options are in-the-money with dividends distribution when

$$1 + R_f > K_0\gamma\sigma^2.$$

Combining the two criteria obtained, under $(\mu > 2K_0\gamma\sigma^2)$, if options in the repurchases setting are in-the-money, the ones in the dividends' setting also are.

Conversely $(\mu < 2K_0\gamma\sigma^2)$, the region of parameters over which $\hat{P}_0(\theta_{BB}^*) > \kappa$ strictly encompasses the one where $\hat{P}_0(\theta_D^*) > \kappa$.

$$J_D^m := (1 + R_f)(R_m + \theta_m(\hat{P}_0(\theta_D^*) - \kappa)^+) + R_m,$$

$$J_{BB}^m := (1 + R_f)(R_m + \theta_m(\hat{P}_0(\theta_{BB}^*) - \kappa)^+) + R_m.$$

Let us consider the setting where both sets of options, whether θ_D^* or θ_{BB}^* is implemented, are in-the-money. Then:

$$J_{BB}^m > J_D^m \iff (\hat{P}_0(\theta_{BB}^*) - \kappa) > (\hat{P}_0(\theta_D^*) - \kappa),$$

$$\iff 0 < K_0 - \frac{\mu - (1+R_f)}{\gamma\sigma^2}$$

$$\iff (1 + R_f) > \mu - K_0\gamma\sigma^2.$$

Which is the condition for the options to be in-the-money when θ_{BB}^* is implemented. Thus, if both of the redistribution methods yield options in the money, $J_{BB}^m > J_D^m$.

□

Optimal Equilibrium Prices and Redistribution Ratios

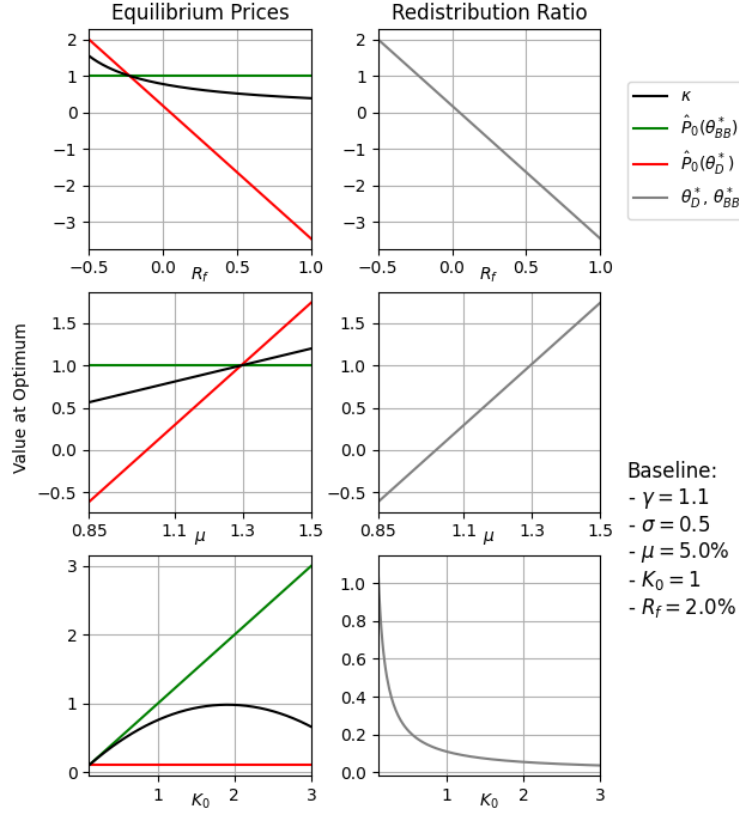


Figure 1.13: Equilibrium price (*Left*) with no redistribution (*Black*), Optimal Dividends (*Red*) and Optimal Repurchases (*Green*) Implemented. On the right is shown the common optimal redistribution ratio θ_D^* , θ_{BB}^* .

Top: varying model parameter is R_f - *Center:* varying model parameter is μ - *Bottom:* varying model parameter is K_0 . Baseline refers to the parameter values used other than the varying feature (*Top:* R_f , *Center:* μ , *Bottom:* K_0).

We observe that the buybacks price is the highest for realistic values of μ, R_f and K_0 . $\hat{P}_0(\theta_{BB}^*) < \kappa$ coincides with $\hat{P}_0(\theta_D^*) < \kappa$, when $\theta_D^* \notin (0, 1)$. That is in the truly remote cases when $R_f < -20\%$, and when $\mu > 30\%$. The case $\hat{P}_0(\theta_D^*)$ and $\hat{P}_0(\theta_B^*)$ smaller than κ does not appear here.

CONCLUSION

We firstly extended a known research framework for firm valuation and optimal dividends allocation to a different form of liquid assets redistribution, share repurchases. While there are multiple ways of executing share buybacks: open tender offers, private negotiations, Dutch auctions or put executions, open market repurchases are the most popular nowadays, representing more than 95% of the volume of buybacks executed by publicly listed firms. This is consequently the method we implemented. Under this setup, we proved that the Modigliani-Miller equivalence for investor utility stands in our representative investor setting. This equivalence is maintained swapping an exogenous risk-free rate for an endogenous one, allowing for multi-period consumption. Focus-

ing on investors' utilities, dividends and share repurchases also are not differentiated when disagreements arise among investors about the company's prospects. However, while investors' equilibria utilities are independent of the redistribution method implemented, even under the simplest and widely used representative investor setting, share buybacks and dividends differ. Indeed, buybacks have a significant effect on share price and earnings per share ratio. Compared to a similar amount spent on dividends, share buybacks increase the EPS ratio, and maintain company price to a level independent of the amount redistributed to investors through repurchases. The setting we used in which these two positive effects of share repurchases arise is very wide. Considering the undeniable importance of stock price and EPS ratio affected, it confirms the theoretical grounding for the long-lasting question in empirical finance: *Why do firms still distribute dividends when buybacks executions are available?* (Bagwell and Shoven, 1989). This question has often been raised regarding the taxation levels in the United States, as mentioned in our introduction. Our first findings point in the same direction, but are even more general, as the tax advantage given to gains from share buybacks, while being true in the US, is not universal (PwC, 2023).

In 1982, under the Ronald Reagan administration the SEC issued rule 10b–18 largely democratising the use of share buybacks by protecting public firms from potential market manipulation claims. As noted, since then, the relevance of dividends distribution has been questioned (Bagwell and Shoven, 1989). Meanwhile, in practice, firms still generally perform dividends. This shows a gap in research's understanding of firms' motivations as to which redistribution method to use. Recently, buybacks have reappeared as a potential topic of interest for new regulation introduction in various major economies such as France and the United States. To predict the impact of such new regulations targeting share repurchases and opine on their future efficacy, it is crucial to understand the effects of buybacks compared to dividends.

Many factors for their difference have been highlighted by empirical studies. Some are marginal and have little impact, such as repurchasing small investors totally in order to diminish the total number of investors, reducing the costs of investors relations or the impact of transaction costs on the firms' choices. Other factors are essential however. We already mentioned the signalling hypothesis and the unequal tax treatment.

Depending on which factors drive a firm's decision and in which context they operate, the effects of redistribution will naturally vary. In practice, since many factors are taken into account for each redistribution decision, their effects blend and it is challenging to determine directly whether they are individually consequential. The signalling hypothesis appears to have opposite consequences depending on the situation. EPS management, while confirmed, does not have systematic outcomes either (Almeida et al., 2016). Thus, to complement empirical examinations, there is need for

more analytical research on share buybacks motivations and consequences. The last sections of this work exhibit the relevance of some factors comparing the distinctions between the two redistribution modes.

In the last section, we displayed settings in which either dividends are beneficial to investors, or share repurchases are.

We introduce an investor whose position in the firm he is banned from reducing. Under this ownership structure, investors' utility is maximised using dividends. Using share buybacks, the firm's behaviour varies substantially for varying investors' the risk-aversion. For increasing risk-aversion levels, different regimes arise. As expected, the firm's level of redistribution initially increase with investors' risk-aversion. Yet, as risk-aversion reaches a high enough point, the firm's optimal redistribution level decreases substantially. This is an instance of the non-linear effect of share repurchases. It can also be seen as a first explanation as to why firms still execute dividends. This extension is motivated by investors who cannot, or will not, sell their stake in a firm. Some are legally bound to hold positions, the most important of such being asset managers whose portfolios can often only be changed with their clients approval, or through an Exchange Traded Fund (ETF) composition modification. Other investors will not reduce their stake in a firm as it carries corporate control over the firm. Our setup replicates an investor considering first maintaining a majority of a firm's voting rights. This effect could be shown empirically too by future work comparing redistribution methods and ownership structures.

Finally, we extend our model to take managerial interests into account. In particular, we consider an executive, whose compensation includes stock options. We then weight the executive's utility alongside investors' within the firm's decision process. Indeed, other studies ([Lambrecht and Myers, 2017](#)) have highlighted the central role that firms' executives play in their management, specifically on their redistribution policies. Executives compensated through salary exclusively have opposite interests to the investors'. Without managerial exposure to their firm's equity value, any non risk-free growth potential has negative value as the firm's bankruptcy risk would increase, even marginally. Thus, the likelihood of managers not being compensated would be enlarged, and their expected utility reduced. As a consequence, stock options are widely used to align the managers' utility with the investors' ([Gao and Kronlund, 2020](#); [Weisbenner, 2000](#)). Implementing managerial stock options in the scope of our model, we find that managers always benefit from implementing share repurchases instead of dividends. This is also a strong suggestion that the inclusion of stock options within executives' compensation packages has been a driver of the adoption of share repurchases by listed firms, which could also be shown empirically.

A APPENDIX

1 REMOVING ASSUMPTION 1.7: NEGATIVE Δ

In this Appendix, we study the behaviour of Δ . Δ 's sign controls for the existence of the company's price in a pure share repurchases scenario. We recall:

$$\Delta(\theta_{BB}) = ((1 + R_f)(1 - \theta_{BB}) + \mu \theta_{BB})^2 - 4K_0\gamma\sigma^2(1 + R_f)(1 - \theta_{BB})\theta_{BB}.$$

BOUNDS FOR $\Delta < 0$

For $\hat{P}_0(\theta_{BB})$ to be real-valued, $\Delta(\theta_{BB}) \geq 0$ is required. This is equivalent to:

$$(1 + R_f)(1 - \theta_{BB}) + \mu \theta_{BB} \geq 2\sqrt{K_0\gamma\sigma^2(1 + R_f)\theta_{BB}(1 - \theta_{BB})}.$$

The next figure A.14 plots the values of θ_{BB} which yield positive Δ in a typical setting.

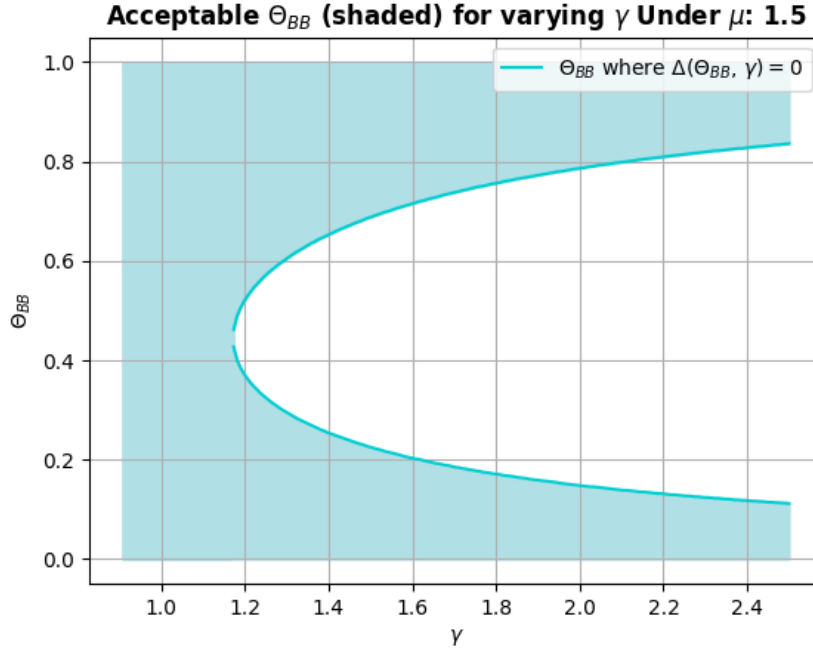


Figure A.14: Acceptable Ratios Bought Back (*Blue*) for Varying γ

Theorem A.1. *We find that in the buybacks setting, depending on model parameters, two regimes arise:*

1) $\mu \geq \gamma K_0 \sigma^2$:

In the regime where risk-adjusted growth rate of the firm is positive, Δ is always positive, and the firm's price always real.

2) $\mu < \gamma K_0 \sigma^2$:

$$\exists \theta_{BB}^1, \theta_{BB}^2 \in (0, 1) \text{ such that } \text{sgn}(\Delta) = \begin{cases} -1, & \text{if } \theta_{BB} \in (\theta_{BB}^2, \theta_{BB}^1), \\ 1, & \text{else.} \end{cases}$$

Succinctly:

$$\min_{\theta_{BB} \in \mathbb{R}} (\text{sgn}(\Delta)) = \begin{cases} 1, & \text{if } \mu \geq K_0 \gamma \sigma^2, \\ -1, & \text{if } \mu < K_0 \gamma \sigma^2 \text{ and } \theta_{BB} \notin (\theta_{BB}^2, \theta_{BB}^1). \end{cases}$$

Where we defined θ_{BB}^1 and θ_{BB}^2 as:

$$\theta_{BB}^1 = \left(1 - \frac{\mu - 2K_0 \gamma \sigma^2 + 2\sqrt{K_0 \gamma \sigma^2 (K_0 \gamma \sigma^2 - \mu)}}{1 + R_f} \right)^{-1},$$

$$\text{and } \theta_{BB}^2 = \left(1 - \frac{\mu - 2K_0 \gamma \sigma^2 - 2\sqrt{K_0 \gamma \sigma^2 (K_0 \gamma \sigma^2 - \mu)}}{1 + R_f} \right)^{-1}.$$

Proof. The determinant Δ is a quadratic polynomial in θ_{BB} .

We trivially have

$$\lim_{|\theta_{BB}| \rightarrow \infty} \Delta = +\infty.$$

Thus there are only two potential behaviours for $\text{sgn}(\Delta(\theta_{BB}))$. Either $\text{sgn}(\Delta(\theta_{BB})) = 1$ for all values of θ_{BB} , or there exists a range $(\theta_{BB}^2, \theta_{BB}^1)$ such that $\Delta(\theta_{BB}) < 0$ is equivalent to $\theta_{BB} \in (\theta_{BB}^2, \theta_{BB}^1)$. Firstly, let us determine the repurchases ratio minimizing Δ . We have:

$$\frac{\partial \Delta(\theta_{BB})}{\partial \theta_{BB}} = 4(2\theta_{BB} - 1)K_0 \gamma \sigma^2(1 + R_f) + 2(\mu - (1 + R_f))((1 - \theta_{BB})(1 + R_f) + \theta_{BB}\mu).$$

Setting $\frac{\partial \Delta(\theta_{BB})}{\partial \theta_{BB}} = 0$, Δ 's single minimum is attained for buybacks ratio θ_{BB}^* , which satisfies:

$$\begin{aligned} 0 &= 4(2\theta_{BB}^* - 1)K_0 \gamma \sigma^2(1 + R_f) + 2(\mu - (1 + R_f))((1 - \theta_{BB}^*)(1 + R_f) + \theta_{BB}^*\mu), \\ &= \theta_{BB}^* \left(8K_0 \gamma \sigma^2(1 + R_f) + 2(\mu - (1 + R_f))\mu - 2(1 + R_f)(\mu - (1 + R_f)) \right) \\ &\quad - 4K_0 \gamma \sigma^2(1 + R_f) + 2(\mu - (1 + R_f))(1 + R_f), \\ 0 &= \theta_{BB}^* \left(4K_0 \gamma \sigma^2 + \frac{1}{1 + R_f}(\mu - (1 + R_f))^2 \right) - 2K_0 \gamma \sigma^2 + \mu - (1 + R_f), \end{aligned}$$

And thus,

$$\theta_{BB}^* = \frac{2K_0\gamma\sigma^2 - (\mu - (1 + R_f))}{4K_0\gamma\sigma^2 + \frac{(\mu - (1 + R_f))^2}{1 + R_f}}$$

Replacing θ_{BB} by this expression in $\Delta(\theta_{BB})$:

$$\begin{aligned}\Delta(\theta_{BB}^*) &= \frac{4K_0\gamma\sigma^2(1 + R_f)^2(\mu - K_0\gamma\sigma^2)}{4K_0\gamma\sigma^2(1 + R_f) + ((1 + R_f) - \mu)^2}, \\ \Delta(\theta_{BB}^*) &= (1 + R_f)(\mu - K_0\gamma\sigma^2) \left(1 + \frac{(\mu - (1 + R_f))^2}{4K_0\gamma\sigma^2(1 + R_f)}\right)^{-1}.\end{aligned}$$

Since $(1 + R_f) \left(1 + \frac{(\mu - (1 + R_f))^2}{4K_0\gamma\sigma^2(1 + R_f)}\right)$ is always non-negative, the sign of $\Delta(\theta_{BB}^*)$ is the sign of $(\mu - K_0\gamma\sigma^2)$.

As θ_{BB}^* minimizes Δ ,

$$\begin{aligned}\text{sgn}(\min_{\theta_{BB} \in \mathbb{R}} \Delta(\theta_{BB})) &= \text{sign}(\Delta(\theta_{BB}^*)). \\ \text{sgn}(\min_{\theta_{BB} \in \mathbb{R}} \Delta) &= \text{sgn}(\mu - K_0\gamma\sigma^2).\end{aligned}$$

We have the claimed $\mu \geq K_0\gamma\sigma^2$ implies $\Delta(\theta_{BB})$ is non-negative for all values of θ_{BB} . When $\mu < K_0\gamma\sigma^2$, we can check that Δ 's roots are as follows,

$$\Delta = 0 \iff \theta_{BB} = \begin{cases} \left(1 - \frac{\mu - 2K_0\gamma\sigma^2 + 2\sqrt{K_0\gamma\sigma^2(K_0\gamma\sigma^2 - \mu)}}{1 + R_f}\right)^{-1}, \\ \left(1 - \frac{\mu - 2K_0\gamma\sigma^2 - 2\sqrt{K_0\gamma\sigma^2(K_0\gamma\sigma^2 - \mu)}}{1 + R_f}\right)^{-1}. \end{cases}$$

Let's name these roots θ_{BB}^1 and θ_{BB}^2 respectively, with $\theta_{BB}^2 \leq \theta_{BB}^1$.

□

Corollary A.2. *Given $\mu < K_0\gamma\sigma^2$, the investor's expected utility J and the company price $\hat{P}_0$ have finite and well-defined limits when the repurchases ratio θ_{BB} tends towards $(\theta_{BB}^1)^\downarrow$ or $(\theta_{BB}^2)^\uparrow$.*

When $\mu < K_0\gamma\sigma^2$, we cannot evaluate $J(\theta_{BB}, 0)$ in the $(\theta_{BB}^2, \theta_{BB}^1)$ range. However, this corollary ensures that outside this interval, $J(\theta_{BB}, 0)$ exists and is finite.

Proof. We verify that as $\theta_{BB} \rightarrow (\theta_{BB}^2)^\uparrow$ and as $\theta_{BB} \rightarrow (\theta_{BB}^1)^\downarrow$, the company's price $\hat{P}_0$

and the investor's expected utility J converge towards finite values:

$$\begin{aligned}
\lim_{\theta_{BB} \rightarrow \theta_{BB}^2} \hat{P}_0(\theta_{BB}) &= -\sqrt{\gamma K_0 \sigma^2 (\gamma K_0 \sigma^2 - \mu)} \frac{2\gamma K_0 \sigma^2 - \mu + (1+R_f) + 2\sqrt{\gamma K_0 \sigma^2 (\gamma K_0 \sigma^2 - \mu)}}{\gamma \sigma^2 (1+R_f)}, \\
&= -\sqrt{\gamma K_0 \sigma^2 (\gamma K_0 \sigma^2 - \mu)} \cdot \frac{1}{\gamma \sigma^2 \theta_{BB}^2}. \\
\lim_{\theta_{BB} \rightarrow \theta_{BB}^1} \hat{P}_0(\theta_{BB}) &= \sqrt{\gamma K_0 \sigma^2 (\gamma K_0 \sigma^2 - \mu)} \frac{2\gamma K_0 \sigma^2 - \mu + (1+R_f) - 2\sqrt{\gamma K_0 \sigma^2 (\gamma K_0 \sigma^2 - \mu)}}{\gamma \sigma^2 (1+R_f)}, \\
&= \sqrt{\gamma K_0 \sigma^2 (\gamma K_0 \sigma^2 - \mu)} \cdot \frac{1}{\gamma \sigma^2 \theta_{BB}^1}. \\
\lim_{\theta_{BB} \rightarrow \theta_{BB}^2} J_{BB}(\theta_{BB}) &= \frac{3K_0\mu}{2} - K_0^2\gamma\sigma^2 + K_0\sqrt{K_0\gamma\sigma^2(K_0\gamma\sigma^2-\mu)} + \\
&\quad \frac{(\mu-(1+R_f))\sqrt{K_0\gamma\sigma^2(K_0\gamma\sigma^2-\mu)}}{\gamma\sigma^2}, \\
\lim_{\theta_{BB} \rightarrow \theta_{BB}^1} J_{BB}(\theta_{BB}) &= \frac{3K_0\mu}{2} - K_0^2\gamma\sigma^2 + K_0\sqrt{K_0\gamma\sigma^2(K_0\gamma\sigma^2-\mu)} \\
&\quad - \frac{(\mu-(1+R_f))\sqrt{K_0\gamma\sigma^2(K_0\gamma\sigma^2-\mu)}}{\gamma\sigma^2}.
\end{aligned}$$

Thus, the limits of $\hat{P}_0$ and J are well-defined and finite when the repurchase ratio tends towards θ_{BB}^1 and θ_{BB}^2 .

□

Corollary A.3. *We can extend the $\theta_{BB} \in (0, 1)$ framework to $\mathbb{R}$. The equilibrium price and the optimal investor's expected utility remain finite and well-defined continuous functions of θ_{BB} .*

That is, we can allow the repurchases decision to create an aggregate short position in the company, or for the issuance of equity to be decided by the company's board.

Proof. In Remark 1.9, we showed that the equilibrium $(\hat{\theta}_0, \hat{P}_0)$ behaves smoothly around $\theta_{BB} = 0$, and be continuously extended at 0 by the original model with no redistribution. Indeed, as $(\hat{\theta}_0, \hat{P}_0)$ is a consistent equilibrium, $\theta_{BB}\hat{P}_0(\theta_{BB})$ approaches 0 as θ_{BB} does. Analogously, around $\theta_{BB} = 1$, $\hat{P}_0(\theta_{BB})$ tends to K_0 and $J_{BB}(\theta_{BB})$ behave smoothly.

When $\theta_{BB} \notin (0, 1)$, we always have $\Delta > 0$, with and without Assumption 1.7. Indeed, we have $\theta_{BB}^1, \theta_{BB}^2 \in (0, 1)$.

Thus, the market clearing price exists for any $\theta_{BB} \in \mathbb{R} \setminus [0, 1]$ and varies continuously over $\mathbb{R} \setminus [0, 1]$.

Thus, we only need to check that as $\theta_{BB} \rightarrow \pm\infty$, J_{BB} , $\hat{P}_0$, and $\theta_{BB} \cdot \hat{P}_0$ behave smoothly and finite.

$$\begin{aligned}\lim_{\theta_{BB} \rightarrow \pm\infty} \hat{P}_0(\theta_{BB}) &= 0. \\ \lim_{\theta_{BB} \rightarrow \pm\infty} \theta_{BB} \hat{P}_0(\theta_{BB}) &= K_0 + \frac{1}{2\gamma\sigma^2} \left(-(\mu - (1 + R_f)) + \sqrt{(\mu - (1 + R_f))^2 + 4(1 + R_f)K_0\gamma\sigma^2} \right). \\ \lim_{\theta_{BB} \rightarrow \pm\infty} J_{BB}(\theta_{BB}) &= \frac{\mu - 1 + R_f}{4\gamma\sigma^2} \cdot \left(\sqrt{(\mu - (1 + R_f))^2 + 4K_0\gamma\sigma^2(1 + R_f)} - (\mu - (1 + R_f)) \right) \\ &\quad - \frac{1}{2}K_0(-1 + R_f).\end{aligned}$$

□

In practice, the price turning complex is not problematic as long as the behaviour of the model at its optimal level of redistribution is not affected by $\Delta < 0$. In particular, to compare the share repurchases and dividends' settings, we would like to ensure that θ_D^* is also an acceptable buybacks ratio:

Lemma A.4. *Given any tuple of model parameters $(\gamma, R_f, K_0, \mu, \sigma)$:*

In the regime $\mu < \gamma K_0 \sigma^2$, when the price $\hat{P}_0(\theta_{BB})$ is complex in for $\theta_{BB} \in (\theta_{BB}^2, \theta_{BB}^1)$, then the optimal level of dividends θ_D^ is such that,*

$$\theta_D^* \notin [\theta_{BB}^2, \theta_{BB}^1].$$

Precisely,

$$\theta_{BB}^1 < \theta_D^*.$$

In other words, for all values of μ , the optimal dividends spending ratio θ_D^* is also an acceptable ratio of outstanding shares to buy back.

Proof. To this end, let's compare the optimal θ_D^* to θ_{BB}^1 :

$$\begin{aligned}\theta_D^* - \theta_{BB}^1 &= 1 - \frac{\mu - (1 + R_f)}{K_0\gamma\sigma^2} - \frac{1 + R_f}{1 + R_f - \mu + 2K_0\gamma\sigma^2 - 2\sqrt{K_0\gamma\sigma^2(K_0\gamma\sigma^2 - \mu)}}, \\ &= \frac{(K_0\gamma\sigma^2 - \mu + (1 + R_f)) \cdot ((1 + R_f) - \mu + 2K_0\gamma\sigma^2 - 2\sqrt{K_0\gamma\sigma^2(K_0\gamma\sigma^2 - \mu)})}{(1 + R_f)K_0\gamma\sigma^2 - \mu K_0\gamma\sigma^2 + 2K_0^2\gamma^2\sigma^4 - 2K_0\gamma\sigma^2\sqrt{K_0\gamma\sigma^2(K_0\gamma\sigma^2 - \mu)}} \\ &\quad - \frac{(1 + R_f)K_0\gamma\sigma^2}{(1 + R_f)K_0\gamma\sigma^2 - \mu K_0\gamma\sigma^2 + 2K_0^2\gamma^2\sigma^4 - 2K_0\gamma\sigma^2\sqrt{K_0\gamma\sigma^2(K_0\gamma\sigma^2 - \mu)}}.\end{aligned}$$

We have that:

$$\begin{aligned} \text{sgn}(\theta_D^* - \theta_{BB}^1) &= \text{sgn}\left((K_0\gamma\sigma^2 - \mu + (1 + R_f)) \cdot ((1 + R_f) - \mu + 2K_0\gamma\sigma^2 \right. \\ &\quad \left. - 2\sqrt{K_0\gamma\sigma^2(K_0\gamma\sigma^2 - \mu)}) - (1 + R_f)K_0\gamma\sigma^2 \right). \end{aligned}$$

Where we used that

$$\begin{aligned} (1 + R_f)K_0\gamma\sigma^2 - \mu K_0\gamma\sigma^2 + 2K_0^2\gamma^2\sigma^4 - 2K_0\gamma\sigma^2\sqrt{K_0\gamma\sigma^2(K_0\gamma\sigma^2 - \mu)} &= \\ K_0\gamma\sigma^2\left((1 + R_f) + (\sqrt{K_0\gamma\sigma^2} - \sqrt{K_0\gamma\sigma^2 - \mu})^2 \right) &> 0. \end{aligned}$$

Replacing $K_0\gamma\sigma^2$ by Ω and $(1 + R_f)$ by R for clarity:

$$\begin{aligned} \text{sgn}(\theta_D^* - \theta_{BB}^1) &= \text{sgn}\left(\Omega(2R - 3\mu + 2\Omega) + (\mu - R)^2 + 2\sqrt{\Omega(\Omega - \mu)}(-\Omega + \mu - R) \right), \\ &= \text{sgn}\left((\mu - R)^2 + 2(\Omega - \sqrt{\Omega(\Omega - \mu)}) \cdot (\Omega + R - \mu) - \mu\Omega \right). \end{aligned}$$

$$\text{Defining } f(\Omega, \mu, R) = (\mu - R)^2 + 2(\Omega - \sqrt{\Omega(\Omega - \mu)}) \cdot (\Omega + R - \mu) - \mu\Omega.$$

$$\text{We have } \partial_R f(\Omega, \mu, R) = -2(\mu - R) + 2(\Omega - \sqrt{\Omega(\Omega - \mu)}),$$

$$\text{So } \partial_R f(\Omega, \mu, R) = 0 \iff R = \mu - \Omega + \sqrt{\Omega(\Omega - \mu)}.$$

Notice that $\lim_{R \rightarrow \pm\infty} f(\Omega, \mu, R) = \infty$, and that f is a quadratic polynomial in R .

Replacing R by $\mu - \Omega + \sqrt{\Omega(\Omega - \mu)}$ in f 's expression,

$$\begin{aligned} f(\Omega, \mu, \mu - \Omega + \sqrt{\Omega(\Omega - \mu)}) &= (\Omega - \sqrt{\Omega(\Omega - \mu)})^2 + 2(\Omega - \sqrt{\Omega(\Omega - \mu)}) \cdot \sqrt{\Omega(\Omega - \mu)} - \mu\Omega, \\ &= (\Omega - \sqrt{\Omega(\Omega - \mu)}) \cdot (\Omega + \sqrt{\Omega(\Omega - \mu)}) - \mu\Omega, \\ &= \Omega^2 - \Omega(\Omega - \mu) - \mu\Omega = 0. \end{aligned}$$

Consequently, $\forall \Omega, \mu \in \mathbb{R}, \text{ s.t. } \Omega > \mu, f(\Omega, \mu, R) > 0$. Equivalently,

$$\theta_D^* > \theta_{BB}^1.$$

□

EQUIVALENCE BETWEEN SHARE BUYBACKS AND DIVIDENDS WITHOUT ASSUMPTION 1.7

In this appendix, we extend the proof of equivalence between pure share repurchases and dividends strategies in the representative investor setting.

Firstly, we had V_D the maximal investor utility implementing dividends only, and V_{BB} the optimal one executing repurchases only. In this Appendix, the optimisation is run over buybacks strategies that yield a real price: the admissible set of values for θ_{BB} is $\mathbb{R} \setminus (\theta_{BB}^2, \theta_{BB}^1)$. That is,

$$V_{BB} = \sup_{\theta_{BB} \notin (\theta_{BB}^2, \theta_{BB}^1)} \hat{J}_{BB}(\theta_{BB}).$$

The proof outlined in Section 2 that $V_D \geq V_{BB}$ does not rely on Assumption 1.7. Thus, we only need to demonstrate $V_D \leq V_{BB}$ to extend the proof of the economic equivalence of share repurchases and dividends distribution.

We recall the definition of $f : \mathbb{R} \mapsto \mathbb{R}$:

$$\forall \theta_{BB} \in \mathbb{R}^*, f(\theta_{BB}) := \theta_{BB} \hat{P}_0(\theta_{BB}),$$

extended at 0 to $f(0) = 0$ by continuity, since we have that $(\hat{\theta}_0, \hat{P}_0)$ is a consistent equilibrium over $\mathbb{R} \setminus (\theta_{BB}^2, \theta_{BB}^1)$, and $\theta_{BB}^2 > 0$.

Lemma A.5. *Assume $1 + R_f > 0$.*

1. *If $\mu \geq \sigma^2 \gamma K_0$, under Assumption 1.7, the equation $f(\theta_{BB}) = K_0 \theta_D^*$ defined on $\theta_{BB} \in \mathbb{R}$ admits a unique solution θ_{BB}^* .*
2. *Suppose instead $\mu < \sigma^2 \gamma K_0$.*

If $K_0 \gamma \sigma^2 (\mu - 2(1 + R_f)) \leq (\mu - 1 - R_f)^2$ holds, the equation $f(\theta_{BB}) = K_0 \theta_D^$ defined on $\theta_{BB} \in \mathbb{R} \setminus (\theta_{BB}^2, \theta_{BB}^1)$ admits a unique solution θ_{BB}^* with $\theta_{BB}^* \geq \theta_{BB}^1$.*

Otherwise, there is no $\theta_{BB} \in \mathbb{R} \setminus (\theta_{BB}^2, \theta_{BB}^1)$ such that $f(\theta_{BB}) = K_0 \theta_D^$.*

In either case, when the unique solution exists, it must be given by $\theta_{BB}^ = \theta_D^*$.*

Case 1 is equivalent to Proposition 2.2. We focus the proof on Case 2.

Proof. Since $\mu < K_0 \gamma \sigma^2$, f is locally continuous on $(-\infty, \theta_{BB}^2]$ and $[\theta_{BB}^1, \infty)$ respectively. We show that, given $K_0 \gamma \sigma^2 (\mu - 2(1 + R_f)) \leq (\mu - 1 - R_f)^2$, there still exists a unique solution to $\theta_D^* K_0 = f(\theta_{BB})$.

From Proposition 2.2, we still have:

$$\begin{aligned} \lim_{\theta_{BB} \rightarrow +\infty} f(\theta_{BB}) &> K_0 \theta_D^*, \\ \lim_{\theta_{BB} \rightarrow -\infty} f(\theta_{BB}) &< K_0 \theta_D^*. \end{aligned}$$

And we recall from Appendix 1,

$$\begin{aligned}\lim_{\theta_{BB} \rightarrow (\theta_{BB}^2)^\uparrow} \theta_{BB}^2 \hat{P}_0(\theta_{BB}) &= -\frac{\sqrt{\gamma K_0 \sigma^2 (\gamma K_0 \sigma^2 - \mu)}}{\gamma \sigma^2}, \\ \lim_{\theta_{BB} \rightarrow (\theta_{BB}^1)^\downarrow} \theta_{BB}^1 \hat{P}_0(\theta_{BB}) &= \frac{\sqrt{\gamma K_0 \sigma^2 (\gamma K_0 \sigma^2 - \mu)}}{\gamma \sigma^2}.\end{aligned}$$

We also recall that:

$$f'(\theta_{BB}) = \frac{1 + R_f - \Delta(\theta_{BB})^{-\frac{1}{2}}(\Delta(\theta_{BB}) - \frac{\theta_{BB}}{2} \Delta'(\theta_{BB}))}{2\sigma^2 \gamma \theta_{BB}^2}.$$

Defining $g : \mathbb{R} \setminus (\theta_{BB}^2, \theta_{BB}^1) \mapsto \mathbb{R}$ as

$$g(\theta_{BB}) := \Delta(\theta_{BB})^{\frac{1}{2}}(1 + R_f) - (\Delta(\theta_{BB}) - \frac{\theta_{BB}}{2} \Delta'(\theta_{BB})),$$

we have

$$\text{sgn}(f'(\theta_{BB})) = \text{sgn}(g(\theta_{BB})).$$

Which gives,

$$g(\theta_{BB}) = 0 \iff (\mu - K_0 \gamma \sigma^2) \theta_{BB} = 0 \iff \theta_{BB} = 0.$$

That is, f' can only change signs at $\theta_{BB} = 0$. Hence, locally, f must be strictly monotonic on $(-\infty, 0)$, $(0, \theta_{BB}^2)$ and (θ_{BB}^1, ∞) respectively. We also know that $f(0) = 0$ and $\lim_{\theta_{BB} \rightarrow \theta_{BB}^2} f(\theta_{BB}) < 0$, as we recalled above. When $\mu < K_0 \gamma \sigma^2$, we have $\theta_D^* > 0$. By the strict monotonicity and continuity of f on $(0, \theta_{BB}^2)$, $(f(\theta_{BB}) = K_0 \theta_D^*)$ does not admit any solution on $(0, \theta_{BB}^2)$.

Likewise, the strict monotonicity of f and the facts that $\lim_{\theta_{BB} \rightarrow -\infty} f(\theta_{BB}) < K_0 \theta_D^*$ and $f(0) = 0$ rule out the existence of solution to the equation $(f(\theta_{BB}) = K_0 \theta_D^*)$ on $(-\infty, 0)$.

$$g(\theta_{BB}^1) = \frac{\theta_{BB}^1}{2} \Delta'(\theta_{BB}^1) > 0,$$

from which we conclude f is strictly increasing on (θ_{BB}^1, ∞) . With $\lim_{\theta_{BB} \rightarrow \infty} f(\theta_{BB}) > K_0 \theta_D^*$, the necessary and sufficient condition for $f(\theta_{BB}) = K_0 \theta_D^*$ to admit a unique

solution is

$$\begin{aligned} \lim_{\theta_{BB} \rightarrow (\theta_{BB}^1)^\downarrow} f(\theta_{BB}) \leq K_0 \theta_D^* &\iff \frac{\sqrt{\gamma K_0 \sigma^2 (\gamma K_0 \sigma^2 - \mu)}}{\gamma \sigma^2} \leq K_0 - \frac{\mu - 1 - R_f}{\sigma^2 \gamma} \\ &\iff K_0 \gamma \sigma^2 (\mu - 2(1 + R_f)) \leq (\mu - 1 - R_f)^2. \end{aligned}$$

If instead $\lim_{\theta_{BB} \rightarrow (\theta_{BB}^1)^\downarrow} f(\theta_{BB}) > K_0 \theta_D^*$, then $f(\theta_{BB}) > K_0 \theta_D^*$ for all $\theta_{BB} \geq \theta_{BB}^1$ and the equation $f(\theta_{BB}) = K_0 \theta_D^*$ cannot admit any solution on $[\theta_{BB}^1, \infty)$.

We now show that $\theta_{BB}^* = \left(1 - \frac{\mu - (1 + R_f)}{K_0 \gamma \sigma^2}\right)$ is the admissible solution to $(f(\theta_{BB}) = K_0 \theta_D^*)$ given $K_0 \gamma \sigma^2 (\mu - 2(1 + R_f)) \leq (\mu - 1 - R_f)^2$.

Since $\theta_D^* > \theta_{BB}^1$, we have that $\theta_{BB}^* = \theta_D^* = 1 - \frac{\mu - (1 + R_f)}{K_0 \gamma \sigma^2}$ is an admissible strategy with $\Delta(\theta_{BB}^*) > 0$. Actually, from Proposition 2.2, we already have that $\Delta(\theta_{BB}^*)$ can be expressed as:

$$\Delta(\theta_{BB}^*) = \left(\frac{(\mu - (1 + R_f))^2}{\gamma K_0 \sigma^2} - (\mu - 2(1 + R_f)) \right)^2.$$

Thus, since we assumed that $(K_0 \gamma \sigma^2 (\mu - 2(1 + R_f)) \leq (\mu - 1 - R_f)^2)$ holds,

$$\sqrt{\Delta(\theta_{BB}^*)} = \frac{(\mu - (1 + R_f))^2}{\gamma K_0 \sigma^2} - (\mu - 2(1 + R_f)).$$

Plugging this expression back in the expression for the company's price, $\hat{P}_0(\theta_{BB}^*)$, as in Proposition 2.2:

$$\hat{P}_0(\theta_{BB}^*) = K_0.$$

With the chosen θ_{BB}^* , $\hat{P}_0(\theta_{BB}^*)$ precisely equals K_0 . Thus, with $\theta_{BB}^* = \theta_D^*$, $\theta_{BB}^* \hat{P}_0(\theta_{BB}^*) = \theta_D^* K_0$.

□

Combining Lemma A.5 with the above, we get:

Corollary A.6. *If $K_0 \gamma \sigma^2 (\mu - 2(1 + R_f)) \leq (\mu - 1 - R_f)^2$, then*

$$\hat{J}(0, \theta_D^*) = J(\theta_{BB}^* = \theta_D^*, 0),$$

And thus

$$V_D = V_{BB}.$$

The proof is identical to Proposition 2.2's proof. Indeed, we already showed that any

θ_{BB}^u such that

$$f(\theta_{BB}^u) = \theta_D^* K_0,$$

also automatically satisfies, $\hat{J}(0, \theta_D^*) = J(\theta_{BB}^u, 0)$.

Proof. From the previous Lemma, let us finally show that $J(\theta_{BB}^*, 0) = J(0, \theta_D^*)$.

$$\begin{aligned} J(\theta_{BB}^*, 0) &= (1 + R_f)(W_{-1} + (\theta_{-1} - \hat{\theta}_0)\hat{P}_0(\theta_{BB}^*)) \\ &\quad + \mu \frac{K_0 - \hat{P}_0(\theta_{BB}^*)\theta_{BB}^*}{1 - \theta_{BB}^*} \hat{\theta}_0 - \frac{\sigma^2 \gamma}{2} \frac{(K_0 - \hat{P}_0(\theta_{BB}^*)\theta_{BB}^*)^2}{(1 - \theta_{BB}^*)^2} \hat{\theta}_0^2. \end{aligned}$$

Using that $\theta_{-1} = 1$ and that, given θ_{BB}^* , at equilibrium $\hat{\theta}_0 = (1 - \theta_{BB}^*)$:

$$\begin{aligned} J(\theta_{BB}^*, 0) &= (1 + R_f)(W_{-1} + \theta_{BB}^* \hat{P}_0(\theta_{BB}^*)) + \mu (K_0 - \hat{P}_0(\theta_{BB}^*)\theta_{BB}^*) \\ &\quad - \frac{\sigma^2 \gamma}{2} (K_0 - \hat{P}_0(\theta_{BB}^*)\theta_{BB}^*)^2, \\ &= (1 + R_f)(W_{-1} + K_0 \theta_D^*) + \mu K_0 (1 - \theta_D^*) - \frac{\sigma^2 \gamma}{2} (K_0 (1 - \theta_D^*))^2, \\ &= J(0, \theta_D^*). \end{aligned}$$

And by generality of θ_{BB}^* , $V_{BB} \geq V_D$.

□

Remark A.7. *We have that*

$$K_0 \gamma \sigma^2 (\mu - 2(1 + R_f)) \leq (\mu - 1 - R_f)^2$$

is satisfied automatically if $\mu \leq 2(1 + R_f)$ or $\mu \geq K_0 \gamma \sigma^2$.

Else, that is when

$$2(1 + R_f) < \mu < K_0 \gamma \sigma^2,$$

the condition becomes equivalent to

$$K_0 \gamma \sigma^2 \in \left(\mu, \mu + \frac{(1 + R_f)^2}{\mu - 2(1 + R_f)} \right].$$

In summary, the equation $(f(\theta_{BB}) = \theta_D^ K_0)$ has a unique solution unless*

$$(2(1 + R_f) < \mu) \quad \text{and} \quad \left(\mu + \frac{(1 + R_f)^2}{\mu - 2(1 + R_f)} < K_0 \gamma \sigma^2 \right), \quad (12)$$

in which case it does not have a solution.

Then under (12),

$$V_D > V_{BB}.$$

In economic terms, if the company's risky technology is expected to be highly profitable relative to the bank account growth rate, then for the dividends and buybacks to be economically equivalent, there is a ceiling over the values the company's per-dollar risk charge can take.

Proof. Under (12), let us assume there exists $\theta_{BB}^u \in \mathbb{R}$ such that

$$J(\theta_{BB}^u, 0) = J(0, \theta_D^*).$$

Let $\theta_D^u \in \mathbb{R}$ be a pure dividends strategy defined as:

$$\theta_D^u = \frac{\theta_{BB}^u \hat{P}_0(\theta_{BB}^u)}{K_0}.$$

That is, implementing θ_D^u has the firm redistribute the same amount to investors as θ_{BB}^u .

We already showed that this implies

$$J(0, \theta_D^u) = J(\theta_{BB}^u, 0).$$

Thus,

$$J(0, \theta_D^u) = J(0, \theta_D^*).$$

By optimality of θ_D^* , we get that θ_D^u is an optimal dividends strategy. As we know θ_D^* is the unique pure dividends optimal strategy, we must have

$$\theta_D^* = \theta_D^u.$$

However, under (12), we have

$$\theta_D^* K_0 \neq \theta_{BB}^u \hat{P}_0(\theta_{BB}^u).$$

So

$$\theta_D^* \neq \theta_D^u.$$

Thus, the claimed θ_{BB}^u cannot exist and we have the claimed

$$V_D > V_{BB}.$$

□

$\Delta < 0$ WITH AN ENDOGENIZED RISK-FREE RATE

In the previous sections of this appendix, we studied Δ 's positivity as a function of θ_{BB} . Within the context of an endogenized risk-free rate, we proceed with an analogous analysis of Δ 's positivity as a function of R_f .

With an endogenized risk-free rate, we cannot obtain a closed form formula for $\hat{R}_f$ and in turn cannot get one for θ_{BB}^* . Given θ_{BB} exogenously, we can numerically find $\hat{R}_f$ however, as the value of the risk-free rate that ensures no cash is carried between periods by the investor. Optimising over values of θ_{BB} , we can find its optimal repurchases ratio numerically.

However, to be able to carry such an analysis in general, for each pair (R_f, θ_{BB}) , we need the price of the company $\hat{P}_0$ to be real. That is, we need Δ to be non-negative. Assumption 1.7 guarantees $\Delta \geq 0$. In the absence of such assumption, in the non-endogenized risk-free rate setup, we had a conditioning on θ_{BB} for $\Delta \geq 0$ to hold.

Here however, we need the condition to be placed on R_f instead, as a variable over which variations determine the model's equilibrium.

Given a set of model parameters and a repurchase ratio θ_{BB} , let us determine whether Δ turns negative for values of R_f .

$\Delta(R_f)$'s minimum is attained when $\frac{\partial \Delta(R_f)}{\partial R_f} = 0$.

$$\text{That is when } (1 + R_f) = \frac{\theta_{BB}}{1 - \theta_{BB}}(2K_0\gamma\sigma^2 - \mu).$$

Replacing $(1 + R_f)$ by this value in Δ yields:

$$\begin{aligned} \min_{R_f} \Delta(R_f) &= \left(\mu \theta_{BB} + \theta_{BB}(2K_0\gamma\sigma^2 - \mu) \right)^2 - 4\theta_{BB}^2 K_0\gamma\sigma^2 (2K_0\gamma\sigma^2 - \mu), \\ \min_{R_f} \Delta(R_f) &= 4\theta_{BB}^2 K_0\gamma\sigma^2 (\mu - K_0\gamma\sigma^2). \end{aligned}$$

And, as for the minimization of Δ for varying repurchase ratio θ_{BB} ,

$$\text{sgn}(\min_{R_f} \Delta(R_f)) = \text{sgn}(\mu - K_0\gamma\sigma^2).$$

In the case where Δ turns negative for certain values of R_f , we have that:

$$\Delta = 0 \iff R_f \in \begin{cases} \frac{\theta_{BB}}{1 - \theta_{BB}} \left(2K_0\gamma\sigma^2 - \mu + 2\sqrt{K_0\gamma\sigma^2(K_0\gamma\sigma^2 - \mu)} \right) - 1, \\ \frac{\theta_{BB}}{1 - \theta_{BB}} \left(2K_0\gamma\sigma^2 - \mu - 2\sqrt{K_0\gamma\sigma^2(K_0\gamma\sigma^2 - \mu)} \right) - 1. \end{cases} \quad (13)$$

Let these two roots be R_f^1 and R_f^2 respectively. We directly have $R_f^1 \geq R_f^2$.

$$\Delta < 0 \iff R_f \in (R_f^2, R_f^1) \text{ and } (\mu < K_0 \gamma \sigma^2).$$

2 IRRELEVANCE OF THE ORDER OF ENDOGENIZATION IN THE CASE WITH ENDOGENIZED RISK-FREE RATE

In the main part of this manuscript, endogenizing the risk-free rate, the final optimisation was run over θ_{BB} / θ_D . That is, for varying redistribution levels, the firm first takes into account the impact its decision would have on the risk-free rate at the optimal consumption level. Only then did the firm decide on which redistribution level to use, fixing the risk-free rate at the same time.

In this appendix, we present an alternative endogenization methodology, assuming that a large number of indistinguishable firms, each having a negligible impact on the risk-free rate, first individually decide on a level of dividends / share repurchases to operate.

Once this first optimisation is done, a central entity fixes the risk-free rate so that the rate chosen by the individual firms leaves the total cash between period $t = 0$ and $t = 1$ in zero net supply.

EQUIVALENCE BETWEEN RISK-FREE RATE ENDOGENIZATION METHODS

NOTATIONS OF EXPECTED UTILITY

We focus on the dividend model. Fix the initial cash and shares endowment as $W_{-1} = 0$ and $\theta_{-1} = 1$. The representative investor's expected utility is a function of $(\theta_0, C_0, \theta_D, P_0, R_f)$ defined via

$$J(\theta_0, C_0, \theta_D, P_0, R_f) := E[U(C_0) + e^{-\beta} U((1 + R_f)[\theta_D \cdot K_0 - C_0 + (1 - \theta_0)P_0] + \theta_0 K_0(1 - \theta_D) \Pi)].$$

In the above, θ_0 is the investor's shares holding at time 0 post-distribution, C_0 is the investor's consumption level at time 0, $\theta_D \cdot K_0$ is the dividend payout, R_f is the risk-free rate, and $\Pi \sim N(\mu, \sigma^2)$ is the stochastic growth rate of the firm.

Note that along $\theta_0 = 1$, J no longer depends on P_0 and we define

$$H(C_0, \theta_D, R_f) := J(1, C_0, \theta_D, P_0, R_f) \\ = E[U(C_0) + e^{-\beta} U((1 + R_f)[\theta_D \cdot K_0 - C_0] + K_0(1 - \theta_D) \Pi_1)].$$

Note also that along $C_0 = \theta_D \cdot K_0$, H no longer depends on R_f and we define

$$\begin{aligned} h(\theta_D) &:= H(\theta_D, \theta_D, R_f) \\ &= E \left[U(\theta_D \cdot K_0) + e^{-\beta} U(K_0(1 - \theta_D) \Pi_1) \right]. \end{aligned} \quad (14)$$

INVESTOR'S OPTIMAL INVESTMENT-CONSUMPTION STRATEGY

Given θ_D , P_0 , and R_f , the investor seeks the optimal investment and consumption strategy θ_0 and C_0 to maximize their utility. Formally, the optimal strategy is denoted by the pair $(\theta_0^*(\theta_D, P_0, R_f), C_0^*(\theta_D, P_0, R_f))$ such that

$$J(\theta_0^*(\theta_D, P_0, R_f), C_0^*(\theta_D, P_0, R_f), \theta_D, P_0, R_f) \geq J(\theta_0, C_0, \theta_D, P_0, R_f) \quad (15)$$

for all $\theta_0, C_0, \theta_D, P_0, R_f$.

Under the choice of equilibrium share price $\hat{P}_0 = \hat{P}_0(\theta_D, R_f)$ such that $\theta_0^*(\theta_D, \hat{P}_0, R_f) = 1$, (15) becomes

$$J(1, C_0^*(\theta_D, \hat{P}_0, R_f), \theta_D, \hat{P}_0, R_f) \geq J(\theta_0, C_0, \theta_D, \hat{P}_0, R_f),$$

which can be restated as

$$H(C_0^*(\theta_D, \hat{P}_0, R_f), \theta_D, R_f) \geq J(\theta_0, C_0, \theta_D, \hat{P}_0, R_f). \quad (16)$$

Inequality (16) holds for all θ_0 . Putting $\theta_0 = 1$ together with an abuse of notation that $C_0^*(\theta_D, R_f) := C_0^*(\theta_D, \hat{P}_0(\theta_D, R_f), R_f)$, (16) leads to

$$H(C_0^*(\theta_D, R_f), \theta_D, R_f) \geq H(C_0, \theta_D, R_f). \quad (17)$$

In other words, the optimal consumption under equilibrium share price (but with exogenous R_f and θ_D) is the maximizer of $H(C_0, \theta_D, R_f)$.

“FULL” EQUILIBRIUM

Under a given θ_D , an equilibrium interest rate is first determined under the money market clearing condition $C_0^* = \theta_D \cdot K_0$. We define $\hat{R}_f = \hat{R}_f(\theta_D)$ as the solution to the equation

$$C_0^*(\theta_D, \hat{R}_f) = \theta_D \cdot K_0. \quad (18)$$

Along the equilibrium interest rate $\hat{R}_f(\theta_D)$, managers determine the optimal dividend to maximize investors' utility, i.e. the optimal dividend θ_D^* satisfies

$$\begin{aligned} H^* &:= H(C_0^*(\theta_D^*, \hat{R}_f(\theta_D^*)), \theta_D^*, \hat{R}_f(\theta_D^*)) \\ &\geq H(C_0^*(\theta_D, \hat{R}_f(\theta_D)), \theta_D, \hat{R}_f(\theta_D)), \text{ for all } \theta_D. \end{aligned} \quad (19)$$

But with (18),

$$\begin{aligned} H(C_0^*(\theta_D, \hat{R}_f(\theta_D)), \theta_D, \hat{R}_f(\theta_D)) &= H(\theta_D \cdot K_0, \theta_D, \hat{R}_f(\theta_D)) \\ &= E [U(\theta_D \cdot K_0) + e^{-\beta} U(K_0(1 - \theta_D) \Pi_1)] \\ &=: h(\theta_D). \end{aligned}$$

Hence (19) can be restated as

$$H^* = h(\theta_D^*) \geq h(\theta_D), \text{ for all } \theta_D. \quad (20)$$

“PARTIAL” EQUILIBRIUM

As another formulation, managers can first optimize dividend by assuming that R_f is fixed. They seek $\tilde{\theta}_D = \tilde{\theta}_D(R_f)$ which maximizes $H(C_0^*(\theta_D, R_f), \theta_D, R_f)$. Formally, for any R_f , $\tilde{\theta}_D(R_f)$ is defined such that

$$H(C_0^*(\tilde{\theta}_D(R_f), R_f), \tilde{\theta}_D(R_f), R_f) \geq H(C_0^*(\theta_D, R_f), \theta_D, R_f) \text{ for all } \theta_D. \quad (21)$$

Finally, the risk-free rate is endogenized via the money market clearing condition $C_0^*(\tilde{\theta}_D, R_f) = \tilde{\theta}_D \cdot K_0$. The “partial” equilibrium interest rate in this case is defined as $\tilde{R}_f^*$ such that

$$C_0^*(\tilde{\theta}_D(\tilde{R}_f^*), \tilde{R}_f^*) = \tilde{\theta}_D(\tilde{R}_f^*) \cdot K_0. \quad (22)$$

The equilibrium dividend payout is $\tilde{\theta}_D^* := \tilde{\theta}_D(\tilde{R}_f^*)$, and the corresponding utility level is $\tilde{H}^* := H(C_0^*(\tilde{\theta}_D^*, \tilde{R}_f^*), \tilde{\theta}_D^*, \tilde{R}_f^*)$.

EQUIVALENCE OF THE TWO VERSIONS OF EQUILIBRIUM

Theorem A.8. *The two versions of the equilibrium are equivalent. Specifically, $\theta_D^* = \tilde{\theta}_D^*$, $\hat{R}_f^* = \tilde{R}_f^*$ and $H^* = \tilde{H}^*$.*

Proof. We first show that $\theta_D^* = \tilde{\theta}_D^*$. By definition, θ_D^* is the maximizer of $h(\theta_D)$, it is sufficient to show that $h(\tilde{\theta}_D^*) \geq h(\theta_D)$ for all θ_D . Then the result will follow from uniqueness of the maximizer, proved in Proposition 3.8.

Now,

$$\begin{aligned}
h(\tilde{\theta}_D^*) &= H(\tilde{\theta}_D^* \cdot K_0, \tilde{\theta}_D^*, x) && \text{(def. of } h, x \text{ can be arbitrary here)} \\
&= H(\tilde{\theta}_D(\tilde{R}_f^*) \cdot K_0, \tilde{\theta}_D(\tilde{R}_f^*), \tilde{R}_f^*) && \text{(def. of } \tilde{\theta}_D^* \text{ and choose } x = \tilde{R}_f^*) \\
&= H(C_0^*(\tilde{\theta}_D(\tilde{R}_f^*), \tilde{R}_f^*), \tilde{\theta}_D(\tilde{R}_f^*), \tilde{R}_f^*) && \text{(clearing condition (21))} \\
&\geq H(C_0^*(\theta_D, \tilde{R}_f^*), \theta_D, \tilde{R}_f^*) && \text{(def. of } \tilde{\theta}_D \text{ in (20))} \\
&\geq H(C_0, \theta_D, \tilde{R}_f^*) && \text{(def. of } C_0^* \text{ in (17)).}
\end{aligned}$$

We therefore conclude

$$h(\tilde{\theta}_D^*) \geq H(C_0, \theta_D, \tilde{R}_f^*)$$

for all arbitrary C_0 and θ_D . In particular, the inequality also holds along $C_0 = \theta_D \cdot K_0$. Hence for all θ_D we have

$$h(\tilde{\theta}_D^*) \geq H(\theta_D \cdot K_0, \theta_D, \tilde{R}_f^*) = h(\theta_D).$$

This proves $\theta_D^* = \tilde{\theta}_D^*$. $H^* = \tilde{H}^*$ follows from

$$H^* = h(\theta_D^*) = h(\tilde{\theta}_D^*) = H(C_0^*(\tilde{\theta}_D^*, \tilde{R}_f^*), \tilde{\theta}_D^*, \tilde{R}_f^*) = \tilde{H}^*.$$

Finally, recall that $\hat{R}_f^*$ and $\tilde{R}_f^*$ respectively satisfy

$$C_0^*(\theta_D^*, \hat{R}_f^*) = \theta_D^* \cdot K_0, \quad C_0^*(\tilde{\theta}_D^*, \tilde{R}_f^*) = \tilde{\theta}_D^* \cdot K_0.$$

Then with $\theta_D^* = \tilde{\theta}_D^*$, $\hat{R}_f^*$ and $\tilde{R}_f^*$ are both solutions to the same equation (in R_f) that

$$C_0^*(\theta_D^*, R_f) = \theta_D^* \cdot K_0. \tag{23}$$

We thus expect $\hat{R}_f^* = \tilde{R}_f^*$.

□

“I am not really inclined to think there is any very effective regulation of the derivative securities markets that would be useful. People who go into it essentially ought to know what risks they’re taking and I don’t see any useful regulation.”

Kenneth Arrow

2

INFREQUENT MARKING-TO-MARKET DOES NOT REDUCE VAR NOR BACKTESTING EXCEPTIONS

Following thesis submission guidelines at Imperial College London, I acknowledge that part of this work, written by myself, was published in a peer-reviewed journal at the time of the thesis submission ([Epperlein and Herskovits, 2024](#)). Some parts are verbatim to the article.

How does VaR behave when data is scarce? We assess the effects of infrequent mark-to-market updates on VaR calibration and on the portfolio against which the VaR is backtested. We then relate our observations to the Non-Modellable Risk Factor framework adopted by the Fundamental Review of the Trading Book.

BACKGROUND

Global Markets divisions of modern investment banks now have to follow extensive regulation to undertake their derivatives market making business.

Derivatives markets are ancient, with instances dating back to the pre-Socratic era in Greece, where Thales de Miletus gathered large positions in the olive oil futures market ([Aristotle, BC](#)). Before that even, physical remains of options from the Babylonian era were exhumed (*see fig. 2.1*) ([Ferguson and Prebble, 2008](#)). Derivatives exchanges have emerged multiple times independently in autarkic nations, as the rice futures exchanges, the Dojima, in Edo Japan, from the seventeenth century onwards ([Schaede, 1989](#); [West, 2000](#)).

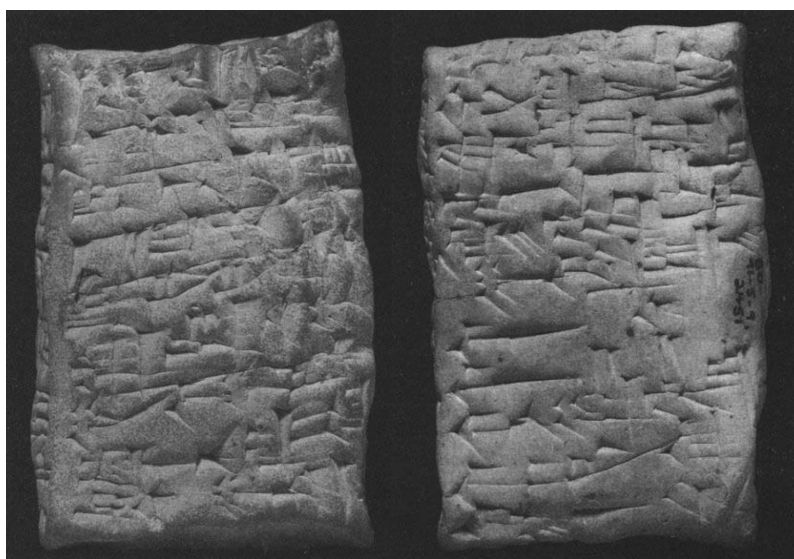


Figure 2.1: A clay tablet from second millennium BC, Mesopotamia, front (*Left*) and rear (*Right*). The inscription states that Amil-mirra will pay 330 measures of barley to the bearer of the tablet at harvest time. Image taken from ([Ferguson and Prebble, 2008](#)).

However, the derivatives markets expanded substantially across the last decades - the global volume of FX and interest rates derivatives alone is nowadays over 700 trillions US Dollars ([BCBS, 2023](#)) - rendered the permanent availability of liquidity a necessary condition for our globalized economy to function. Subsequent crises have engendered multiple sets of international financial regulations, which are the core driver of the global behaviour of market making investment banks.

The Basel Committee on Banking Supervision (BCBS) was founded in 1974 by the G10's central bank governors ([Goodhart, 2011](#)), and iterations of its “*Basel*” regulations started with the *Basel 1.0* rules in July 1988. The Barings bank implosion in 1995 motivated the British authorities to restructure financial regulation, with the creation of the Financial Services Authority (FSA) five years later ([Proctor, 1997](#)). But it also launched the rework of BCBS regulations, from consultations in 1997, to releasing the

“*Basel 2.0*” framework in 2004. The scope of “*Basel 2.0*” and reviews up to *Basel 2.5* enlarged their predecessor’s, in particular expanding on operational risk, as an answer to the Barings collapse, largely led by positions taken by a single trader.

Basel 2.0 also introduced the Internal Models Approach (IMA)¹ (BCBS, 2022b; BCBS, 2022a). Sell-side actors could then create their own models, following BCBS guidelines, to obtain their own risk limits at desk, entity and company levels (BCBS, 2009; BCBS, 2019a).

The 2008 financial crisis (GFC) impelled the BCBS to review its methodologies again, especially over its credit risk measurement² (PRA, 2023a; PRA, 2023b; BCBS, 2009; BCBS, 2019b), and in parallel led up to the dismantling of the FSA³. In the UK, changes also included a provision to tackle scenarios in which Value-at-Risk (VaR) (Best, 2000; Duffie and Pan, 1997) and Expected Shortfall (ES), the two main market risk measures outlined since “*Basel 1.0*”, were unable to capture underlying risk adequately. Named Risks Not in VaR (RNiV) (PwC, 2020), this provision is usually implemented as an add-on to internal VaR models. In the *Basel 3.1* framework, the Fundamental Review of the Trading Book (FRTB), RNiV add-ons were formally encompassed under the newly introduced Non-Modellable Risk Factors (NMRF). This is not mentioned in the official FRTB publications, but is something that was confirmed off-line with regulators responsible for designing NMRF. In fact, the ‘Near-final part 1 policy statement 17/23’ from the PRA states:

“3.54 The PRA proposed to adjust the Basel 3.1 standards to require firms to include NMRFs in their model, for the purpose of performing back-testing at the trading desk level. This would provide a less burdensome way to demonstrate that the model was appropriately conservative compared to requesting separate permissions for each exception on a potentially large number of trading desks to be disregarded (as required under the Basel 3.1 standards’ approach). However, to ensure a firm’s approach remains prudent and aligns with Basel 3.1 standards, firms would be required to exclude NMRFs from back-testing at the overall portfolio level for the purpose of calculating any capital multiplier. At the portfolio level, aligned with Basel 3.1 standards, they would be permitted to disregard exceptions that are caused by NMRFs if approved by the PRA.” Hence, we argue that there is a direct relationship between RNiV (or NMRF) and the capital multiplier.

Understanding the implications of the BCBS regulation modifications and how the FRTB is currently being implemented by sell-side institutions is tantamount to com-

¹IMA is also referred as Internal Ratings Based approach (IRB) by the bank of England (BoE, 2022a).

²IRC was formally developed prior to the GFC. Yet, its scope, and thus importance, were drastically expanded when the CDO market collapsed in the US.

³Other scandals: Northern Rock, LIBOR and PPI also played a role in its replacement.

prehend recent changes in market structure. In particular this study examines changes brought by the replacement of *Basel 2.5* by *Basel 3.1* and subsequent iterations. Under *Basel 2.5*'s framework, 97.5% of relevant financial organizations intended to use the IMA. The introduction of *Basel 3.1* has seen this share reduced to 52.5% (ISDA and for Financial Markets in Europe, 2024; PwC, 2016). This drop has material consequences for risk limits used everyday by banks. Various corporations opting for the Standard Approach (SA) instead of the IMA have the same relative risk limits, and potentially have to follow the same liquidation paths in stressed market scenarios. This leads to increased systemic risk in markets. The simplicity of the SA's models⁴ also cannot capture risk comprehensively. In particular, risks grow linearly with positions' sizes. The findings presented in this chapter support the opinion that the substantial reduction observed in the use of IMA among banks is in large part due to the unplanned and outsized impact of NMRF.

INTRODUCTION

The Fundamental Review of the Trading Book (FRTB) introduced new capital charges for assets whose trades are not observed frequently enough (BCBS, 2019a). In particular, these additional capital charges are intended to be a complement to Value-at-Risk (VaR), or more precisely, to the equivalent capital charge expressed as Expected Short-Fall (ES). To calibrate a VaR model typically requires a large amount of good quality synchronous data, ideally including stressed markets. Reflecting this requirement, regulatory guidelines mandate VaR calibration to include data going back as far as January 1st, 2007, at a daily frequency, thus including multiple financial crisis periods.

These data requirements for VaR to be efficiently calibrated naturally raise the question: how to capture risk carried by assets for which reliable regular data is not available? This problem concerns not only illiquid assets, which might only trade weekly, monthly, or even less frequently, but also liquid assets where trades are simply not observed by the investors attempting to model their risk charge.⁵ Assets traded Over The Counter (OTC) in particular are relevant to this setting, as the only trades observed by an actor calibrating its risk are ones acted by the actor itself. For instance, many local government saving bonds, but also US savings bonds only trade OTC, and are widely held by banks as collateral (see Chapter 3 for a detailed presentation of banks' exposure to sovereign debt).

Within the Internal Model Approach (IMA) of the FRTB framework, the Risk Factor Eligibility Test (RFET) determines for each asset / Risk Factor (RF) whether regular

⁴See Chapter 3 for an outline of SA's risk pricing for sovereign bonds.

⁵It should be noted that FRTB already introduced an illiquidity charge in the ES capital formula, based on a table of liquidity horizons.

VaR is acceptable as a risk charge or if it has to be included in the additional Non-Modellable Risk Factors (NMRF) charge. Generally, for an asset to be omitted from NMRF, its Mark-to-Market (MtM) value must be updated at least 24 times a year, or approximately every 10 traded days on average. *“RFET’s definition implies that VaR is not able to conservatively capture market risk for a risk factor updated less than bi-weekly”* (PRA, 2023a). RFET implies that for any asset, as price observations become rarer, the risk figures reported by our models are not able to adequately capture the risk carried by the asset. As a consequence, as market observations reach a bi-monthly frequency, RFET assumes that the model understates the asset’s risk enough that a capital multiplier larger than 1 needs to be applied to risk figures.⁶ The multiplier is applied in order to re-align the “illiquid” risk figures with the “real” risk the asset carries, that our risk models would capture were daily price updates be available.

This article attempts to demonstrate that this assumption often does not stand, neither for a single risk factor nor for correlated portfolios. Indeed, our findings demonstrate that even in situations where observed liquidity is scarce enough to trigger the RFET, classical risk models still capture risk at an adequate level. The RFET’s capital multiplier is applied to risk figures which are not deflated compared to risk figures obtained on standard time series, updated daily. As mentioned, recent ISDA surveys highlighted a material reduction in IMA usage from 97.5% (under the previous Basel 2.5 framework) to 52.5% (under FRTB) (ISDA, 2024). The results presented in this article thus hint that a major cause in this reduction might be due to the unviability of NMRF and RFET, which multiply risk figures on rarely traded assets. Assets over which VaR and ES already capture risk conservatively.

Specifically, we demonstrate that, rather counter-intuitively, VaR can initially increase as the MtM frequency decreases. To this end, we start with a single stock portfolio. With daily market updates of its value, we obtain a reference level for its 1-day 99% VaR. We then compare this VaR level with figures obtained as marking-to-market only happens periodically every T days. That is, the stock returns are 0 for on average $(T - 1)$ days before a price change happens. While the frequency of price jumps is reduced T -fold, we expect, based on an assumed ‘random walk’, that the volatility of non-zero jumps to be $\sqrt{T}$ times larger than the average daily price return. When the price jumps, T days’ worth of price moves are aggregated in one.

As a general example, let us now consider one particular case. A stock price is updated every day and its returns follow a $N(\mu, 1)$ distribution, i.e. with mean drift μ and unit variance. However, an investor holding the stock is only updated on the stock’s movements bi-weekly: $T = 10$. Let’s then generate 10 days of price returns and compare the 1-day 90% VaR for this given infrequent observer and with the full time

⁶See Section 3 for an outline of RFET’s implementation of this capital multiplier.

series available. In this given setting, with only 10 days, the 90th percentile of returns will be simply be the highest observed figure in the time series. For the full original time series, using moment generating functions (MGF), the expected value for the maximum daily return is bounded upwards by $(\mu > \sqrt{2 \log(10)} / 9)$. As for the infrequent time series, its maximum expected return is 10μ . In this toy example, we observe that the VaR figure is considerably higher for the stock-holder with infrequent marking-to-market (given $\mu > \sqrt{2 \log(10)} / 9$). Yet, the stock in this instance would not meet to the RFET.

Actually, these results typically stand without drift, when μ is null: let's now derive more general results and make our argumentation more rigorous.

1 INFREQUENT MARKING TO MARKET FOR A SINGLE RISK FACTOR

First, let S be the time series of daily risk factor values $S := \{S_t\}_{t \in \mathbb{N}}$, whose return follows a $N(0, 1)$ process. We assume that S 's increments are all independent identically distributed (*iid*) according a normal law.⁷

Let's now introduce an investor who holds one unit of this risk factor. However, this investor can only observe the risk factor's value periodically. As mentioned in the introduction, such an asset can represent non-marketable local government bonds, municipal and state bonds in particular, but also more general illiquid securities, e.g. deep out-of-the-money single stock options. Exactly once every T -days period, the investor has access to the risk factor. Let χ^T be the daily observed returns of our investor. We then have, setting $\chi_0^T = 0$:

$$\forall t \in \mathbb{N} - \{0\}, \chi_t^T = \begin{cases} \sum_{i=t-(T-1)}^t (S_i - S_{i-1}) & \text{if } t \equiv 0 \pmod{T}, \\ 0 & \text{else.} \end{cases}$$

In practice, this setup is equivalent to having an illiquid asset which only trades every T days.

For varying values of T , we are interested in understanding how the 1-day VaR of an investor observing $\{\chi_t^T\}_{t \in \mathbb{N}}$ behaves. We are thus first interested in finding the cumulative distribution function (*CDF*) of returns uniformly sampled from the full time series $\{\chi_t^T\}_{t \in \mathbb{N}}, \chi^T$.

χ^T is a mixture of two distributions. Essentially, χ^T has “fatter tails” the larger T is. One is a constant random variable (*RV*) $C_0 \equiv 0$. Its *CDF* is:

⁷While our analysis focuses on normally distributed returns, we have confirmed that our findings remain true for *t* - *student* distributed returns. We expect further our results to remain valid for usual returns' distributions, centered at 0 and symmetric.

$$F_{C_0}(x) = \begin{cases} 0 & \text{if } x < 0, \\ 1 & \text{if } x \geq 0. \end{cases}$$

The second component R_T of χ^T 's mixture represents T days of aggregated risk factor returns:

$$R_T \sim N(0, T).$$

We can now express χ^T 's CDF:

$$F_{\chi^T}(x) = \frac{1}{T} F_{R_T}(x) + \frac{T-1}{T} F_{C_0}(x).$$

In practice, χ^T is composed of sequences of null returns followed by a large periodic Profit & Loss (P&L) marking update. This large update has frequency $1/T$. Hence, every quantile value of χ^T in the $(1/T, 1 - 1/T)$ range is 0. Yet, the remaining tail quantiles are drafted from a $N(0, T)$ distribution. For returns marked-to-market daily, the same tail quantiles follow a $N(0, 1)$ distribution. That is, χ^T has fatter tails the higher T is (*see fig. 2.2*).

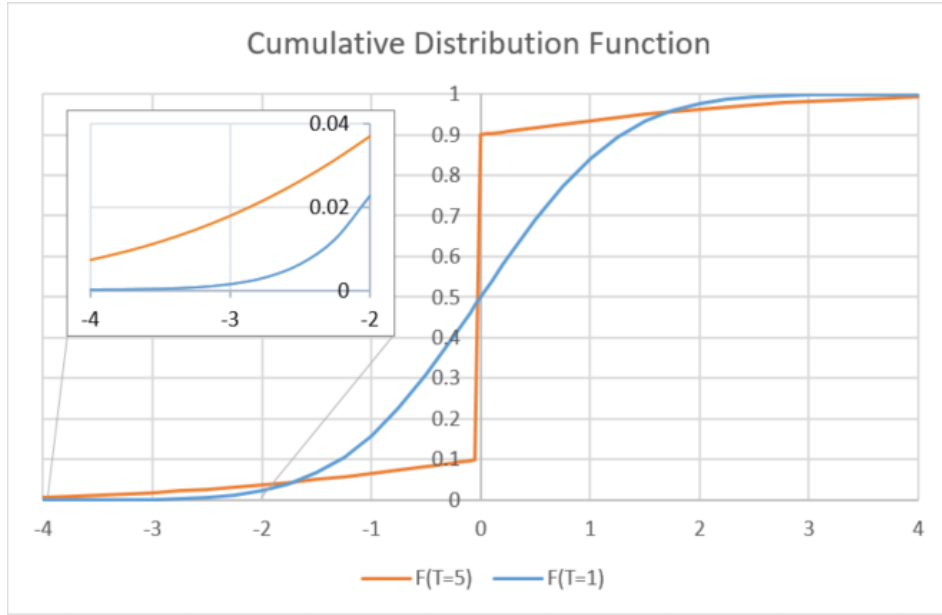


Figure 2.2: Cumulative distribution functions of observed P&L for risk factor returns observed daily (Blue) and weekly (Amber). *Detail:* Tail Distributions of observed P&Ls. With weekly observations, the frequency of null returns creates a CDF jump at 0. The tails are however driven by the large infrequent returns, which are then “fatter”, as observed in the detail.

For varying observation periods T , we want to compute the 1-day VaR associated the investor's portfolio, at a certain confidence level $\alpha \in (0, 1)$. We focus specifically on the regulatory mandated 99% VaR and 97.5% ES levels, i.e. for $\alpha = 0.99$ and $\alpha = 0.975$.⁸

⁸We focus on the 99% upwards VaR and 97.5% ES for simplicity. As VaR and ES are symmetric, our results stand for downwards 1% VaR and 2.5% ES too.

We want to find x such that it solves:

$$\alpha = \frac{1}{T} F_{R_T}(x) + \frac{T-1}{T} F_{C_0}(x).^9$$

Here, as we are looking at tail distributions, we automatically get that $F_{C_0}(x)$ is 1. The previous equation accordingly becomes:

$$1 - T(1 - \alpha) = F_{R_T}(x).$$

R_T is a normal distribution with variance 1, so:

$$F_{R_T}(x) = F_{R_1}\left(\frac{x}{\sqrt{T}}\right).$$

As R_1 follows a standard normal distribution, we get¹⁰:

$$\text{VaR}_{1D}^T(\alpha) = \sqrt{T} \Phi^{-1}(1 - T(1 - \alpha)).$$

Analogously, we can derive a more general formula for the M -days VaR at level α , $\text{VaR}_{MD}^T(\alpha)$.

We first focus on the regime where the risk measure's horizon M is shorter than the observation frequency T : $M < T$. Analog to the $T = 1$ case, we obtain:

Theorem 1.1.

$$\text{VaR}_{MD}^T(\alpha) = \sqrt{T} \Phi^{-1}\left(1 - \frac{T(1 - \alpha)}{M}\right).$$

Proof of Theorem 1.1. Let us first define the probability density function of a more general mixture of normals than described above. Given $T \in \mathbb{N}^*$, let the random variable χ be sampled with probability $1/T$ from an $N(0, \sigma_T)$ distribution and with probability $(T-1)/T$ from an $N(0, \sigma_0)$ distribution.

We can then write χ 's probability density function as:

$$f_\chi(x) = \frac{1}{T} \frac{1}{\sigma_T \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma_T^2}\right) + \frac{T-1}{T} \frac{1}{\sigma_0 \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma_0^2}\right).$$

Let us now assume that we have historical time series that is *iid* and normally distributed with unit variance but only observed every T -days period. χ^T is the resulting

⁹The right-hand side defines the *cdf* of the infrequently observed returns, for a randomly selected day (uniformly sampled from the set of days). While the observation period is fixed, the resulting returns' distribution is a mixture of two distributions.

¹⁰ $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2}\right) dt$ is the standard normal law's *CDF*.

infrequently marked-to-market time series of returns. Informally, we obtain χ^T 's probability density function by letting σ_0 tend to 0 in the previous $f_\chi(x)$ definition, as

$$\delta(x) = \lim_{\sigma_0 \rightarrow 0} \frac{1}{\sigma_0 \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma_0^2}\right). \text{ Formally, we express } \chi^T\text{'s pdf as:}$$

$$f_{\chi^T}(x) = \frac{1}{T} \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{x^2}{2T}\right) + \frac{T-1}{T} \delta(x),$$

where $\delta(x)$ is the Dirac measure, and where $\sigma_T = \sqrt{T}$. Indeed, daily returns follow an $N(0, 1)$ distribution. T days of independent returns accumulated thus follow an $N(0, \sqrt{T})$ distribution.

The previous pdf can be generalized for calculating the distribution over an M -day horizon (given that $M \leq T$) as follows:

$$f_{\chi_M^T}(x) = \frac{M}{T} \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{x^2}{2T}\right) + \frac{T-M}{T} \delta(x).$$

Considering M -days rolling returns, with $M \leq T$, each M -days window contains at most 1 non-zero entry. That is, any observed return at date t appears as the sole non-zero observed return for the next M days, from date t to date $(t + M - 1)$. As in the $M = 1$ case, each observed return follows an $N(0, \sqrt{T})$ distribution. After M days since the last non-zero return has been observed, at date $(t + M)$, the M -days rolling window contains null returns only: the M -days returns reported is 0.

We can now calculate the VaR at an α confidence level by solving for:

$$\begin{aligned} \alpha &= \int_{-\infty}^{\text{VaR}} f_{\chi_M^T}(x) dx, \\ &= \int_{-\infty}^{\text{VaR}} \left[\frac{M}{T} \frac{1}{\sqrt{2\pi T}} \exp\left(-\frac{x^2}{2T}\right) + \frac{T-M}{T} \delta(x) \right] dx, \\ &= \frac{M}{T} \Phi\left(\frac{\text{VaR}}{\sqrt{T}}\right) + \frac{T-M}{T}, \end{aligned}$$

where we used $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x \exp\left(-\frac{t^2}{2}\right) dt$ and $1 = \int \delta(x) dx$.

□

Corollary 1.2. *Let $R = \{R_t\}_{t \in \mathbb{N}}$ be a time series of normally distributed iid returns.*

Then we have that given a return window $M \in \mathbb{N}$ and an observation period $T \in \mathbb{N}$:

$$VaR_{MD}^T(\alpha) = \begin{cases} VaR_{MD}^1(\alpha), & \text{if } T < M \text{ and } M \equiv 0 \pmod{T}, \\ \sqrt{T} \Phi^{-1} \left(1 - \frac{T(1-\alpha)}{M} \right), & \text{else, if } T \geq M. \end{cases}$$

Proof of Corollary 1.2. We already proved the case where $T \geq M$, we complete the proof for $T < M$.

Without loss of generality, let's assume that $R_t \sim N(0, 1)$, $\forall t \in \mathbb{N}$.

Let $\{R_t\}_{t \in \mathbb{N}}$ be the original time series and $\{\chi_M^T(t)\}_{t \in \mathbb{N}}$ the rolling M -days returns constructed from infrequent returns, observed every T days.

$\forall t \in \mathbb{N}$, $\chi_M^T(t)$ is the sum of the past M values of the χ_1^T time series: $\{\chi_1^T, t \in [t - M + 1, t]\}$. In this case, since M is a multiple of T , for all dates t , there are always exactly M/T observed returns in the past M days $\{\chi_1^T, t \in [t - M + 1, t]\}$. Thus each entry of the M -days rolling returns χ_M^T is distributed according to an $N\left(0, \sqrt{M/T \cdot T^2}\right)$ distribution. χ_M^T 's *pdf* is:

$$f_{\chi_M^T}(x) = \frac{1}{\sqrt{2\pi M}} \exp\left(-\frac{x^2}{2M}\right).$$

We also have that if $M \equiv 0 \pmod{t}$, then $\chi_M^T(t) = \chi_M^T(\tilde{t})$, $\forall \tilde{t} \in [t; t + M - 1]$ as no new returns will be observed in the $(M - 1)$ days following t .

We had that $\forall t \in \mathbb{N}$, $\chi_1^T(t) \sim N(0, M)$:

$$f_{\chi_1^T}(x) = \frac{1}{\sqrt{2\pi M}} \exp\left(-\frac{x^2}{2M}\right).$$

As these the two time series of returns have the same distribution, we get that the distribution of their quantiles is identical and thus:

$$\forall \alpha \in (0, 1), \text{VaR}_{MD}^T(\alpha) = \text{VaR}_{MD}^1(\alpha).$$

□

Remark 1.3. *Contrasting with the previous setting, the case $M \not\equiv 0 \pmod{T}$, with $T < M$ does not have the desired $\forall \alpha \in (0, 1)$, $\text{VaR}_{MD}^T(\alpha) = \text{VaR}_{MD}^1(\alpha)$.*

Indeed, if $M \not\equiv 0 \pmod{T}$, with $T < M$, each window of M consecutive terms of χ_1^T contains either $\lceil M/T \rceil$ or $\lfloor M/T \rfloor$ non-null entries. On average, the aforementioned window still contains M/T observed returns, but $\chi_M^T(t)$'s *pdf* now is a mixture of two distinct normal distribution.

The window contains $\lceil M/T \rceil$ non-null entries with probability $p = M/T - \lfloor M/T \rfloor$, since p solves $\lceil M/T \rceil \cdot p + \lfloor M/T \rfloor \cdot (1 - p) = M/T$.

When the sum of M individual returns composing $\chi_M^T(t)$ contains $\lceil M/T \rceil$ observed days, $\chi_M^T(t) \sim N(0, \lceil M/T \rceil \cdot T)$.

Analogously, when the sum of M individual returns composing $\chi_M^T(t)$ contains $\lfloor M/T \rfloor$ observed days, $\chi_M^T(t) \sim N(0, \lfloor M/T \rfloor \cdot T)$.

Thus, χ_M^T 's distribution is a mixture of normals, and its *pdf* is:

$$f_{\chi_M^T}(x) = (M/T - \lfloor M/T \rfloor) \frac{1}{\sqrt{2\pi \lceil M/T \rceil \cdot T}} \exp\left(-\frac{x^2}{2 \lceil M/T \rceil \cdot T}\right) + (\lceil M/T \rceil - M/T) \frac{1}{\sqrt{2\pi \lfloor M/T \rfloor \cdot T}} \exp\left(-\frac{x^2}{2 \lfloor M/T \rfloor \cdot T}\right).$$

As $f_{\chi_M^T} = f_{\chi_M^1}$ fails, we do not have the desired VaR equality.

Going back to the $M \leq T$ case, from the analytical formula for the infrequent Value-at-Risk, we can also obtain an analytical formula for the Expected Shortfall (ES) at level α , $ES_{MD}^T(\alpha)$. From the Expected Shortfall's definition as the average of the returns' distribution tail beyond the VaR at level α , we get:

Theorem 1.4.

$$ES_{MD}^T(\alpha) = \frac{M}{(1 - \alpha)\sqrt{2\pi T}} \exp\left(-\frac{1}{2} \left[\Phi^{-1}\left(1 - \frac{T(1 - \alpha)}{M}\right) \right]^2\right).$$

Proof of Theorem 1.4. The expected shortfall at level α of a distribution with *pdf* f and *cdf* F is defined as the average level of the distribution above its α^{th} quantile. That is:

$$ES_f(\alpha) = \frac{\int_{F^{-1}(\alpha)}^{\infty} x f(x) dx}{\int_{F^{-1}(\alpha)}^{\infty} f(x) dx} = \frac{\int_{F^{-1}(\alpha)}^{\infty} x f(x) dx}{1 - \alpha},$$

For any random distribution with *cdf* F , the VaR at level $\alpha \in (0, 1)$ is precisely defined as $F^{-1}(\alpha)$:

$$\begin{aligned} ES_{MD}^T(\alpha) &= -\frac{\int_{-\infty}^{-\text{VaR}_{MD}^T(\alpha)} x f(x) dx}{1 - \alpha}, \\ &= -\frac{\int_{-\infty}^{-\text{VaR}_{MD}^T(\alpha)} \left[\frac{M}{T} \frac{x}{\sqrt{2\pi T}} \exp\left(-\frac{x^2}{2T}\right) + \frac{T - M}{T} x \delta(x) \right] dx}{1 - \alpha}. \end{aligned}$$

Using $0 = \int x \delta(x) dx$, and making the change of variable $u = x^2/2T$, such that $x dx = T du$, we get:

$$\begin{aligned} \text{ES}_{MD}^T(\alpha) &= -\frac{\int_{\infty}^{\left(\frac{\text{VaR}_{MD}^T(\alpha)^2}{2T}\right)} \left[\frac{M}{\sqrt{2\pi T}} \exp(-u) \right] du}{1 - \alpha}, \\ &= \frac{\left[\frac{M}{\sqrt{2\pi T}} \exp(-u) \right]_{\infty}^{\left(\frac{\text{VaR}_{MD}^T(\alpha)^2}{2T}\right)}}{1 - \alpha}, \\ &= \frac{M}{(1 - \alpha)\sqrt{2\pi T}} \exp\left(-\frac{\text{VaR}_{MD}^T(\alpha)^2}{2T}\right). \end{aligned}$$

We replace $\text{VaR}_{MD}^T(\alpha)$ by its analytical formula derived previously to get the claim:

$$\text{ES}_{MD}^T(\alpha) = \frac{M}{(1 - \alpha)\sqrt{2\pi T}} \exp\left(-\frac{1}{2} \left[\Phi^{-1}\left(1 - \frac{T(1 - \alpha)}{M}\right) \right]^2\right).$$

□

By comparing the formulas for the two risk measures, $\text{ES}_{MD}^T(\alpha)$ and $\text{VaR}_{MD}^T(\alpha)$, we observe the well-known:

$$\text{ES}_{MD}^T(0.975) \approx \text{VaR}_{MD}^T(0.99).$$

Fig. 2.3 below shows 1-day 99% VaR and 10-days 99% VaR as functions of T , normalized by $\text{VaR}_{1D}^1(99\%)$ and $\text{VaR}_{10D}^1(99\%)$, respectively. We also display the normalized 1-day and 10-days ES for $\alpha = 0.975$, as they correspond to near identical values to 99% VaR at $T = 1$.

The same figure shows the average VaR surcharge values on actual CDS returns, across the 2007 – 2022, and to the 2008 and 2020 crises periods. The dataset is comprised of 1138 5-year maturity CDS time series (*See Appendix*).¹¹ For the empirical analysis we chose the CDS markets as they are more likely to display lack of observability/liquidity, especially during stress periods.

For each empirical CDS time series, observed every T days, the α^{th} percentile of the 15 years of available M -days returns is determined to obtain the $\text{VaR}_{MD}^T(\alpha)$. Graphs

¹¹The empirical figures presented are computed using absolute returns. Implementing log or relative returns instead yields qualitatively and quantitatively similar VaR changes when MtM frequency is reduced. Using equity returns time series instead also does not change the directional impact of infrequent MtM (*See Appendix*).

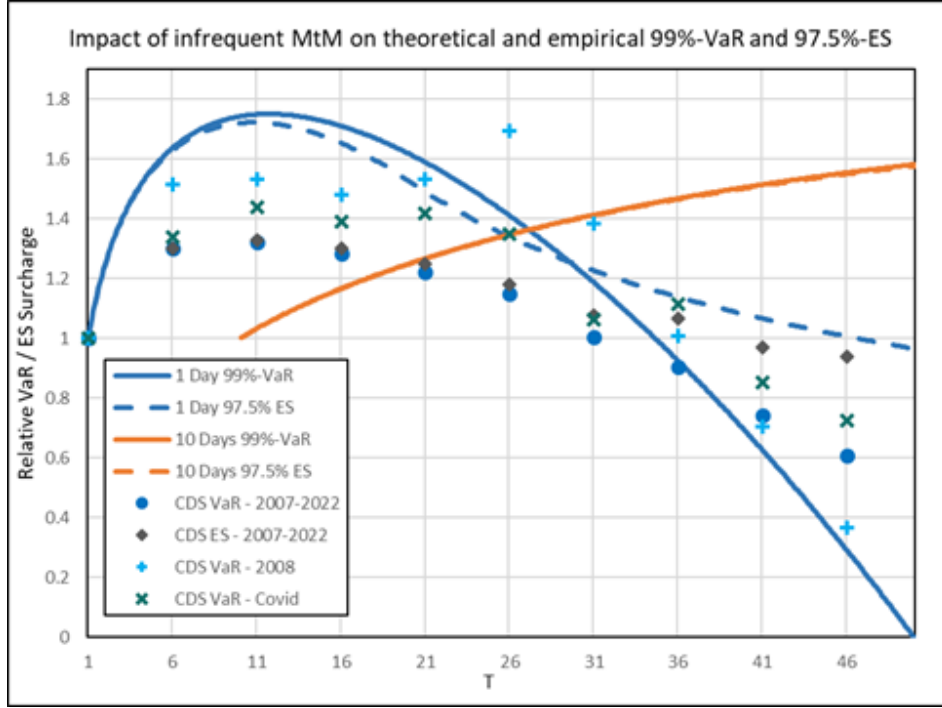


Figure 2.3: VaR and ES average levels for varying observation periods T . Normalized by figures obtained for $T = 1$. In *Solid and Dashed Blue*, 99% 1-day VaR and 97.5% 1-day ES, respectively. In *solid and dashed amber*, 99% 10-days VaR and 97.5% 10-days ES, respectively. Empirical CDS results for downwards 99% 1-day VaR are shown for 2007-2022 period, Credit Crisis and Covid Crisis (*Scatter: Blue*). Empirical downwards 97.5% 1-day ES (*Scatter: Grey*) for the 2007-2022 period.

below present the average resulting from this process performed over the 1138 available CDS names.

It is commonly assumed that infrequent market updates can lead to lower VaR figures. However, our analytical model shows that a 10 days market update leads to a 74% increase in 1-day VaR, going down to 62% for 20 days and approaching parity for 35 days. Similarly, the augmentation for 10-days VaR persists even beyond 35 days, for up to 350 days. However, we should also note that both 1-day VaR and 10-days VaR eventually approach zero when T reaches 50 days and 500 days, respectively, as the tail becomes depleted.

This undesirable behaviour is not seen with a coherent risk measure such ES, where the risk doesn't go to zero for large T . Also, we observe that the 97.5% 1-day ES only approaches parity at $T = 47$. For smaller values of T both 99% VaR and 97.5% ES converge, and so for simplicity we continue to focus our analysis on the VaR metric.

If we now look at the empirical CDS data on *fig. 2.3*, we see qualitatively similar results regarding VaR's amplification as a function of T . In particular, under the stressed market conditions, the average empirical autocorrelation increases (*see fig. A.14*), and the impact of infrequent MtM on VaR & ES is stronger than in regular market conditions. Indeed, the volatility of infrequent returns is larger than in normal market

times: consecutive daily returns are correlated, thus their sum χ^T 's distribution is more volatile. Hence, we continue our analysis based on the analytical model.

On a global level, reducing the frequency of price observations increases drastically the magnitude of the price jumps when they happen. Since 1 day-VaR is calibrated over single day returns only, the infrequent large returns obtained through rare MtM thus increase the final VaR figures.

For an observation frequency of T days, the same reasoning applies to M -days VaR, whenever $M < T$. Indeed, M -days VaR essentially relies on the distribution of returns computed over M -consecutive days. In general, M -days VaR scales with $\sqrt{M}$ times the 1-day VaR. As returns constructed from infrequent price measures aggregate T days of price changes at once, the magnitude of these price jumps analogously scales with $\sqrt{T}$ times the 1-day VaR. As long as $M < T$, the M -day VaR is also increased when MtM is infrequent.

Definition 1.5. *Generally, given (T, M, α) , $\forall T, M \in \mathbb{N}$ and $\alpha \in (0, 1)$, we define the relative surcharge incurred to the M -day VaR by the infrequent observations:*

$$RelVaR(T, M, \alpha) = \frac{VaR_{MD}^T(\alpha)}{VaR_{MD}^1(\alpha)}.$$

Using this definition, we can obtain the relative VaR surcharge incurred by the infrequent observation frequencies:

Lemma 1.6.

$$RelVaR(T, M, \alpha) = \sqrt{\frac{T}{M}} \frac{\Phi^{-1}\left(\frac{T}{M}(1 - \alpha)\right)}{\Phi^{-1}(1 - \alpha)}.$$

In particular, we have the scaling invariance of RelVaR over the VaR window M :

$$RelVaR(T, M, \alpha) = RelVaR(T/M, 1, \alpha).$$

Proof of Lemma 1.6. From Theorem 1.1: $VaR_{MD}^T(\alpha) = \sqrt{T} \Phi^{-1}\left(\frac{T}{M}(1 - \alpha)\right)$.

In parallel, $VaR_{MD}^1(\alpha) = \sqrt{M} VaR_{1D}^1(\alpha)$.

$$RelVaR(T, M, \alpha) = \sqrt{\frac{T}{M}} \frac{\Phi^{-1}\left(1 - \frac{T}{M}(1 - \alpha)\right)}{\Phi^{-1}(\alpha)},$$

$$RelVaR(T, M, \alpha) = \sqrt{\frac{T}{M}} \frac{\Phi^{-1}\left(\frac{T}{M}(1 - \alpha)\right)}{\Phi^{-1}(1 - \alpha)},$$

where we used Φ^{-1} 's symmetry.

$RelVaR(T, M, \alpha) = RelVaR(T/M, 1, \alpha)$ follows immediately.

□

Changing the VaR horizon M does not change the structure of the relative surcharge incurred by the low MtM frequency. Increasing M simply slides the surcharge per observation frequency T to the right. Consequently, it is enough to study the 1-day VaR, as it is straightforward to obtain results for the M -days VaR.

In particular, for a given confidence level α , we are interested in the maximum VaR (*resp.* ES) surcharge incurred by infrequent MtM, and by the maximal observation frequency increasing the VaR (ES). For varying tail levels α , *fig.* 2.4 presents the maximal observation periods T for which the VaR is increased by the infrequent MtM.

We should note for $T = 2$, the smallest infrequent observation periodicity, we need $\alpha > 91\%$ for us to observe a VaR amplification. Indeed, removing every other observation, we get the strict $VaR_{1D}^2(\alpha) = 0, \forall \alpha \in [0.25, 0.75]$. α needs to increase to above 91% to exhibit any VaR increase.

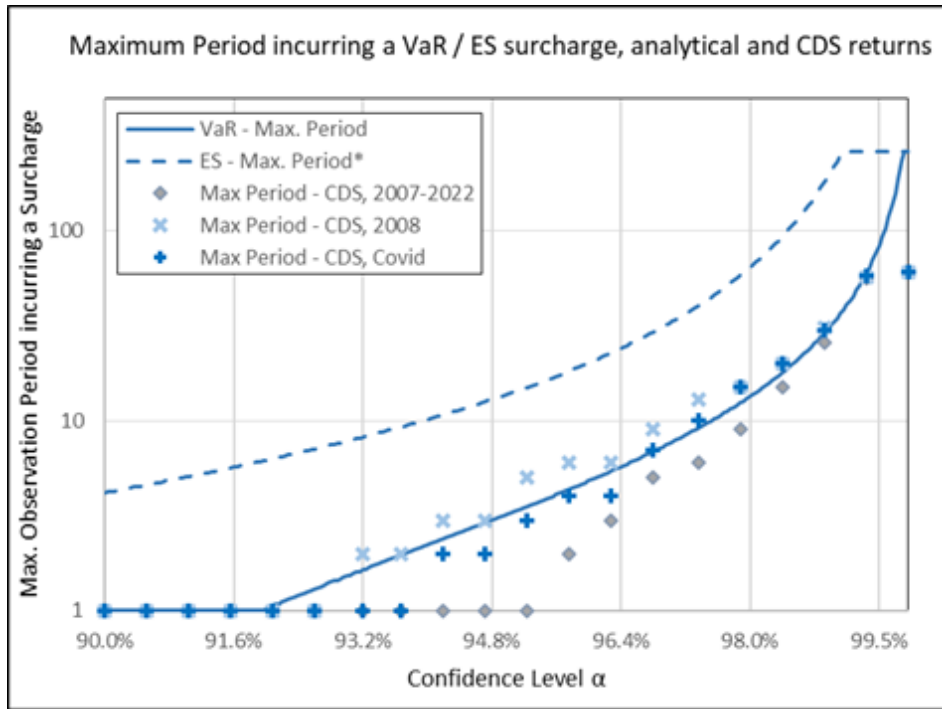


Figure 2.4: For varying confidence levels α , is shown the maximum observation period (in days) above which infrequent MtM does not amplify the VaR (*Solid*) and ES (*Dashed*) at level α (*Empirical CDS values: scatter plots, for the 15 years window and stressed market periods*). For all VaR figures generated with more regular MtM, the VaR is heightened. *: capped at yearly observation frequency ($T = 260$).

For all typical tail levels α (above 99%) however, VaR and ES swell for low values of T . Even monthly observation frequencies yield an increase in VaR and ES. The further in the tail, the greater the magnitude of the surcharge and the larger the range of frequencies for which surcharges are observed.

In short, we showed that in usual VaR and ES computations, infrequent market updates

enlarge reported figures. Even with half the risk factor updates required to meet the RFET, we see higher VaR and ES levels than based on daily updates. In fact, even with only 8 (i.e. $> 250/35$) observations per year the 1-day VaR would on average be greater than the equivalent 1-day VaR with daily observations.

In point of fact, we now demonstrate that these rises are maintained even within hedged portfolios.

2 INFREQUENT MTM'S EFFECT ON CORRELATED RISK FACTORS

To capture the impact of scarce MtM updates on structured positions, we institute an index $\mathbb{I} := \{\mathbb{I}_t\}_{t \in \mathbb{N}}$, with $N(0, 1)$ returns. The single risk factor S has correlation ρ with the index. As a regularly traded risk factor on public markets, the investor always has daily access to the index. Contrary to the individual risk factor, MtM every T days. As we let the number of days T between each of the single risk factor's observations increases, the apparent absolute correlation to the index daily time series decreases:

$$\text{Corr}(\chi^T; T) = \frac{\text{Cov}(\chi^T; T)}{\sigma_{\chi^T} \sigma_T} = \frac{(\rho/T)}{(1 \cdot 1)} = \frac{\rho}{T}.$$

Using,

$$\text{Cov}(\chi^T; T) = \mathbb{E} \left(\frac{1}{T} \sum_{i=1}^T S_i \cdot \sum_{j=1}^T \frac{1}{T} T_j \right) = \frac{T}{T^2} \rho.$$

As in the single risk factor setup, given T , to obtain the 1-day VaR at level α is equivalent to solving for x in:

$$\alpha = \frac{T-1}{T} \Phi(x) + \frac{1}{T} \Phi \left(\frac{x}{\sqrt{1 + T + 2\rho/T}} \right).$$

Equivalently, lowering the risk factor's MtM frequency decorrelates the investor's risk factors. With long-short positions (e.g., the investor hedges his risk factor with the index), a lower absolute correlation level effectively nullifies the hedged structure of the portfolio.

As a consequence, the VaR level drastically increases. Intuitively, this augmentation would be mirrored by an identical reduction in VaR for positively correlated risk factors. However, the results from the previous section still apply: the VaR level for the single risk factor is typically higher when MtM frequency is lessened. So the baseline risk charge for holding the risk factor is enlarged: for independent risk factors (i.e. $\rho = 0$), the portfolio's VaR is raised (*see fig. 2.5*) (until T reaches 35 for $\alpha = 99\%$). Only for highly correlated risk factors does the portfolio's VaR slightly subside when

the risk factor is not observed daily.¹²

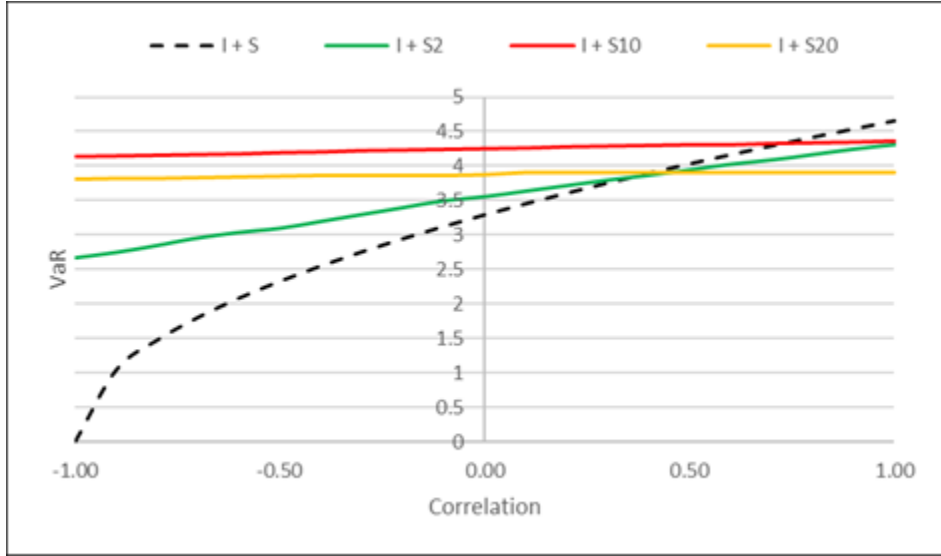


Figure 2.5: 1-day 99% VaR of portfolio containing one unit of the risk factor and of the index, for varying correlation levels between the index and the single risk factor. *Solid:* The index's value is observed daily, while the risk factor's is available every second day (*Green*), bi-weekly (*Red*) and monthly (*Amber*). *Dashed:* Both the index and the single risk factor are observed daily.

Periodic updates imply a rise in VaR charges, both for single risk factors and for typical structured portfolios. Separately, regulatory capital multipliers are applied whenever portfolio returns exceed VaR too often.

3 VaR BACKTESTING EXCEPTIONS WITH INFREQUENT P&L

Here, our starting point is that of a perfectly calibrated VaR model¹³ where its input data is observed daily. However, this time around, we are going to gauge the accuracy of the VaR model by conducting backtesting where the portfolio P&L is only observed every T days.

When backtesting, we count any instance of P&L surpassing VaR as a VaR exception. We show that given a VaR model, when a backtest is conducted with an infrequently observed time series, the rate of exceptions increases.

Specifically, under current *Basel* regulation, a VaR model exhibiting too many exceptions automatically triggers an investigation leading to a model's disqualification.

¹²As in the previous section, we confirmed that our results remain valid on real structured equity portfolios (*See Appendix*). Single risk factors hedged by the average market returns, the index, exhibit a VaR surcharge at all typical VaR levels.

¹³In this section, $\text{VaR}(\alpha)$ will systematically refer to $\text{VaR}_{1D}^1(\alpha)$ using the previous sections' notations.

Indeed, focusing on daily returns, for 10 or more VaR exceptions to be observed every year, the model would necessarily be missing a consequent proportion of the risk carried by risk factor. Such a model would thus be an unsuit candidate for a conservative measure of risk.

Let us consider a risk factor which available returns do not meet the RFET. Then, NMRF mandates that a supplementary charge is to be applied to this risk factor's VaR figures, assuming the lack of data results in the model missing some risk. The value of the supplementary charge mandated by regulations is designed to match the level of missing risk by the VaR model, yielding final VaR figures in line with the real underlying risk. However, we show in Section 4 that this is not the case. Indeed, in typical scenarios triggering the RFET, the charge applied is very high, assuming that the model is omitting a vast share of the real risk. Actually, we demonstrate that for a VaR model to leave out such a large share of the risk, the number of exceptions observed in the first place during its backtesting would have been enough to get the model rejected. That is, even when backtesting the model with infrequent observations.

With daily MtM updates, we formally define the probability $p(\alpha)$ of a $\text{VaR}(\alpha)$ exception from P&L by:

$$p(\alpha) := \mathbb{P}(\chi^1 > \text{VaR}_{1D}^1(\alpha)).$$

For independent and normally distributed daily returns, $p(\alpha)$ is given by:

$$p(\alpha) = \Phi(-\text{VaR}(\alpha)) = \Phi(\Phi^{-1}(1 - \alpha)) = 1 - \alpha.$$

However, if the MtM occurs every T days, the expression for $p(\alpha, T)$ becomes¹⁴,

$$p(\alpha, T) = \Phi\left(\frac{\Phi^{-1}(1 - \alpha)}{\sqrt{T}}\right).$$

Hence:

MtM Period T	$p(99\%, T)$	$p(97.5\%, T)$
1	0.01	0.025
2	0.05	0.08
5	0.15	0.19
21	0.31	0.33
125	0.42	0.43
250	0.44	0.45

Table 2.1: For varying MtM periods T , we show the probability of a VaR exception being observed, both for 97.5% and 99% 1D-VaR. As observations get scarce, the likelihood of a VaR exception increases.

For a 99% VaR, we get that $p(\alpha, T)$ increases fivefold from 0.01 to about 0.05 as T is changed from 1 to 2. It then continues to grow more gradually to 0.15 for $T = 5$

¹⁴By assuming P&L to be *iid*, its volatility scales by $\sqrt{T}$ when MtM frequency is T days.

(weekly update). For all tail levels α , as T tends towards ∞ , $p(\alpha, T)$ is monotone. It first grows rapidly then slowly converges to $1/2$. However, while the probability of a backtesting exception grows as the MtM update period T increases, the number of effective backtesting trials decreases as $1/T$. Indeed, the frequency of non-zero returns, which can trigger an exception, is $1/T$. Hence, the average VaR exceptions become:

$$E(\alpha, T) = \frac{250}{T} \cdot p(\alpha, T).$$

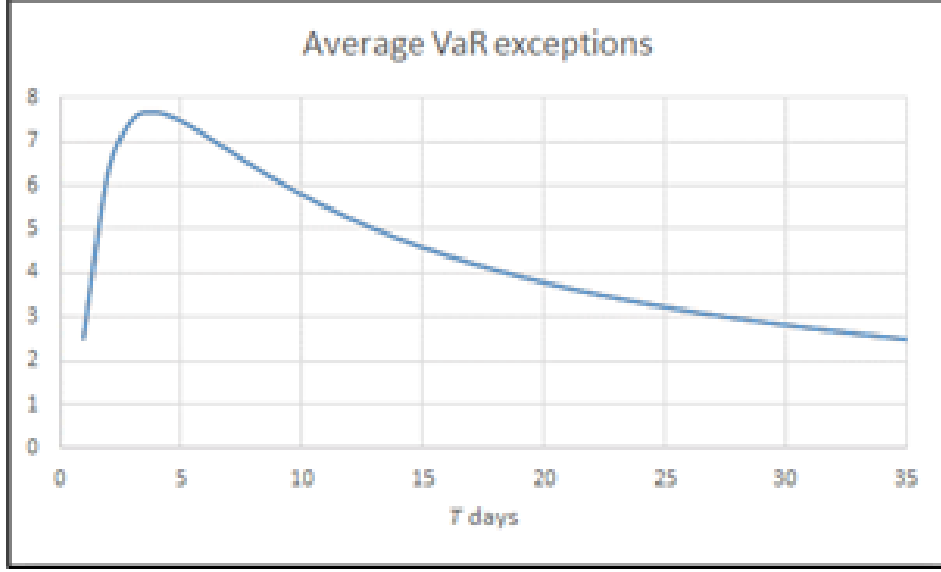


Figure 2.6: $E(0.99, T)$. Yearly expected number of 1-day 99% VaR exceptions as a function of MtM period T . While $p(0.99, T)$ increases monotonically, $E(0.99, T)$ decreases slowly weekly updates are reached, since the number of backtesting trials is reduced.

From *fig. 2.6*, we get that with $T = 1$, there is about one in ten (11%) chance of getting 5 or more exceptions even with a perfect model. However, by the time we get to 10 exceptions, the probability is only one in four thousand (0.025%). The regulators have used this fact to impose penalties through capital multipliers as the number of VaR exceptions increases from 5 to 10+.

Fig. 2.6 shows explicitly that $E(99\%, T)$ is nearly maximal at about 7.5 for a weekly update frequency. It can also be shown that the probability of getting 5 or more exceptions increases to 88%, with 11% probability of attaining 11 or more exceptions.

In parallel, given $T > 1$, to satisfy $\frac{E(\alpha, T)}{E(\alpha, 1)} = 1$ yields:

$$\sqrt{T} \cdot \Phi^{-1}(T(1 - \alpha)) = \Phi^{-1}(1 - \alpha),$$

which is equivalent to satisfying $RelVaR(T, 1, \alpha) = 1$.¹⁵ The MtM periods T across which the VaR is increased coincide with the periods where VaR exceptions are more

¹⁵Or equivalently: $VaR_{1D}^T(\alpha) = VaR_{1D}^1(\alpha)$.

likely.

Hence, like in *fig. 2.3*, we find that when $T = 35$ the average expected number of VaR exceptions returns to 2.5, which is the number obtained when $T = 1$.

4 VaR AND BACKTESTING UNDER DAILY MTM

In the absence of observability issues, it is important to review the relationship between ‘true’ missing risk, i.e. Risk-Not-in-VaR (RNiV), and backtesting.

In order to ensure capital adequacy, regulators impose a capital multiplier k when the number of 1-day 99% VaR exceptions exceeds 4, as follows¹⁶.

$$k = \frac{\Phi^{-1}(1 - \alpha)}{\Phi^{-1}\left(\frac{\text{VaR exceptions}}{250}\right)} \left(\equiv 1 + \frac{\text{RNiV}}{\text{VaR}} \right).$$

The multiplier (or RNiV add-on) corrects for the missing risk to the VaR, so as to bring it back to 99% confidence level.¹⁷ Hence, we can turn the problem around and estimate the expected number of VaR exceptions for a given quantum of missing risk. This is shown in *fig. 2.7*, where we extend amount of missing risk up to 100%. *Fig. 2.7* also shows (*Black*) the probability of avoiding a Type-2 error (i.e. statistical power): should a VaR model be missing a large amount of ‘true’ risk, with high probability, the model would have been rejected in the first place.

The power of the test to reject a bad model becomes weaker as the presumed amount of missing risk becomes small in the amber zone. Conversely, if 1/3 of the VaR were ‘missing’ we would expect an average of 10 exceptions, with only a 3% chance of landing on the green zone (Type 2 Error), and wrongly assuming the model behaved correctly. Moreover, the probability of a Type 2 error becomes truly remote, $O(10^{-10})$, when VaR is 100% understated.

Consequently, if NMRF was to be interpreted as RNiV, it should be capped at 33%, similar to how the regulatory backtesting multiplier k is capped at 10+ exceptions.¹⁸

¹⁶Within Basel 2.5’s framework, 4 backtesting exceptions a year automatically trigger an investigation for missing risks.

¹⁷In reality, the FRTB capital multiplier is set at $1.5 \cdot k$, with a floor of 1.5 in the green zone and cap of 2 in the red zone.

¹⁸Typical FRTB implementations can generate NMRF/ES in excess of 100%, which is truly incommensurate with the risk.

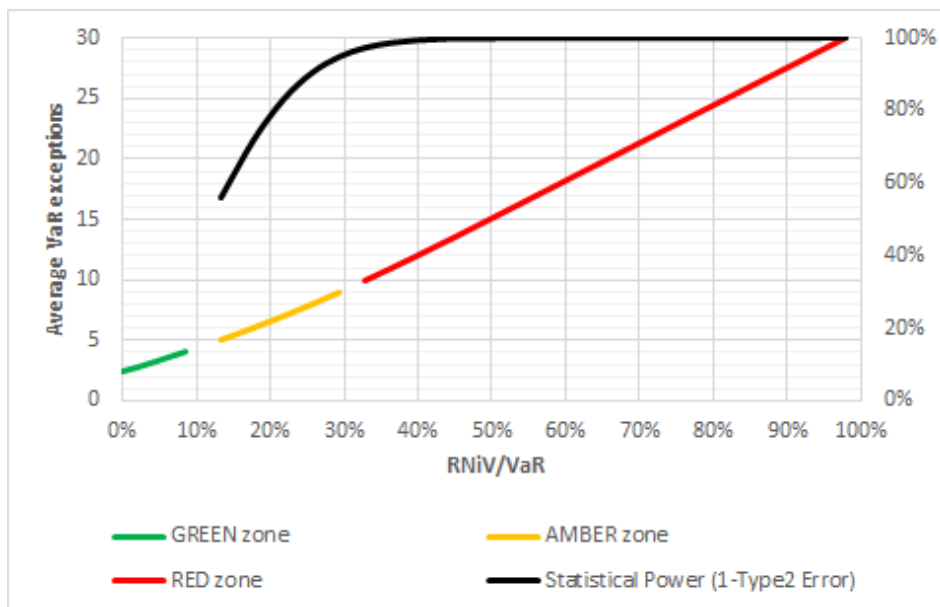


Figure 2.7: Expected number of 99%-VaR exceptions as a function of % of missing risk, across *GREEN*, *AMBER* and *RED* zones. Also, shown is the power of the backtest (*RHS axis, Black*).

5 CONCLUSION

This chapter shows infrequent MtM can dramatically increase VaR and ES. For both 1-day 99% VaR and 1-day 97.5% ES this increase peaks at about 74% for a bi-weekly observation period. For VaR, the surcharge disappears at a 35-days observation period, which is 25 days above the threshold required to pass the RFET. ES remains heightened for less frequent observation frequencies. We also demonstrated that this tendency remains true for most structured portfolios. A similar effect is also observed for 10-days 99% VaR and 10-days 97.5% ES, but where the 74% VaR / ES increase occurs with 10×10 days observation period. We showed that the VaR increase incurred by infrequent MtM is also significant for empirical CDS, both in regular market conditions and in extreme volatility periods. Hedged single risk factor empirical portfolio also exhibit the increase in typical VaR levels (> 99%).

In practical terms, we showed that for a wide range of risk factors that do not meet the RFET, 1-day 99% VaR and 97.5% ES are actually conservative measures of risk. NMRF thus applies an additional risk charge to assets for which risk is already adequately captured in VaR and ES. For such risk factors, NMRF capital is in effect a liquidity risk surcharge, which was not intended to be. We opine that the substantial drop in IMA use observed when implementing the FRTB from Basel 2.5 is largely motivated by NMRF's double counting of risk across the aforementioned class of assets. Therefore, we proposed a simple solution that NMRF charge should be capped at 33% of the VaR (or equivalent ES charge), similar to how the regulatory backtesting multiplier k is capped after 10 exceptions. In parallel and without remediation, as a result of NMRF's

adoption, to artificially pass the RFET, banks will buy data from providers, incurring high upfront costs. In turn, the heightened frequency of the aggregated data would reduce risk figures as no capital multiplier would be applied. The real risk however, would not be modified. Indeed, the RFET only focuses on the last 2 years, while VaR and ES engines are required to contain more than a decade of data, including stressed market periods, generally the great financial crisis. The acquired data permitting the exclusion from NMRF would thus strictly not impact the risk figures reported.

Finally, we also demonstrated that if our starting point is a ‘perfect’ 1-day 99% VaR, yet the backtesting portfolio is updated infrequently, the backtesting performance deteriorates thus leading to increased likelihood of getting capital multipliers.

A APPENDIX

In the previous sections, we presented the analytical and empirical impact of infrequent marking to market (MtM) on Value-at-Risk (VaR). We showed that typically, and contrary to regulators' assumptions, reducing the frequency at which trade data is available actually increases VaR. This results holds for the VaR levels of a single asset, but also for the VaR charge associated to a generic structured correlated portfolio of assets. On a global level, reducing the frequency of price observations increases drastically the magnitude of the price jumps when they happen. Since 1 day-VaR is calibrated over single day returns only, the infrequent large returns obtained through rare MtM thus increase the final VaR figures.

For an observation frequency of T days, the same reasoning applies to M -days VaR, whenever $M < T$. Indeed, N -days VaR essentially relies on the distribution of returns computed over M -consecutive days. In general, M -days VaR scales with $\sqrt{M}$ times the 1-day VaR. As returns constructed from infrequent price measures aggregate T days of price changes at once, the magnitude of these price jumps analogously scales with $\sqrt{T}$ times the 1-day VaR. As long as $M < T$, the M -day VaR is also increased when marking to market is infrequent.

However, for the described scalings in $\sqrt{M}$ and $\sqrt{T}$ to hold, we assumed that subsequent returns are independent identically distributed. In other words, the time series of returns has no autocorrelation. As a toy example of the opposite, let us consider instead an asset which returns $\{R_t\}_{t \in \mathbb{N}}$ exhibit complete price reversal.

$$\forall t \in \mathbb{N}^*, R_t = -R_{t-1} : \text{returns are entirely anti-autocorrelated at lag level 1.}^{19}$$

In this particular example, given any even observation frequency T , the measured returns are null, and thus effective VaR is 0 at all tail levels. The theoretical VaR however, obtained with daily price updates, equals $|R_t|$ at all tail levels. In this extreme example, VaR disappears as a result of infrequent MtM. As a whole, our results are dependent on the autocorrelation level witnessed. For large classes of financial assets, it is well-known that autocorrelation between returns is not negligible. For publicly traded equity stocks in particular, price reversal at different horizons is well documented. At very low horizons, the rise of high-frequency actors has made central the effect of price reversals. At daily horizons too, price reversal is a defining feature of equity prices (Akbas et al., 2022). Therefore, let us in these sections study the impact of infrequent MtM on the VaR levels exhibited by empirical equity and CDS prices, as well as their structure, in order to ensure that the results outlined above in this chapter remain true

¹⁹The space of fully anti-autocorrelated time series at lag 1 is wider than the space series satisfying this condition. Infrequent MtM of any time series in this larger space will however normally yield the same directional impact on VaR.

for empirical financial data.

We firstly extend the validity of our results to equity prices, and then thoroughly expand analog results for CDS, confirming the results from *fig.* 2.3, 2.4.

The third section shows that during various stressed market periods, the most important periods as to the validity of VaR and ES, that infrequent MtM still surcharges the risk. Actually, infrequent MtM's effect on VaR and ES is more pronounced during market downturns. This can be explained by the changing nature of the autocorrelation of financial returns during crises times.

The main section of this chapter argued that analytical hedged portfolios saw an increase in their associated VaR as observed liquidity is reduced. We confirm this result empirically for typical long/short single equity stock portfolios hedged by the average returns.

Throughout this study, we built infrequent returns from only allowing market observations on specific dates, independently of the behaviour of the underlying price process. While this is realistic for many financial assets, which by construction only rarely trade publicly, other mechanisms impact risk factors' observability negatively. In particular, we confirm that VaR and ES increase with lower visibility for volatility based infrequent returns.

In this scenario, we assume that front office traders only report a change in a risk factor's valuation if it has moved enough. That is, trades executed near the previous price are not observed, until the executed price diverges enough. In major markets, to which market-making banks are exposed, this behaviour is widespread. As a general example, one can focus on some options markets, such as single-stock options, where prices and risk factors are driven by options' Δ s or prices. Meanwhile, the quoted variable is an option's implied volatility as a function of its strike. The implied volatility, σ , is only updated relatively rarely by brokers, when their market sentiment changes. As a consequence, even a liquid market with theoretically easily observable data by the firm have little observations reported.

Finally, we show that our empirical results are not a consequence of the methodology applied to construct returns: all main types and rolling windows exhibit qualitatively the same tail distribution behaviours when market observations disappear.

1 EMPIRICAL EQUITIES, THE EFFECTS OF INFREQUENT MTM

Findings presented in this section confirm that for typical VaR levels, infrequent MtM expands equity Value-at-Risk figures, in accordance to the analytical results presented previously in chapter 2.

We have 1155 equity names available from January 2007 onwards. Most of the data is not available until early 2008. Our dataset is composed of the equity names representing the ASX 200, FTSE 100, HSI, NKY 225, S&P 500, SMI and SX5E indexes as of December 2022.²⁰ Data is available at a daily frequency, reporting the aggregated daily price. *Fig. A.8* presents the data's geographical and sectorial repartition.

Equity Availability per Region & Sector



Figure A.8: Available equity names' geographical and sectorial repartition.

From equity prices $\{S_t\}_{t \in \mathbb{N}}$, we build infrequent returns for given MtM observation periods T as for the analytical returns. In the following sections, we extend the framework from absolute empirical returns to relative and log-returns. *Fig. 2.3* presented results for relative returns.

We define infrequent returns with MtM period T as, $\forall t \in \mathbb{N}^*$,

$$\chi_t^T = \begin{cases} \sum_{i=t-(T-1)}^t (S_i - S_{i-1}) & \text{if } t \equiv 0 \pmod{T}, \\ 0 & \text{else.} \end{cases}$$

For varying T , we then compare the VaR at level α for each equity stock with regular and infrequent returns. Firstly we present the distribution of the VaR surcharge incurred by infrequent MtM.

$$\text{For each } (\alpha, T, S), \text{ Surcharge}(\alpha, T, S) = \text{VaR}_{1D}(\alpha, \chi^T) - \text{VaR}_{1D}(\alpha, \chi^1).$$

²⁰Both CDS and equity datasets are obtained through Nomura International Securities' proprietary systems.

To assess infrequent MtM's impact on equities' VaR, we show the distribution of Surcharge (*see fig. A.9*) across equities for standard values of α and T . The higher the VaR level computed, the larger the surcharge from infrequent MtM. For a 99% VaR, we can see that weekly MtM drives the VaR shocks up substantially (*see fig. A.9, Left*). For a lower VaR level, 97.5%, the effect on average null when price data is only accessible every other day (*see fig. A.9, Right*). For a fixed VaR level, the impact of varying MtM frequency on VaR is typically quadratic. VaR surcharges grow for low values of T and subsequently decrease when T reaches a large enough value. *Fig. A.10* presents the average impact of infrequent MtM on 97.5% and 99.5% VaRs. For increasing values of T , are shown the average (and quantiles) Surcharge on equities.

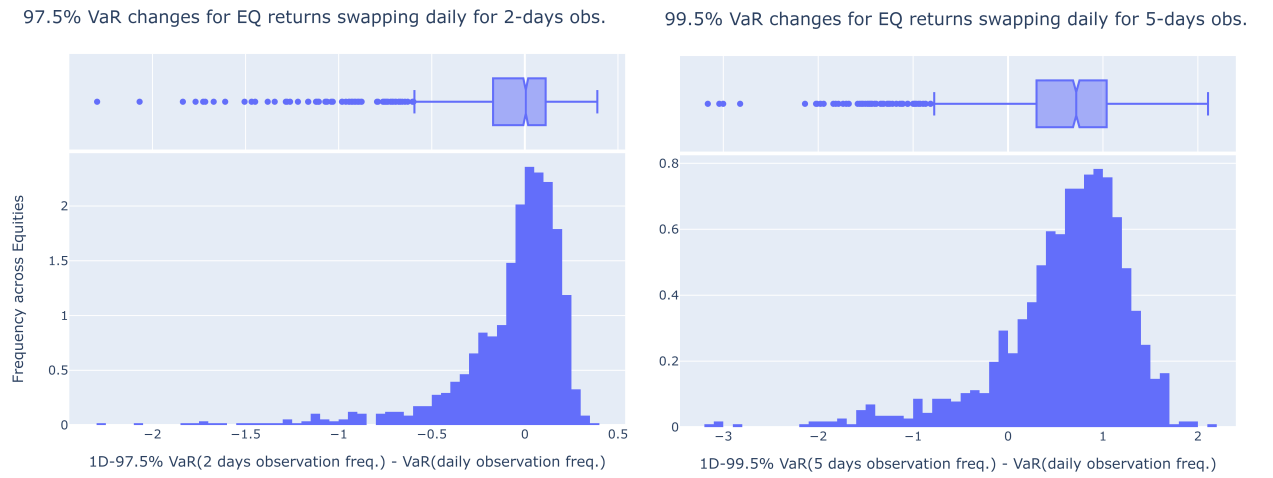


Figure A.9: Distribution across equity names of $(\text{VaR}_{1D}(\alpha, \chi^T) - \text{VaR}_{1D}(\alpha, \chi^1))$. *Left:* for $T = 2$ and $\alpha = 97.5\%$ – *Right:* for $T = 5$ and $\alpha = 99.5\%$

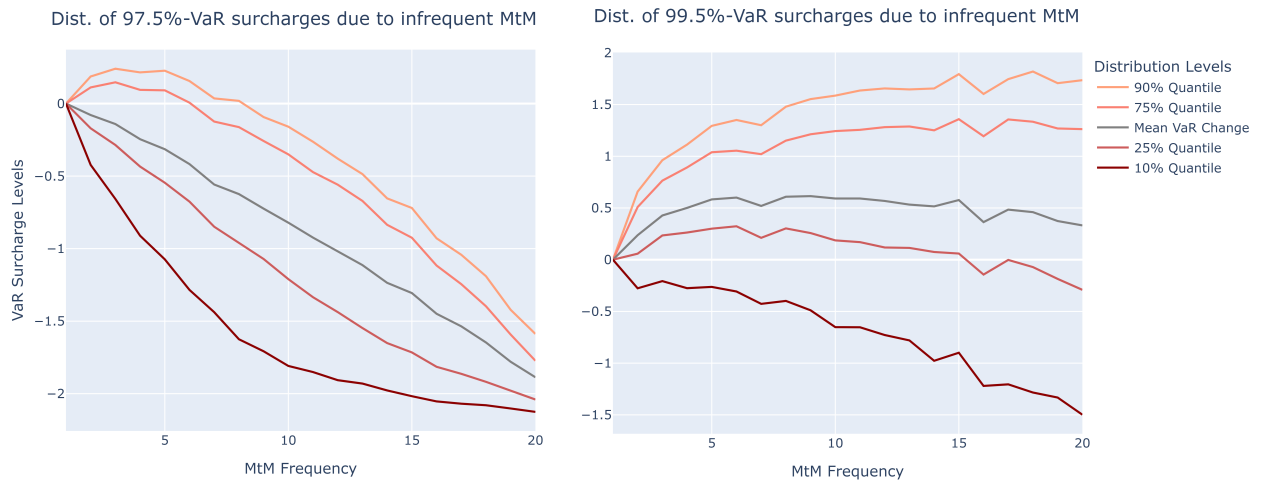


Figure A.10: For varying MtM frequency T , average and quantiles of the surcharges $\{\text{VaR}_{1D}(\alpha, \chi^T) - \text{VaR}_{1D}(\alpha, \chi^1)\}_{\text{Equities } \chi}$. *Left:* $\alpha = 97.5\%$ – *Right:* $\alpha = 99.5\%$. Infrequent MtM reduces 97.5%-VaR, while it increases the 99.5%-VaR. For levels above 99%, infrequent MtM increase VaR for values of T ranging for 2 days to a least two weeks (RFET's threshold is 24 observations a year, one every two weeks). For lower levels of VaR, figures are reduced when daily price observations are not available.

In particular, for VaR levels above and including 99%, we observe a consistent and consequent increase in average VaR levels for MtM frequencies of up to a month. Thus, even with as little as 24 prices observations a year, VaR at usual levels is heightened. For typical VaR levels, the table below reports the maximum MtM frequency for which the average VaR impact on equities is positive. Any smaller MtM frequency on average augments the VaR.

VaR Level	Max. T for which average VaR surcharge is positive
< 97.5%	3 days
99%	14 days
99.5%	38 days
99.9%	3 months+

Table A.2: Maximum observation periods incurring a VaR surcharge for varying α .

As mentioned in the introduction to this appendix, autocorrelation and persistence of price returns strongly affect the structure of infrequent returns. There is a mechanical correlation between the level of autocorrelation in price returns and the volatility of infrequent returns. For equities, we now want to explain the results exhibited in *fig. A.9*, *A.10* and *Table A.2* by understanding further the structure of equity prices time series.

Fig. A.11 (Left) presents the distribution of autocorrelation levels across equities for increasing lag levels, up to a month. As expected, we observe a slight but consistent negative average autocorrelation. Infrequent equity returns for a given MtM frequency T thus have their standard deviation reduced compared to the $\sqrt{T}$ scaling exhibited with *iid* returns. Indeed, since infrequent returns are essentially daily returns summed over the observation frequency, negative autocorrelation shrinks the standard deviation of the summed daily returns, i.e. the infrequent returns' standard deviation. *Fig. A.11, Right* highlights this reduction. For increasing T , the volatility of infrequent returns is reduced compared to the regular daily returns, not only on average but for more than 75% of individual equity names.

As presented in *fig. A.9*, *A.10* and *Table A.2*, the effect of infrequent MtM is still noticeably positive on empirical equities' VaR. The non-zero autocorrelation noticeable for equity time series has a non-negligible impact on the infrequent returns' volatility structure, and while, as said, the directional impact of infrequent MtM is maintained for equity returns, its magnitude is slightly reduced as a result.

We now proceed with an analog analysis for CDS returns. While CDS have a less central place in modern and historical financial markets ([Augustin et al., 2016](#); [Sohnke et al., 2022](#); [Tabassum, 2024](#)), their role during crises periods ([Stulz, 2010](#)), when visible liquidity is the most likely to disappear, make them relevant to our particular study.

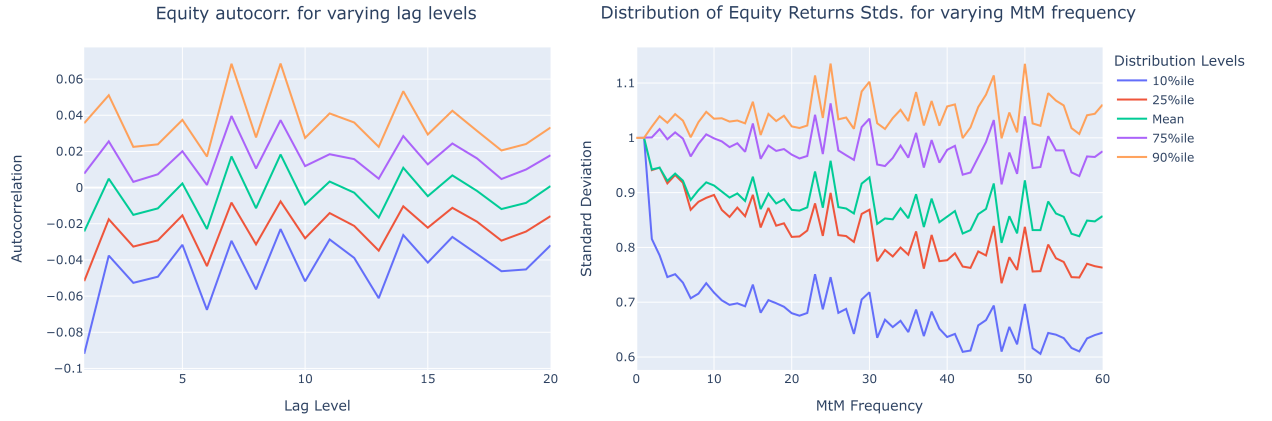


Figure A.11: *Left* – Distribution of Equity Returns autocorrelation, across varying lag levels. Price reversals drive negative autocorrelation at lag $L = 1$. Autocorrelations remain negative on average at a weekly horizon. *Right* – Distribution of the volatility of infrequent equity returns, for varying infrequent observation frequencies. Due to non-zero autocorrelation, infrequent returns' volatility is reduced: subsequent daily returns slightly cancel out: their averaged sum has lower volatility than the original daily returns.

2 EMPIRICAL CDS, THE EFFECTS OF INFREQUENT MTM

Analogously to the previous annex, we want to assess the impact of infrequent MtM on VaR in a wider scope. While the impact over equities was analysed in the previous part, it is important to determine whether the structure of this impact is equity-specific or will remain true in other asset classes. To this end, we replicate the analysis produced previously on Credit Default Swap (CDS) returns. We find that while short-term autocorrelation matters, when only having rare observations, medium-term autocorrelation is also important. Qualitatively, results obtained for equities and CDS exhibit are identical. Quantitatively however, the impact of infrequent MtM is stronger and persists lower in the tail for CDS than for equities.

We have 1138 CDS names available, for which we take the closest to 5 years maturity available for each name.

CDS Availability per Region & Sector



Figure A.12: Available CDS names' geographical and sectorial repartition.

The main global distinction between CDS' behaviour relative to equities is their heightened sensitivity in stressed markets periods. CDS volatility is typically of multiple orders of magnitude higher in the 2008 – 2010 crisis period than in the following decade (*see fig. A.13, A.14 (Top)*). This peculiarity is non-negligible regarding VaR and infrequent MtM. As tail level VaR figures are driven by stressed market periods returns, calibrating over CDS returns is known to yield conservative metrics of risk.

Surprisingly however, while CDS returns' volatilities experience a regime shift after the 2008 – 2011 period, the distribution of CDS autocorrelations do not markedly variate across the whole time-period (*see fig. A.14, Bottom*). Since autocorrelation levels are the main driver of infrequent MtM impact to VaR, global CDS returns volatility distribution does not materially impact infrequent MtM VaR surcharges.

We observed that equities are generally negatively autocorrelated at short time-scales, and that the infrequent MtM incurred VaR surcharge is reduced as autocorrelation becomes negative. Conversely to the previous argument, CDS autocorrelations are visibly different from equities.

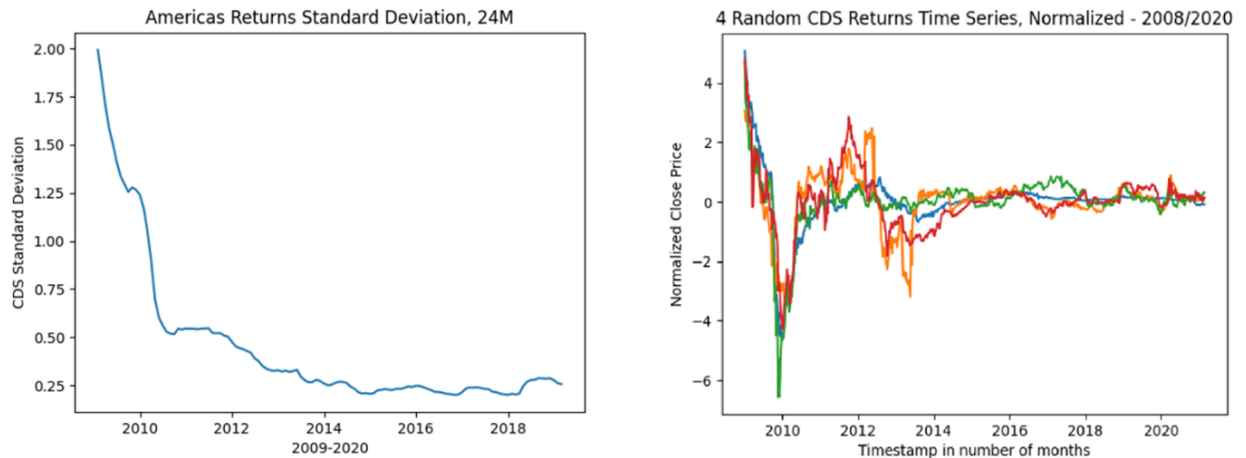


Figure A.13: CDS yearly rolling absolute returns, clustering of volatility around the 2008 stress period, presented for US & Canada region. Compared to equities, VaR and autocorrelation levels for CDS will largely be driven by data across the 2008 – 2010 period and show little sensitivity to data in the 2011 – 2020 range.

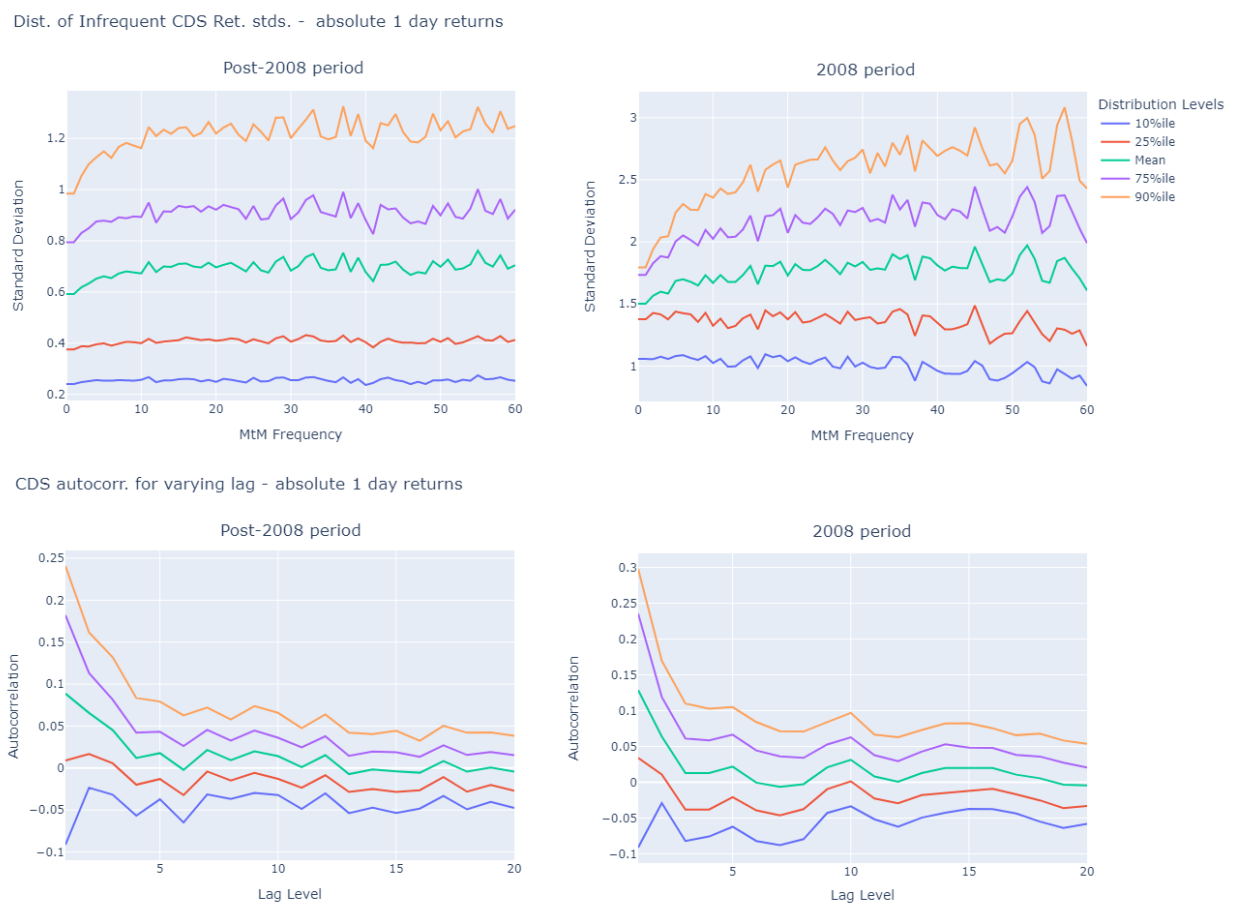


Figure A.14: CDS daily absolute returns. Volatilities are substantively higher around 2008 (*Top-Left*) than later (*Top-Right*). Notice the axis range difference. Values where MtM is 0 (daily synchronous returns) represent volatility of regular CDS returns. *Bottom*: Distribution of CDS autocorrelation for varying lag. We see no sizable difference between the two periods.

The set of histograms above (*see fig. A.15*) presents the impact of infrequent MtM

on CDS VaRs, at various tail levels and various price observation frequencies, ranging respectively from 97.5% to 99.9% and 2-days to monthly. Qualitative observations are similar to the equities' results. The lower the VaR level, the more infrequent MtM understates the VaR. And the least frequent the updates, the bigger the VaR shocks reductions are.

However, across all VaR levels, we observe that independently of the frequency chosen, scarce observations increase CDS VaR more than equity VaR. For equity, most typical ranges of frequencies and tail levels yielded VaR increases. For CDS, many (frequency, tail level) couples see on average a VaR reduction instead, and when even in the equity setup, infrequent MtM reduced VaR, in the CDS setting, the VaR is diminished further. These results confirm that not only the short-term autocorrelation matters, but the medium term one does too. The high positive overnight CDS autocorrelation still matters when considering an update frequency of 2 days. In this scenario, infrequent CDS is always higher than its regular counterpart (*see fig. A.15, Center Left*).

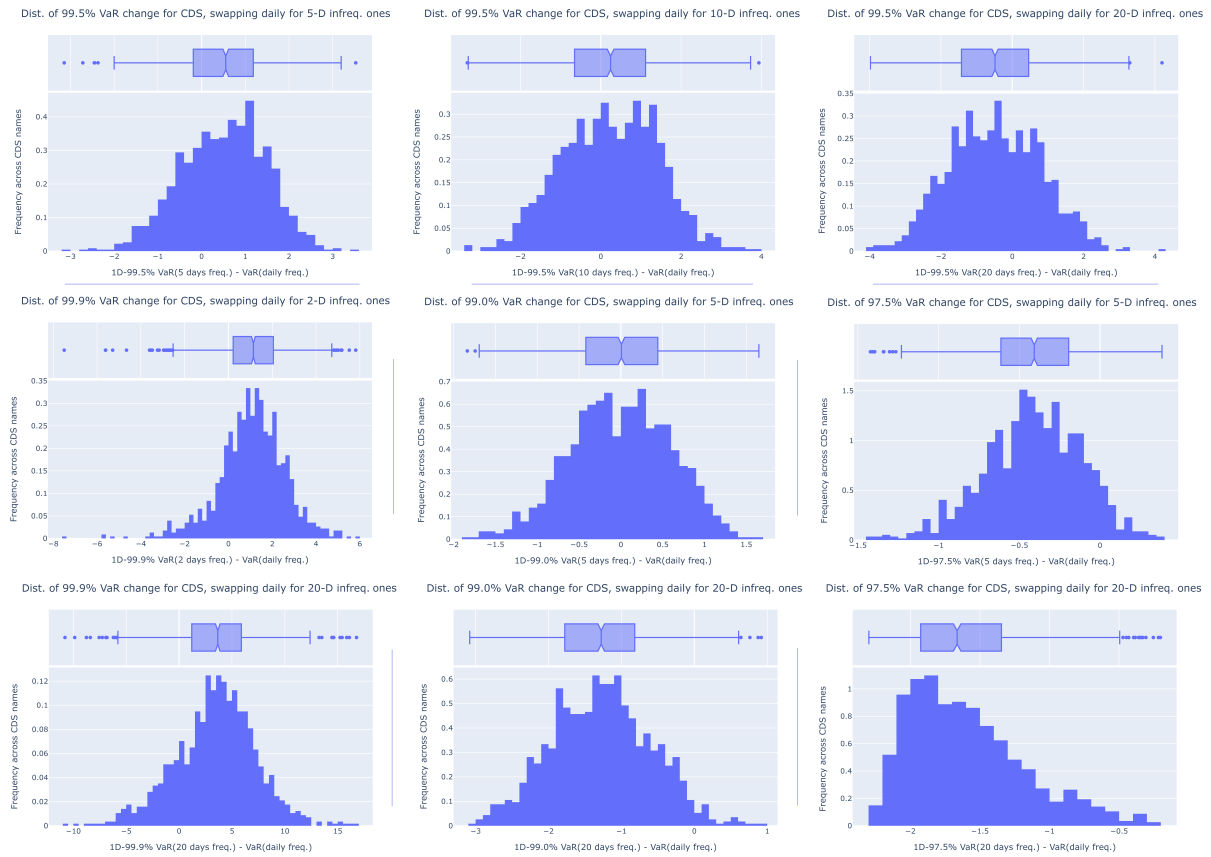


Figure A.15: CDS daily absolute returns. Histograms of impact of infrequent MtM on VaR, for different MtM frequencies and VaR levels. *Top:* 99.5% VaR, with weekly, bi-monthly and monthly observations. *Left:* 99.9% VaR, 2-days and monthly observation frequencies. *Centre:* 99% VaR, weekly and monthly observation frequencies. *Right:* 97.5% VaR, weekly and monthly observation frequencies.

3 STRESSED MARKET PERIODS, 2008 FINANCIAL CRISIS AND COVID

It is tantamount for banks to capture risk adequately in stressed market conditions. By extension, the well functioning of VaR and ES during crises periods, precisely when liquidity can disappear and price updates rarify, is crucial. In this section, we confirm numerically on CDS and equities that infrequent MtM during high volatility periods also increases VaR at typical levels. Actually, the impact on VaR is larger in crisis time than across the whole timeframe.

During the 15 years of data available, two stressed markets episodes are identifiable: March 2020 and September 2008, the Covid and Financial Crisis periods respectively.

We follow a process analogous to the previous section to assess the impact of infrequent MtM on equity and CDS VaR. We however restrict our dataset first to 2008 and then to 2020. During financial downturns, returns' reversals are globally diminished as the medium term market factor is falling sharply (respectively for CDS, growing sharply). Thus, the negative average autocorrelation observed in ordinary market conditions disappears: equity and CDS returns during stressed periods are globally not significantly autocorrelated (*see fig. A.16*). This modification of the empirical autocorrelation function (ACF) is the main driver of the larger infrequent MtM surcharge on risk.

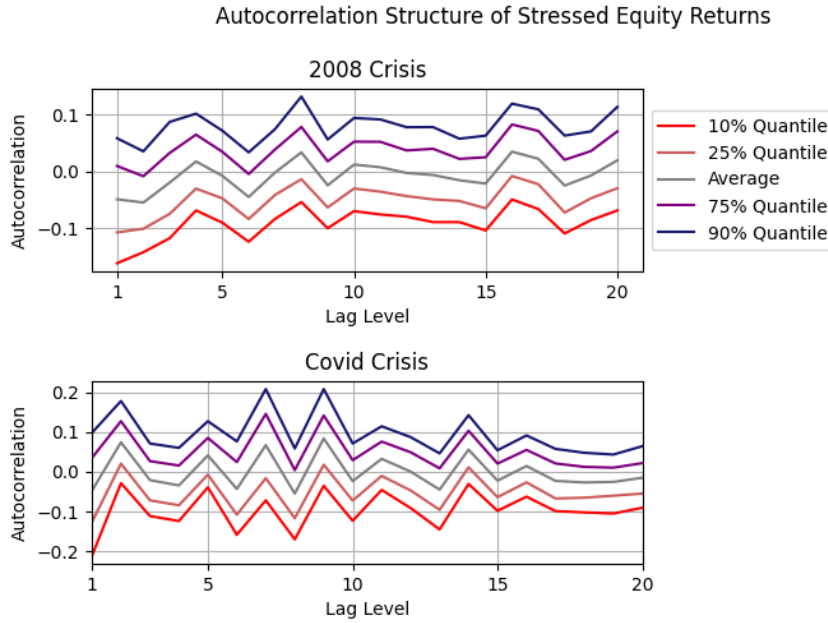


Figure A.16: Distribution of Autocorrelation Levels observed in stressed equity returns for varying lag levels. Average is near 0: the slightly negative autocorrelation subsides in crisis time. Analogous results are observed with CDS data. *Up: 2008 – Down: 2020.*

Fig. A.17 shows in particular that at the most important 99%-VaR level, during crises periods, the average surcharge resulting from infrequent MtM persists for up to



Figure A.17: *Solid:* 2008 Crisis – *Dashed:* Covid Crisis. For varying MtM frequency, distribution of 99%-VaR surcharges across equities (*Left* - absolute surcharge) and CDS (*Right* - relative surcharge). Up to monthly price updates, the impact of infrequent MtM on VaR is significant and positive.

monthly observation frequencies, for both CDS and equities. Across the whole 2007 – 2022 period, the average surcharge is positive for observation frequencies smaller than 3 weeks (15 days) only.

As shown in *Table A.3* the infrequent MtM VaR surcharge persisting longer for crises periods is also present for VaR levels ranging from 97.5% to 99.5%, ensuring that our findings remain true for 97.5% ES too.

VaR Level	Max. T for which average VaR surcharge is positive		
	2008 Crisis	Covid Crisis	2007-2022
< 95%	3 days	3 days	3 days
97.5%	8 days	6 days	3 days
99%	20 days	22 days	14 days
99.5%	34 days	40 days	38 days
99.9%	3 months+	3 months+	3 months+

Table A.3: Maximum observation periods incurring a VaR surcharge for varying α during stressed market periods. The effects of infrequent MtM persist longer during periods of high volatility.

Our analysis has shown that during all stressed market periods considered, exhibiting persistent directional wrong-way market movements, the increase VaR at all usual tail levels is still increased (and by extension, ES too) when observations are scarce.

Despite the consistency of these results, we want to acknowledge that idiosyncratic events could yield a VaR reduction when observations are scarce. In particular, if in between two consecutive price observations, a “flash crash” happens and then entirely subsides, infrequent MtM will diminish VaR.

Yet, some swift and substantial market moves are included in our dataset, such as the Covid stock sell-off in March 2020, during infrequent MtM still results in higher VaR figures.

4 VAR OF INDIVIDUAL EQUITIES HEDGED AGAINST INDEX

As in our previous analytical derivations, we want to extend our empirical assessment of infrequent MtM's impact on VaR to a portfolio comprising more than a single stock (see Section 2). We consider a single-stock portfolio, hedged against average market returns. For a given individual equity name, this portfolio is comprised of one share of this equity name and shorts the average stock returns. The amount shorted is chosen so that the exposition to both sides of this trade are identical. We construct all 1155 such equity-based portfolios and, as in the previous Appendix sections, present the distribution of VaR surcharges entailed by infrequent MtM (see fig. A.18).

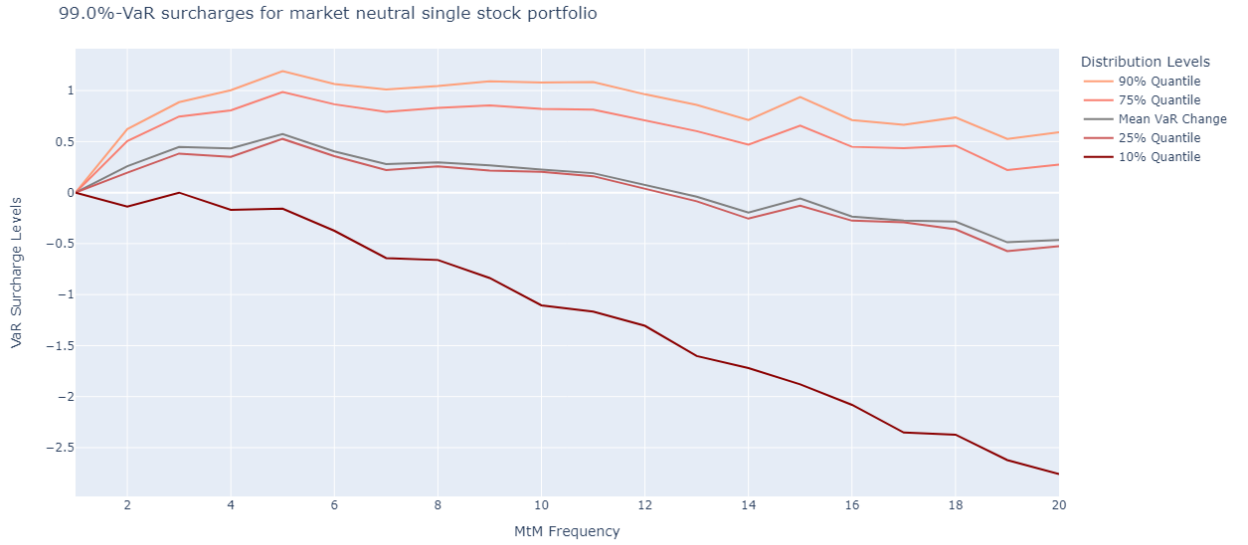


Figure A.18: Distribution of impacts from infrequent MtM on single stock portfolios hedged by the market average. The average surcharge remains positive even for bi-weekly updates, analogously to the simple single stock portfolio. However, infrequent MtM reduces VaR drastically for a small proportion of available equities (10% quantile).

At VaR levels ranging from 85% to 99.9% (see Table A.4), the effect of price data unavailability is not only qualitatively but also quantitatively similar for a single stock exposure than for a hedged portfolio: low market observability increases VaR at all typical tail levels, entailing that ES is also enlarged for straightforward hedged empirical portfolios.

VaR Level	Portfolio VaR Level	
	Hedged Single Stock	Single Stock
< 97.5%	2 days	3 days
99%	12 days	14 days
99.5%	37 days	38 days
99.9%	3 months+	3 months+

Table A.4: Maximum observation periods incurring a VaR surcharge on a structured portfolio for varying tail level α .

5 THRESHOLD BASED INFREQUENT RETURNS

In the previous sections, infrequent returns were built considering the risk factor's market value was observed regularly every T days. For wide classes of risk factors, this definition is natural as it models illiquid risk factors, rarely traded, thus observed infrequently, despite large volumes being held by market participants. For other risk factors however, low observation frequency can be driven by other effects. We find that for other forms of infrequent returns, VaR and ES are drastically increased when liquidity observations become scarce.

Front office traders will only update a risk factor's price shall it have moved enough. Indeed, the high frequency of trades makes it prohibitive to manually report each. Any number of trades happening within a band of the original price P_0 will not be reported. Only when the price P_t of some transaction exits the band $[P_0 - \epsilon, P_0 + \epsilon]$ will the price formally be updated. The value of ϵ is linked to the distribution of the risk factor's returns, e.g. the price change is reported only once a $0.5\text{-}\sigma$ change has happened.

Let the full unobserved price series be $\{P_t\}_{t \in \mathbb{N}}$. As in the main part of this chapter, we assume that returns are normal and *iid*. Given a threshold $\epsilon \geq 0$, we define the infrequently dated price series be $\{P_t^\epsilon\}_{t \in \mathbb{N}}$ iteratively as:

$$P_0^\epsilon = P_0,$$

$$P_t^\epsilon = \begin{cases} P_t, & \text{if } |P_t - P_{t-1}^\epsilon| > \epsilon, \\ 0, & \text{else.} \end{cases}$$

We then define the infrequently observed returns as:

$$\forall t \in \mathbb{N}, R_t^\epsilon := \begin{cases} 0, & \text{if } t = 0, \\ P_t^\epsilon - P_{t-1}^\epsilon, & \text{else.} \end{cases}$$

The threshold's value ϵ depends on the underlying risk factor's returns volatility. In practice, given a distribution for the risk factor's returns, $\epsilon(T)$ represents the average expected absolute return in a chosen time period T (T is expressed in traded days). That is, we assume traders only report a price change once the risk factor's price has

changed by an amount $\epsilon(T)$ that equals the expected absolute price change during any period of length T .

Assuming the risk factor's daily returns follow a normal distribution $N(0, 1)$, its absolute daily returns follow a half-normal distribution of parameter 1. The average value of the absolute returns after T traded days then is the average value of a half-normal distribution of parameter $(T - 1)$:

$$\forall t \in \mathbb{N}, \epsilon(T) = \sqrt{(T - 1) \frac{2}{\pi}},$$

For empirical CDS and normal returns, *Fig. A.19* presents the impact of infrequent market observations on both VaR and ES. It presents the results for the previous “periodically” observed returns and for the aforementioned “threshold” based infrequent returns.

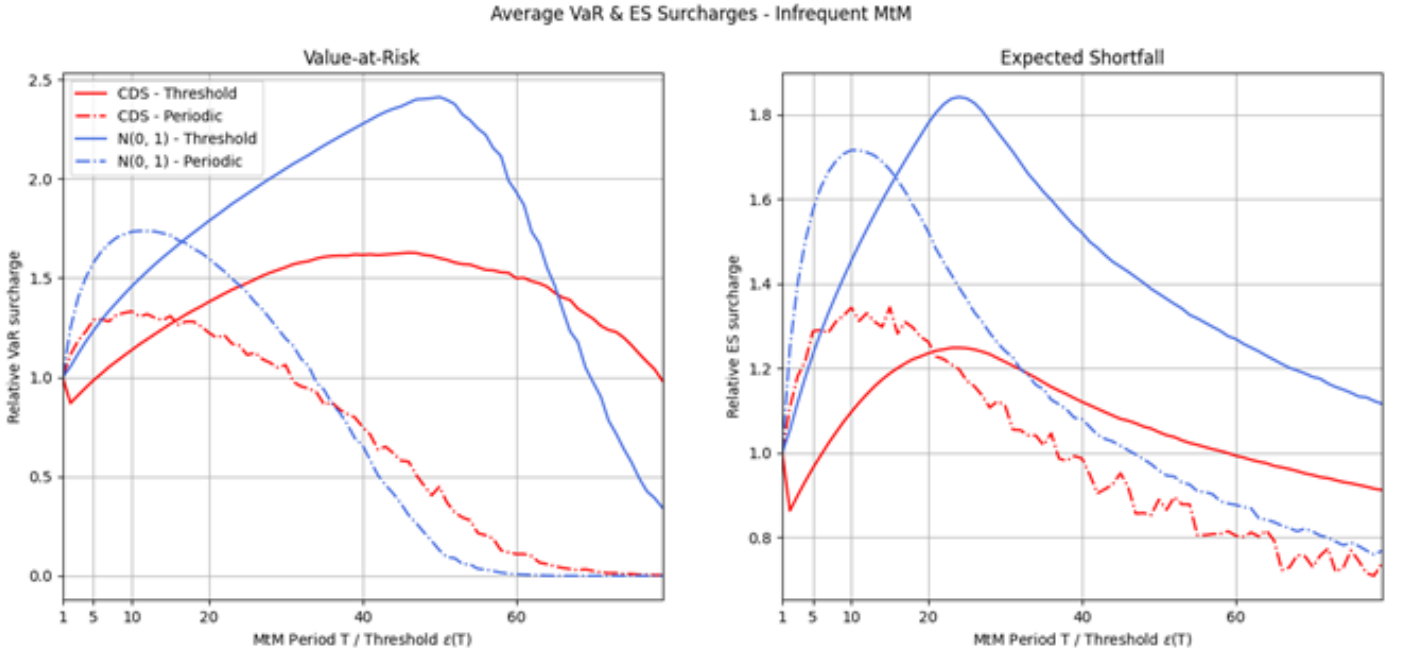


Figure A.19: 99%-VaR (Left) and 97.5%-ES (Right) relative surcharges for normal (Blue) and empirical CDS (Red) returns, for varying threshold $\epsilon(T)$ & index T . For a given T , the two types of infrequent returns are presented: “threshold” (Solid) and “periodic” (Dashed) based. For a given T , the corresponding threshold $\epsilon(T)$ is $\sqrt{T - 1} \cdot \sqrt{2/\pi}$. The maximal surcharge is significantly higher for infrequent “threshold” returns than for “periodic” ones. The largest VaR / ES surcharges also happen for larger values of T for “threshold” based infrequent returns. For empirical CDS returns and low thresholds, ES is reduced (Right, Solid Red).

Fig. A.19 shows that “threshold” based infrequent returns have larger maximal VaR / ES surcharges. This maximal surcharge happens for large $\epsilon(T)$: $T > 25$. Indeed, the surcharge for “threshold” infrequent returns is larger but also shifted to the right: low values of T carry a smaller surcharge, which can even be negative for fat-tailed skewed CDS returns.

Given T , CDS “threshold” returns actually have lower VaR for all levels α such that $\text{VaR}_{\epsilon(1)}(\alpha) > \epsilon(T)$:

$$\text{If } \text{VaR}_{\epsilon(1)}(\alpha) > \epsilon(T), \text{ then } \text{VaR}_{\epsilon(1)}(\alpha) \geq \text{VaR}_{\epsilon(T)}(\alpha). \quad (1)$$

Where $\text{VaR}_{\epsilon(T)}(\alpha)$ is the α -VaR of “ $\epsilon(T)$ -threshold” based infrequent returns.

(1) is the main driver of the ES reduction for low values of T . *Table A.5* shows the resulting VaR reduction for low T and $\alpha = 99.5\%$, 97.5% as well as the resulting reduction in 97.5%-ES for CDS returns.

As T increases, $\epsilon(T)$ grows, and (1), $\text{VaR}_{\epsilon(1)}(\alpha) > \epsilon(T)$, is never satisfied: ES is increased too. For normal returns, since they are not fat tailed (*see Table A.6*), VaR values for daily observations are lower, and the condition $\text{VaR}_{\epsilon(1)}(\alpha) > \epsilon(T)$ is not met even for low values of T .

Period	“Threshold” CDS Returns VaR / ES Surcharge		
T	99.5% VaR	97.5% VaR	97.5% ES
1	1	1	1
2	0.85	0.9	0.86
3	0.88	0.975	0.9
4	0.9	1.04	0.93
5	0.93	1.09	0.96
6	0.95	1.14	0.99
7	0.97	1.19	1.02
8	1.0	1.22	1.05

Table A.5: VaR / ES surcharges for small varying $\epsilon(T)$ at different tail levels. We observe a decaying reduction in 99.5%-VaR and 97.5%-VaR. The 99.5%-VaR reduction persists longer as (1) remains satisfied longer further in the tail.

	Return Type	
	CDS	Normal
VaR 97.5%	1.61	1.96
VaR 99%	2.39	2.32
VaR 99.5%	3.13	2.57
ES 97.5%	2.63	2.34

Table A.6: VaR and ES values for original normalized CDS and normal returns. CDS’s fat tails yield higher VaRs for $\alpha > 99\%$ and lower ones below.

On (1): let us focus on a day t for which a very large tail return χ_t is observed. Since χ_t is large, in general we have $N(t+1) = t$. We also have that on average $N(t) < t-1$. In other words, in the infrequent “threshold” returns, χ_t is combined with returns from previous days. CDS and more largely financial returns are negatively autocorrelated (*see Appendix section 2*): the days preceding day t are on average of the opposite sign.

As we assumed $\chi_t > \epsilon(T)$, $R_t^\epsilon > 0$ still holds. However on average we now mechanically have $\chi_t > R_t^\epsilon$. As we chose χ_t to represent $\text{VaR}_{\epsilon(1)}(\alpha)$ for some α , we get (1).

For normal returns, as they exhibit no fat tails, we already have that $\text{VaR}_{\epsilon(1)}(\alpha) > \epsilon(T)$ generally does not hold. Since normal returns are by definition not autocorrelated, shall the aforementioned condition hold, we would also not observe the VaR reduction from (1). Thus, the 97.5%-ES reduction in *Fig. A.19* for CDS happens as CDS returns are both fat tailed and anti-correlated.

As T and $\epsilon(T)$ increase, $N(t+1) = t$ does not hold as $\chi_t \leq \epsilon(T)$. Thus $R_{N(t+1)}^{\epsilon(T)}$ combines χ_t as well as other tail values, and we get the consistent VaR and ES increases observed for larger T in *fig. A.19* and *Table A.5*.

6 THE IMPACT OF CHANGING THE RETURNS' STRUCTURE

In this appendix, we used daily absolute returns. We previously presented results for relative returns (*see fig. 2.2, 2.4*) and for absolute returns above in this appendix. While to VaR it is natural to use daily returns in a chosen specific form, log, absolute or relative, it is not usual for many other types of risk measures. Other risk measures which however still centre on measuring the tails distributions of past returns in varying fashions. More precisely, widely used risk measures such as Incremental Risk Charge (IRC) or the calibration of FX shocks can rely on yearly overlapping returns. As these measures are central to calibrate risk on private and sovereign bond portfolios, arguably the most important measures to investment banks, it is important for the relevance of this study to ensure that our results remain true for other forms of returns.

Even for equities' analysis, the standard is not absolute returns. We presented results above for absolute returns, as our analytical derivations are not tractable starting with relative or log-returns. In this annex, we provide numerical confirmation that our results remain valid for relative & log returns. Also, lengthening the returns' window only increases the effects of infrequent MtM on tail returns.

As we showed that the impact of infrequent MtM is directly determined by the autocorrelation structure of returns, understanding the impact on autocorrelation of modifying the returns' type will contribute towards generalizing the analytical results established in previous sections. Let us then compare results obtained for different return types:

For given equity prices $\{P_t\}_{t \in \mathbb{N}}$, the rolling daily log-returns are:

$$\forall t \geq 1, S_t = \log \left(\frac{P_t}{P_{t-1}} \right).$$

Analogously, the rolling daily relative returns are: $\forall t \geq 1, S_t = \frac{P_t}{P_{t-1}} - 1$.

And as used previously, the rolling daily absolute returns are: $\forall t \geq 1, S_t = P_t - P_{t-1}$.

We find that there is no impact from swapping the return type between the three described above. Both autocorrelation structures and VaR impacts from infrequent MtM are identical (*see fig. A.20, A.21, A.22*).

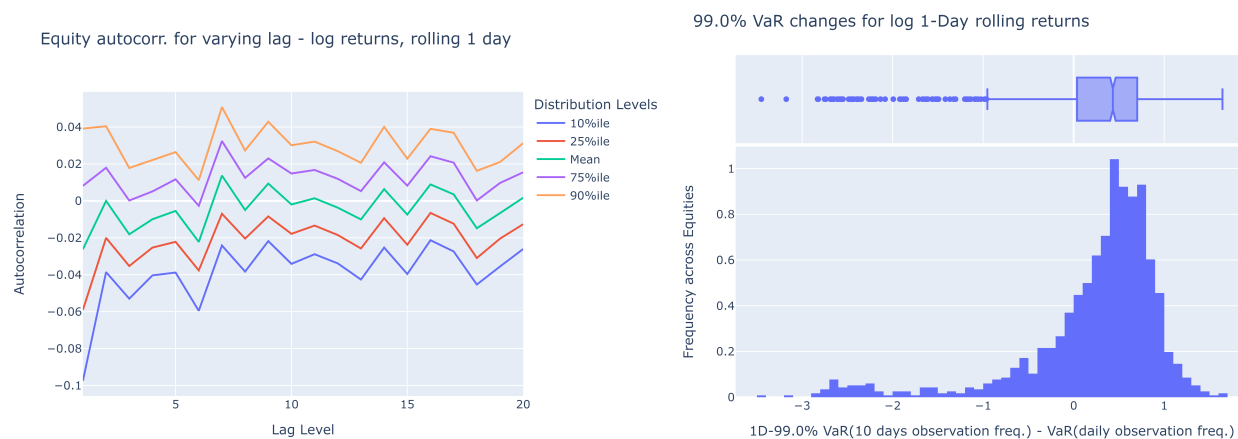


Figure A.20: Equity daily log-returns – Left: Autocorrelation for increasing lag, from 1 to 20 days. *Right:* Distribution of the impact on 1-Day 99% VaR of bi-monthly price updates.

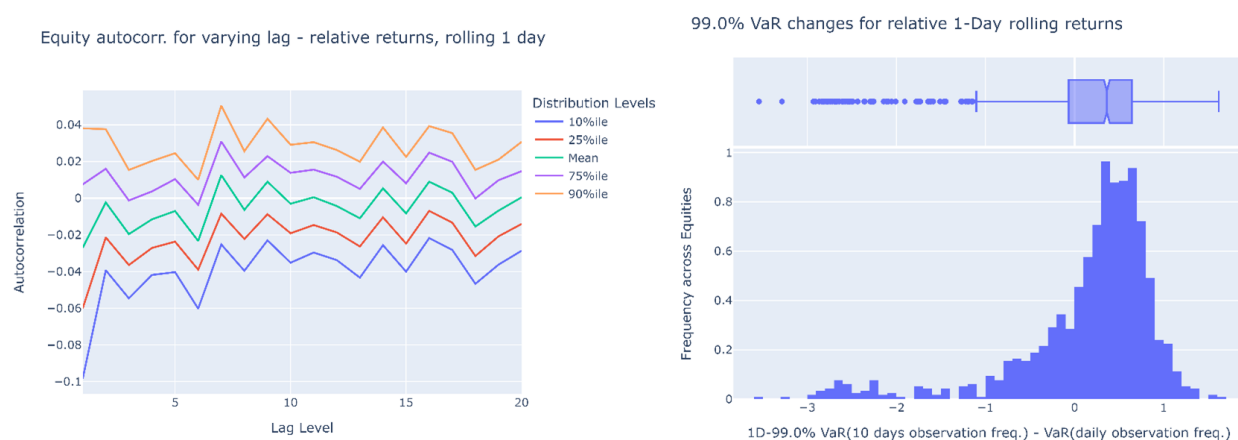


Figure A.21: Equity daily relative returns – Left: Autocorrelation for increasing lag, from 1 to 20 days. *Right:* Distribution of the impact on 1-Day 99% VaR of bi-monthly price updates.

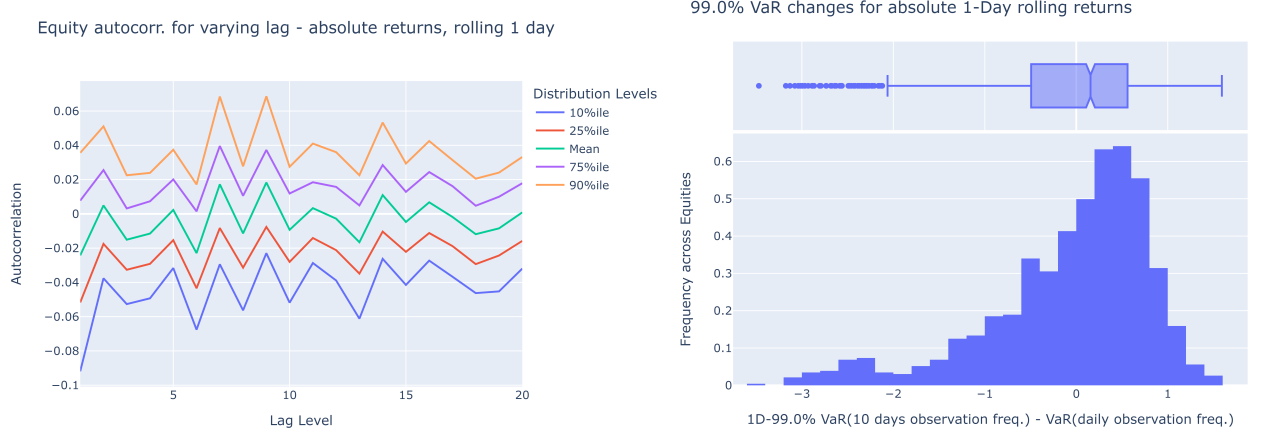


Figure A.22: Equity daily absolute returns – Left: Autocorrelation for increasing lag, from 1 to 20 days. Right: Distribution of the impact on 1-Day 99% VaR of bi-monthly price updates.

We also carry the same analysis on CDS returns (*see fig. A.23*) and obtain the identical conclusion: changing the return type does not matter towards the effect of infrequent MtM on VaR, and by extension ES.

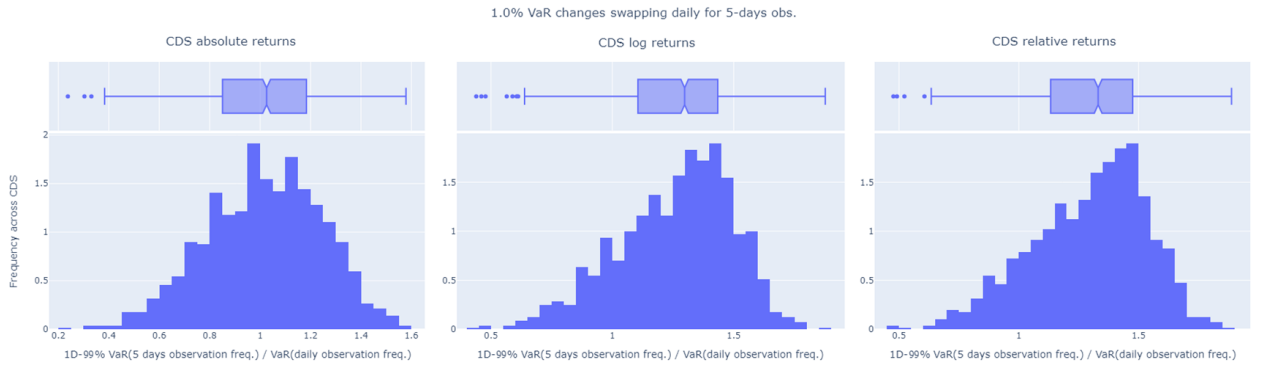


Figure A.23: CDS returns – Distribution of the impact on 1-Day 1% VaR of bi-monthly price updates. Left: Absolute Returns. Center: Log>Returns. Right: Relative Returns.

As there is no impact from swapping return type in our returns' construction, we can extend the qualitative results obtained for absolute returns in this study to relative and log-returns. Next, we perform an analogous comparison between various return windows, from the daily one presented in the main part of the study, to the maximum yearly rolling window, and compare the effects on their volatility structures and VaR levels when only infrequent MtM is available.

Given a time horizon for the returns' window H , the H -days rolling returns are defined as:

$$\forall t \geq H, S_t^H = P_t - P_{t-H}.$$

Specifically, we compare returns with rolling window horizons H in $\{1, 5, 20, 252\}$ representing respectively daily, weekly, monthly and yearly rolling returns. Naturally,

using larger return windows increasing the autocorrelation in returns at low lags (lags smaller than the return window), as subsequent returns are impacted by the same daily price levels:

$$\forall t \geq H, S_t^H = \sum_{i=t-H+1}^t (P_i - P_{i-1}).$$

We automatically have that $(P_t - P_{t-1})$ impacts equally all elements of $\{S_u^H\}_{u \in [t, t+H-1]}$. Thus, autocorrelation is artificially heightened over the first H lag levels (*see fig. A.24*). We instead focus on analysing the volatility structure of infrequent returns using different return windows. The volatility of infrequent returns for varying observation period T increases as T approaches H . As T overtakes H , the volatility structure stabilizes. Therefore, the larger H is, the larger the volatility of infrequent returns (when $H < T$) (*see fig. A.25*). These significantly elevated volatility levels drive a substantial increase in infrequent MtM VaR. The returns' horizon window amplifies the VaR surcharge incurred by infrequent MtM (*see fig. A.26*).

Summarizing the results of this annex, the results presented in the main part of this document not only stand independently of the return type used but are also stronger when swapping the 1 day return window to a longer one. For a more comprehensive explanation of the drivers of this increase, see the next section.

Equity Autocorrelation at Different Lag Levels - absolute returns, rolling 5 day

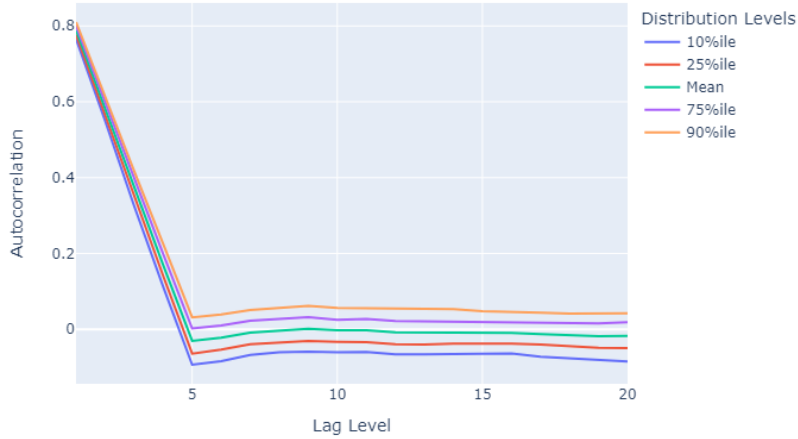


Figure A.24: Distribution of autocorrelations, rolling weekly returns. Over lag levels of 1 to 5 days, the autocorrelation is artificially increased. It decreases linearly to daily returns autocorrelation levels as the common dependency between original and lagged returns disappears. It however smooths the autocorrelation structure for lag larger than 5 days.

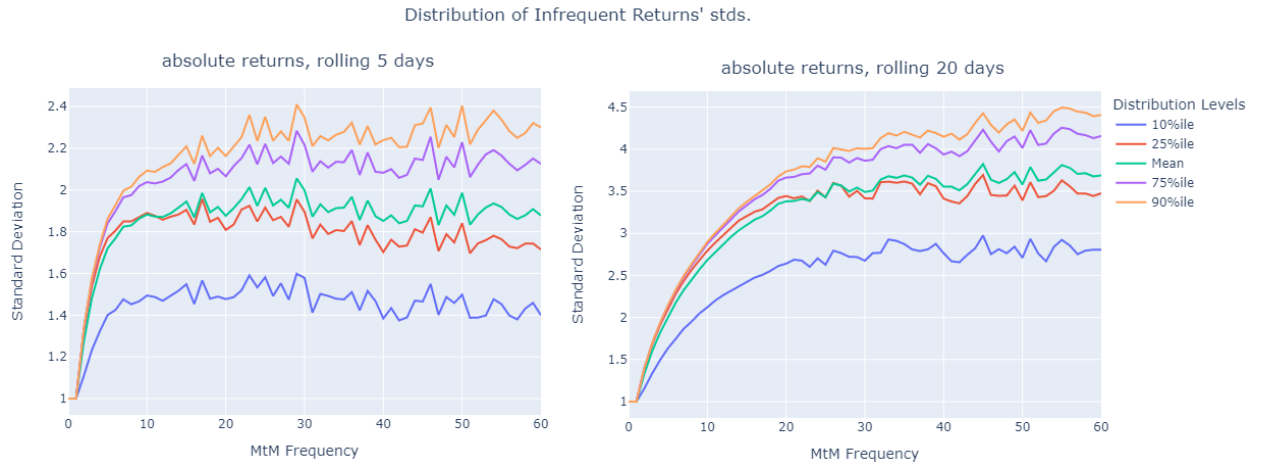


Figure A.25: Distribution of infrequent returns standard deviations, using weekly (*Left*) and monthly (*Right*) rolling returns. The returns' volatility is higher in both cases than with daily infrequent returns.

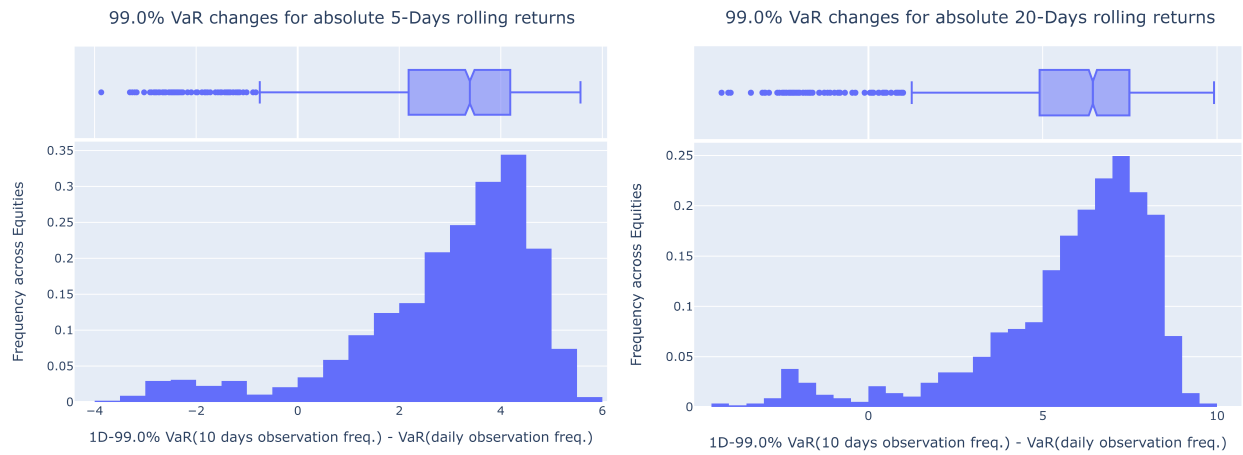


Figure A.26: VaR surcharge incurred by infrequent MtM, on weekly rolling returns (*Left*) and monthly rolling returns (*Right*). Both are substantively higher than the infrequent MtM surcharge on daily returns.

7 THE IMPACT OF THE RETURNS' WINDOW ON THEIR CORRELATION STRUCTURE

We quickly present the influence of swapping return window length on the structure of pairwise correlations. We introduce pairwise correlations, as within any realistic setting for a large financial institution, VaR for multiple assets would be considered, and not only for one individual stock. We showed that changing return window modifies the infrequent MtM's impact on VaR. The longer the rolling window considered, the more significant the influence on VaR is.

We can see below (*see fig. A.27*) that in parallel to the previous points, using longer time windows for returns' computation drives pairwise correlations up considerably. Days of large returns are therefore more likely to be the same for many equity names at once. When considering the VaR generated by all these names at once, the tail returns are hence generated by individual tail returns of many equities at once, since

they are highly correlated.

Thus, using longer spanning return windows, the aggregated VaR across a portfolio would be increased even further than shown above for individual equities.

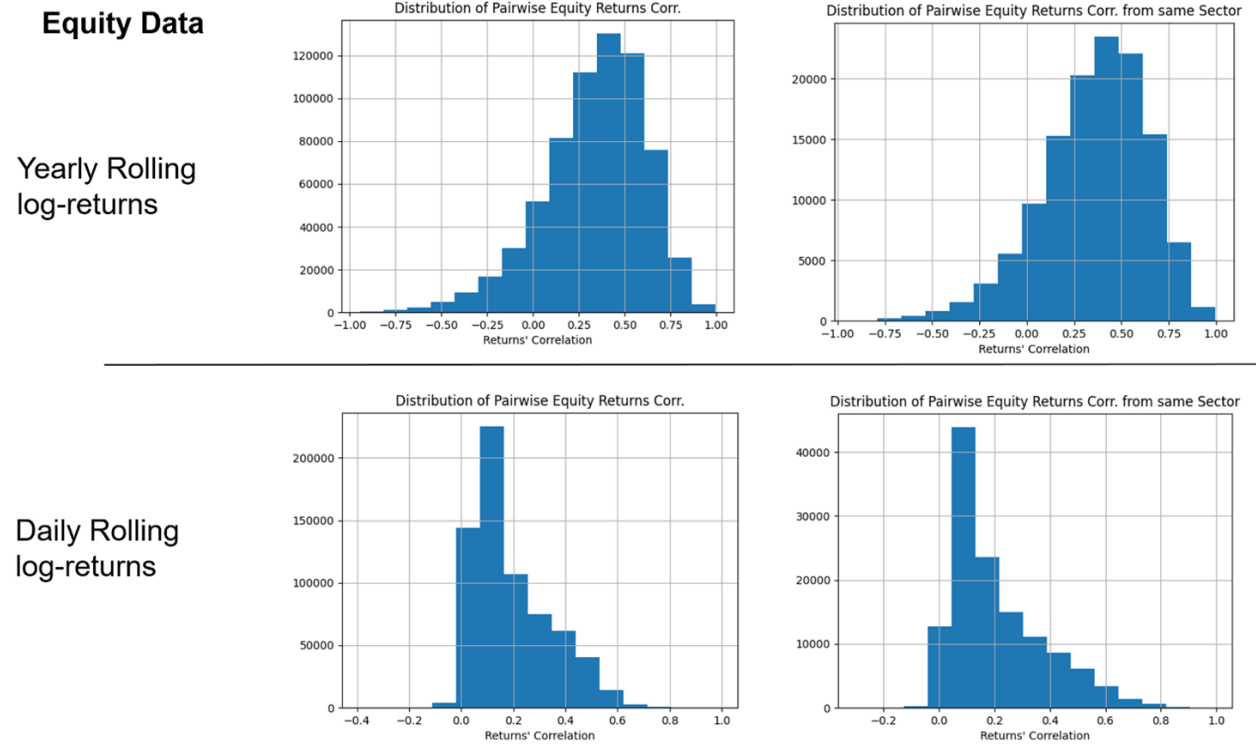


Figure A.27: For yearly (*Top*) and daily (*Bottom*) yearly log-equity-returns, we present the distribution of pairwise correlations, both among general equity names (*Left*) and then only comparing equity names from the same sector (*Right*). The magnitude of the change as well as the change in the distribution's structure demonstrate the inevitable impact of choosing the return type.

“If 80 percent of our work is data preparation, then ensuring data quality is the important work of a machine learning team.”

Andrew Ng

3

Machine Learning for Sovereign Credit Grade Prediction

Content presented in this Chapter was developed within Nomura International Securities. It partly relies on third party data, available for purchase. It also uses Nomura International Securities proprietary data, unavailable publicly.

Parts of this Chapter are redacted, following Nomura International Securities Compliance guidelines.

Following thesis submission guidelines at Imperial College London, I acknowledge that part of this work, written by myself, was presented in a conference at the time of the thesis submission.

According to the Bank for International Settlements, investment banks across the globe now hold more than a trillion US dollars' worth of sovereign bonds¹. Assigned by credit rating agencies, sovereign credit grades' movements largely drive these positions' risk and volatility. Their prediction is thus crucial to better manage sovereign credit portfolios. Available economic data is unreliable, scarce and restricted by the number of sovereign entities with government debt. Established historical models used to set and predict sovereign credit grades center on restrictive linear modelling. Modern machine learning techniques require numerous parameters and vast datasets to converge over noisy data.

Through thorough data processing, motivated by both economic and statistical insight, we put forward classifiers which novelly demonstrate that small neural networks and random forests can calibrate accurately even on poor, lopsided macroeconomic datasets. Their accuracy outperforms all known industry and published implementations, both linear and neural network based. Our results are cross-validated on carefully isolated data, both temporally and geographically, replicating a production situation. This architecture is the first capable of predicting sovereign credit grades' evolution with accuracy high enough to surpass the intuitive and easily implemented "*constant classifier*" benchmark. Unlocking the use of modern statistical methods to small, low quality economic and financial datasets, our predictors underscore the core importance of extensive problem specific data pre-processing in machine learning for macroeconomic classification.

¹Source: Bank for International Settlements, International debt securities publications (BIS-compiled), 2023

INTRODUCTION

Financial institutions' portfolios are heavily exposed to sovereign debt. A financial institution such as Nomura holds sovereign bonds, in particular Japanese government bonds, through multiple branches of its business (Alletzhauser, 1990). First, Nomura is one of the only 223 companies allowed to bid for newly issued bonds by the Japanese government (as of December 2023) (Coggan, 2009; MoF, 2023a). As such, bidders for new bond issuances act as brokers but also in practice as shadow bankers (Agirman et al., 2013). That is, bond market makers buy a share of newly issued bonds higher than the total current demand for the bonds orders they have received. This excess supply in turn allows the firm to act as a liquidity provider for these most recent bonds. The total yearly value for these new bonds is large. For the Japanese government in 2023 alone, new issuance will reach 9.3 Trillion Yen, while a seventh of the already outstanding 206 Trillion Yen will be renewed through new issuances (MoF, 2023b).

In practice, as continuous actors in sovereign bond markets, investment banks also keep positive inventories of fluctuating size in most categories of sovereign bonds they frequently trade. There can be dozens of such categories for a single government, as maturity, interest rate and borrowing purpose all need to be in line for two bonds to be equivalent (MoF, 2023b). The continuous demand from banks for sovereign bonds in a context of high demand for safe credit assets is challenging from market making banks: governments regularly issue relatively small amounts of bonds to ensure liquidity remains available and that banks can continue operating as market makers for riskier assets (ECB, 2023; MoF, 2023c).

Finally, Retail and Prime Brokerage Departments of Wholesale Divisions within investment banks hold sovereign bonds in the name of their clients, institutional ones (e.g. mutual and pension funds) but also in the name of buy-side counterparties (e.g. hedge funds and high-wealth individual investors). The volumes involved are vastly larger than the previous exposures described. Bonds held in the name of other financial actors carry a different pricing of their risk than ones within the banks' own portfolios. The bank acts as a third party in these positions. This role still carries a substantial risk (BoE, 2022b). Nomura specifically has recently suffered multiple large losses from the risk carried by such positions, as counterparties default and do not meet margin calls precisely when their collateral's value is affected (Bouveret and Haferkorn, 2023; Griffin and De Paoli, 2024). These particular positions largely consisted of equity exposures, but other recent events have called attention to analog risks in credit markets (Alonso Olmedo et al., 2021).

In short, banks own broad volumes of sovereign bonds within various parts of their activity. This Chapter provides the description of three models developed internally

at Nomura, improving on their predecessors at predicting the movements of the risk charge carried by sovereign bonds.

We attempt to model and predict the movements of sovereign credit grades using data available to Nomura. We have access to dozens of financial, social and economic time series tracking the evolution of countries across 2 decades.

To assess the quality of our model, we first build an intuitive and straightforward benchmark, the “*constant predictor*”. This comparison metric simply assumes that a country’s grade never changes: for any timestamp and country, it outputs the last available observed credit grade for that country. The low rate of change of sovereign credit grades makes this predictor perfectly accurate in almost 80% of cases. When wrong, this predictor still is often close to the true value.

While carrying no predictive power, this benchmark is high. To our knowledge, no prior published model meets its accuracy (Overes and van der Wel, 2023; Ozturk et al., 2016). Typical industry legacy models, presented in the first section of this Chapter (see Section 1) also are not as precise as this benchmark.

The two final sections of this Chapter present two classifiers. Our neural network classifier aims at accuracy solely, and its out of sample results outperform the “*constant predictor*”. Our Random Forest model intends to carry explainability and more importantly homogenous predictive accuracy across classes. That is, we want this model to be good at identifying grade changes, the minority class in our dataset. This latter random forest classifier achieves 80% of accuracy over the “*grade change*” class, while maintaining high accuracy ($> 90\%$) over the larger “*no-change*” class. All our claims are based on cross-validation performed over separated entries. To ensure absence of data leakage (Wang and Ruf, 2022), the cross-validation is carried over either new countries or the most recent years of data accessible, strictly splitting the test and train sets geographically or temporally.

The accuracy is a substantial improvement over previous implementations. More importantly, the two predictors novelly demonstrate that machine learning models can calibrate accurately on macroeconomic datasets limited both in size and quality. Other studies have prescribed the use of machine learning for the prediction of sovereign credit grades (Anand et al., 2023; Ekmekcioglu et al., 2023). Yet, published models never met a level of precision suitable to be used in practice instead of simple benchmarks such as the “*constant predictor*”. In finance, especially in pricing, machine learning has already proven its high value, yet these models are usually calibrated over vast datasets (Buehler et al., 2019a). To the best of our knowledge, the present study seems to be the first to attain a satisfactory level of accuracy calibrating over small and low-quality macroeconomic datasets towards credit grading. In particular, it opens the door to the future use of machine learning in larger macroeconomic prediction tasks for which voluminous datasets are by definition not available. Moreover, we highlight the impor-

tance of our problem-specific data pre-processing framework, motivated by statistical insights but also by the valuable collaboration with credit professionals.

We first provide a brief overview of the special treatment that sovereign bonds hold in modern sell-side banking. This summary motivates better sovereign credit grades prediction and highlights their centrality in pricing sovereign bonds' risk.

In particular, we focus on the regulatory framework, and how international banking regulation indirectly entices banks to be exposed to vast amounts of investment grade sovereign bonds, concentrated over *AAA* and *AA* creditors.

The rest of this Chapter contains five sections. The first section provides an outline of a typical current industry model. Its accuracy provides a benchmark for the subsequent models presented. In parallel, this model is not only a typical industry model but also similar to credit rating agencies' models. As such, its relatively low ability to predict S&P sovereign credit grades highlights the centrality of the human input in agencies' decisions.

The second section studies the available dataset and outlines the main difficulties of this modelling task. In particular, it focuses on two such challenges. Firstly, the available dataset has a high dimension. Yet, the number of entries in the training set is capped by a multiple of the number of countries with credit grades. As legacy and agency models rely on proprietary implementations of regressions (Duan et al., 2004; Metz and Cantor, 2006; Poor's, 2011; Wilde, 1997), the dimensionality problem entailed cannot be ignored (Chatterjee and Simonoff, 2012; Christensen, 2019).

The third section first describes a small-sized neural network model, with few layers and weights, which inputs are similar to legacy models'. We show that it substantially outperforms legacy model at predicting grade outputs. Its performance was assessed on independent entries, geographically / temporally. That is, the test set contains either countries that the network has never seen, or data for years that the network has never seen. This model confirms the soundness of a machine learning approach to predict sovereign credit grades. While machine learning has been proven to work both in financial (Buehler et al., 2019a,b; Israel et al., 2020) and macroeconomic (Biagini et al., 2022) settings (*see the following survey on machine learning in finance* (Gogas and Papdimitriou, 2021) *and neural networks specifically* (Ozbayoglu et al., 2020; Ruf and Wang, 2020)). Other studies have given analytical support to ML results in finance (Han et al., 2018).

In the fourth section, we describe the most accurate neural network model developed, using multiple years of data at once. This approach satisfies Credit Risk's first demand for a model targeting accuracy only. Over an independent testing set, this model outperforms all previous known implementations. More importantly, it is to our knowledge the first model to be more accurate than the "*constant predictor*". The "*constant pre-*

dictor” is benchmark presented in the second section, carrying no predictive power but highly accurate.

This “*multi-year neural network*”, despite its accuracy, functions as a blackbox. The fifth and last section presents a random forest model (Atsushi and Lutz, 2004; Biau, 2012; Yim and Mitchell, 2005) that, though less predictive overall, is more useful to Credit Risk analysts. Indeed, by minimizing Type-II Error, and managing to provide insight into the drivers of a “grade change” prediction, it helps analysts reduce the number of countries they need to focus on when reviewing grades. Due to “fire-sale” behaviours (Pinter, 2023) when sovereign credit grades are updated, it is more important to Credit Risk analysts to spot the countries for which a grade change is imminent than the opposite. Minimizing the rate of false negatives, a country which grade’s will change labelled “*no-change*” is thus the priority in this implementation. Specifically, this implementation manages to reach comparable accuracies for the two classes, despite being largely imbalanced.

The three predictors built use different sets of features. The first (*see Section 3*) uses a set of 9 features designed to replicate the ones used by the legacy models (*see Table 3.5*). Our most accurate predictor (*see Section 4*), uses a proprietary set of approximately 20 features. The random forest predictor (*see Section 5*) uses a set of 13 features (*see Table 3.11*) derived from Nomura’s economists’ recommendations.

Our conclusion outlines potential next steps to enhance the quality of the models further.

SOVEREIGN BONDS RISK PRICING & THE IMPORTANCE OF CREDIT GRADES

The core metric of a bank’s solvency is its Tier 1 capital ratio. It divides a bank’s current bank’s core capital to its total Risk Weighted Assets (RWA) (Cannata et al., 2012; Garcia and Arroyo, 2012; Le Leslé and Avramova, 2012). Regulatory bodies set minimum Tier 1 ratios to ensure banks can withstand losses incurred by future stressed market conditions. This minimum generally varies between 5% and 9% (BCBS, 2022b; BCBS, 2022a; Coggan, 2009).

Riskier assets carry a higher weight in RWA and thus contribute more towards reducing the Tier 1 ratio. As described in the previous Chapter, banks can choose whether they use the *standardised approach* (SA) or the *internal models approach* (IMA) to calibrate their risk measures.

Within the *standardised approach*, publicly available credit rating agencies sovereign grades are the cornerstone of modelling weighted risk carried by sovereign bonds. Indeed, the SA calibrates the risk weights directly from the credit grades (*see fig. 3.1*).

Within IMA’s framework, internally calibrated sovereign credit grades are central too.

Credit ratings and sovereign risk weights under the Standardised Approach						
In per cent						Table B
Credit ratings	AAA to AA–	A+ to A–	BBB+ to BBB–	BB+ to B–	Below B–	Unrated
Risk weight	0	20	50	100	150	100

Source: BIS.

© Bank for International Settlements

Figure 3.1: Direct Correspondence between probability of default and risk weight. Zero risk is assigned to AAA, AA+, AA and AA- bonds. In turn, their *RWA* is also 0.

Source: Bank for International Settlements, Treatment of sovereign risk in the Basel capital framework, December 2013

The RWA of a bond then is:

$$RWA = (Loss\ Given\ Default) \cdot (Default\ Probability),$$

$$\text{Where } (Loss\ Given\ Default) = 1 - (Recovery\ Rate).$$

It is mandated that the Loss Given Default (LGD) of sovereign bonds is 45% (BIS, 2013). The recovery rate of sovereign bonds is capped at 65%: a default results in losing at least 45% of the bond's notional. While realistic, it is not a conservative bound. Greek bonds' recovery rate in 2012, over more than two hundred billion US dollars, was below a third (Kamenis, 2014). In 2001, analog 70% haircuts were imposed on defaulted Argentine debt (Doyran, 2014). In general, recovery is a long and uncertain process (Altman and Kishore, 1996; Cantor and Varma, 2004).

In parallel, IMA provisions that probabilities of default are to be set by banks. In particular, it imposes no minimum on PDs: in practice, banks can exclude sovereign bonds from risk encompassed by the Tier 1 ratio by setting the sovereign PDs to 0 (BIS, 2013). The absence of such floor for sovereign PDs has lead to the widespread belief that sovereign bonds are excluded from “*Basel*” regulations (Lanotte et al., 2016).

However, while “*Basel 1.0*” assigned 0 risk weight to OECD sovereign debt² (BCBS, 2019c), banks typically do not apply a 0% probability of default anymore. Sampled institutions report a 0.1% sovereign PD (BCBS, 2013). Multiple studies developed more advanced models for calibrating default risk (Altman, 1998; Bluhm et al., 2010; Crouhy et al., 2000; Delianedis and Geske, 2003; Jarrow et al., 2005).

The SA setting defined risk-weight directly from grades. By contrast, in the IMA framework, sovereign credit grades drive the calibration of PDs, indirectly impacting the risk weights. In practice, financial institutions often set PDs directly from the credit grades using historically calibrated default matrixes. While the the impact of

²Generally, sovereign bonds emitted by OECD countries are considered among the least risky financial assets. Arguably, no OECD country has defaulted on its debt for 70 years. Actually, wealthy countries have defaulted on their debt, notably the US in 1979 (Reuters, 2011). However these instances are marginal and were not caused by the insolvency of the states backing the bonds.

grades on the bonds' risk is not as straightforward as in the SA, they still are a major driver of the final risk-weights from sell-side bond exposures.

We presented the need for modelling sovereign credit grades with respect to the all-important “*Basel*” Tier 1 Capital Ratio. We now provide a non-exhaustive overview of other drivers of risk associated to sovereign bonds, which in turn motivate for better prediction of sovereign credit grades modifications.

As described in the previous chapter, regulating bodies adapt regulation in the face of new crisis. During the eurozone crisis European regulators specifically started overseeing the level of exposure to sovereign risk banks carried (EBA, 2012). After its 2014 continent-wide stress testing program, the European Banking Authority (EBA) reported: *“The stress test required a specific treatment of sovereign exposures. In particular, fixed valuation haircuts were prescribed for all EU countries in the adverse scenario to take into account the credit and interest rate risk of direct sovereign exposures. Other sovereign exposures [...] were to be stressed consistently based on the market risk scenario. In total, more than EUR 2,700BN of direct sovereign exposure were in scope of the stress test. 48% of direct exposures are held as Available for Sale for which haircuts had to be applied to project the impact on capital via other comprehensive income. The second biggest accounting category with 35% – Loans and Receivables and Held to Maturity was subject to the credit risk methodology.”* (EBA, 2014). Banks failing EBA stress test scenarios need to re-capitalise so that their stressed Tier I Ratio remains within the international *Basel* guidelines.

This firstly shows that sovereign bonds are not entirely exempt from regulatory capital minimum requirements (Lanotte et al., 2016).

It also highlights that the credit securities carry various distinct forms of risk.

Indeed, while being the largest risk carried by bonds, default risk is not the only one.³ In particular, through their market making activities, banks are exposed to the fluctuations of bonds' face value.

While the largest driver of bond price volatility are interest rate changes by the central banks emitting the bonds, substantial price swings occur even in the absence of such rate modifications (Financial Stability Board, 2022). For instance, foreign exchange market changes as well as yield curve movements expectations affect bonds' valuations. Credit grade changes too affect the market value of loans emitted by the concerned country. Thus, as we underlined the value of credit grade predictions to Credit Risk, their modelling is also important to Market Risk.

³Also presented is the distinct risk profiles attached to identical assets held in different portfolios. While in normal market conditions, the separation is clear, it blurs in stressed market conditions (Aharon and Ali, 2024; Hauf and Posth, 2023).

SOVEREIGN CREDIT GRADES

As outlined previously, any sell-side institution aiming at using “*Basel*”’s Internal Model Approach to calibrate its risk limits ought to have its own credit grades model. Non-regulated agencies also calibrate widely used sovereign credit grades. Moody’s, S&P Global Ratings (previously Standard & Poor’s (S&P, 2011; S&P, 2015)) and Fitch are the leading ones, whose grade modifications strongly impact markets (Cantor and Packer, 1996). Inspired by common modelling bases (Bluhm et al., 2010), these models rarely differ substantially (Hill et al., 2010) despite very distinct implementations. The two main general references are Moody’s own KMV model (Duan et al., 2004; Metz and Cantor, 2006), and Credit Suisse’s CreditRisk+ model (Wilde, 1997).

While methodologies of these models are known or even public, all their calibration processes rely at some point on proprietary analysts’ judgment. This adjustment renders grades hard to predict and not a simple output from countries’ various indicators (De Moor et al., 2018; Reusens and Croux, 2016). From this, grades become financial and economic indicators of their own (Cantor and Packer, 1996; Erdem and Varli, 2014; Partnoy et al., 2002), impacting other indicators (Afonso et al., 2012; Cheikh et al., 2021). The importance grades have gained motivated attempts to improve their ability to represent countries’ situations and to widen their scope (Anand et al., 2023; Ekmekcioglu et al., 2023). Their influence reaches macroeconomic importance and studies have exhibited their impact at this scale (Sy, 2009; Utzig, 2010).

Regulated sell-side sovereign credit grades models are usually entirely proprietary, but aforementioned exposure to market risk incentivises convergence with rating agencies’ decisions. Credit grades are usually reviewed every quarter, but agencies reserve the right to modify grades at any time if deemed necessary.

Since there is human intervention at stake, replicating the statistical side of agency models is not enough to replicate their output (Afonso, 2003; Oral et al., 1992; Rowland, 2004; Vij, 2005). The models presented in this Chapter develop a new sovereign credit grade prediction engine. The use of Neural Networks (Russell and Norvig, 2010; Shalev-Shwartz and Ben-David, 2014) and Random Forests (Atsushi and Lutz, 2004; Breiman, 2013; Cutler et al., 2011) aims at being able to take complex interactions between features into account, in contrast with classical financial models (*see Section 1*).

MODEL PREREQUISITES

All major financial institutions have access to extensive private datasets which encompass a wide range of indicators, datasets provided by third parties but also proprietary ones. For every country, decades of data are available for dozens of indicators. These indicators range from financial and capital markets features to social and demographic

descriptors. The full list of features available is reported in *Appendix 1*. We opine, as outlined in above, that complex combinations of these indicators are necessary to capture the credit grades movements.

Combining statistical know-how and Credit Risk analysts' economics experience⁴, our goal is the development of a statistical model which leverages the aforementioned unused sovereign indicators.

In particular, the requirements of the model are the following:

- Accurate against some appropriately defined benchmark, e.g. current production models / simple predictors. Prior legacy and internal models serve as a first comparison metric.
The aforementioned "*constant predictor*", formally defined in *Section 2* is our main comparison measure. While our first models, replicating the input / output structure of current models, do not meet this benchmark, *Section 4* outlines classifiers that do.
- Stable through time so that model maintenance does not require extensive re-work. Any successful model would be used within the bank's systems. Changes to any model's global structure entail going through the process of model validation. Frequent re-validation would render the model unusable as, pending validation, any model's predictions cannot be used. Our models' performance will be continually monitored to assess their quality.
- For every sovereign credit that is changed within a given yearly period, there are on average 3 which are not changed. As a consequence, the classes in our training set are very imbalanced. The desired model performs correctly over the two classes. A low error rate over the dominant class while coupled with a high error rate over the minority class would yield a high overall accuracy. This behaviour would however mimic the "*constant model*"'s behaviour. The random forest predictor presented is designed to have balanced predictive power and manages accuracy on the minority class nearing the majority one.
- Able to identify the main drivers of a prediction, especially when a change is predicted. At first, this predictor is to be operated by Credit Risk analysts as an additional indicator helping to make the final human decision on a country's credit grade perspective. It is thus crucial for any predictor to be able to highlight the main drivers of its estimation, in particular when the predicted outcome is not the expected one.

⁴We thank Alfonso de la Torre and Kevin Duning in particular for their guidance and help throughout this project

1 LEGACY INDUSTRY MODELS AND LIMITATIONS

To illustrate the type of model commonly in use in the industry, we give a summary of a generic legacy model typically being used in production currently⁵.

Sovereign Credit Ratings are commonly displayed as letters from *AAA* (safest grade) to *C / D* (worst grade – attributed to countries currently in default). However, they are calibrated as integers/notches from 1 for *AAA* to 18-20 for countries in Default (*see fig. 3.2*). These integers are the prediction target of all the models presented. All features and sovereign credit grades in our dataset are updated yearly. The economic, social and financial data driving the grade changes are usually published prior to grade modifications. Thus, our model targets to predict countries’ yearly grade updates from the datapoints for that year and optionally previous years.

More frequent data updates and sovereign credit grades are available, up to a quarterly frequency. However, as outlined in *Section 2*, point **C**, grades on average do not change every year. Thus, using quarterly instead of yearly features and grades keeps the total number of grade changes constant. Meanwhile, each frequency increase multiplies the total size of the dataset. Thus, separating the dataset between grade “*change*” and “*no-change*”, augmenting the observation frequency simply accutes the imbalance between the two classes.

A typical model used in financial institutions is a linear regression based model using some set of pre-determined features, such as the nine shown in *Table 3.1*. It is not obvious that linear models are not adequate to tackle this predictive task ([Brunton et al., 2016](#)), as they have been shown to outperform more complex methodologies such as small scale neural networks in financial settings ([Ruf and Wang, 2022](#)).

Government Quality
Labor Market Quality
Soundness of Sovereign Debt
Economic Reliability
History of Defaults
Banking Sector Quality
Income Level
Sovereign Reserves
Debt Load

Table 3.1: Typical List of Features used in predicting the evolution of sovereign credit grades.
Redacted: This list represents broad categories which each encompass actual individual features used.

An entry in the dataset consists of one datapoint for each of these nine features, plus the Credit Grade to predict. Meaning that any country will appear in the dataset

⁵*This model does not represent the model used currently by Nomura, which is proprietary.*

Classification	Companies			Scale
	Moody's	S&P	Fitch	
	Aaa	AAA	AAA	1
	Aa1	AA+	AA+	2
	Aa2	AA	AA	3
	Aa3	AA-	AA-	4
Investment grade	A1	A+	A+	5
	A2	A	A	6
	A3	A-	A-	7
	Baa1	BBB+	BBB+	8
	Baa2	BBB	BBB	9
	Baa3	BBB-	BBB-	10
	Ba1	BB+	BB+	11
	Ba2	BB	BB	12
	Ba3	BB-	BB-	13
	B1	B+	B+	14
	B2	B	B	15
Speculative	B3	B-	B-	16
	Caa1	CCC+	CCC+	17
	Caa2	CCC	CCC	18
	Caa3	CCC-	CCC-	19
	-	CC	CC	20
	-	C	C	21
	Ca	SD	DDD	22
	C	D	DD	
	-	-	D	

Source: Fitch, Moody's and Standard & Poors.

Figure 3.2: Grades mapping and equivalences between three main rating agencies. Grades above 20, (21 and 22), are considered in default, and combined with grade class 20.

Investment banks' exposures are concentrated over investment grade bonds.

Source: Sovereign risk ratings' country classification using machine learning. D. Silva, Thais G. do Rego, B. Frascaroli. 2019. (Silva et al., 2019)

Available at: <https://api.semanticscholar.org/CorpusID:226305710>

Last Accessed: April 27, 2024.

many times, once per period. Economic indicators are often only updated on a yearly basis. Thus, every entry in the dataset is indexed by a *(Country, Year)* pair.

This structure ideally allows for a large enough dataset to be used. Data is reliably available for over 20 years and over 200 countries. Consequently, the effective dataset reaches up to 4000 entries. Problematically, for models more complex than Linear Regression, as many countries do not change substantially year-on-year, the “effective” size of the dataset is reduced and it potentially introduces leakage.

Importantly, this input structure does not capture any evolution in the figures used to predict a country's grade, nor its previous grades. Only a one year snapshot is fed in the model.

The problem then becomes the interpolation of an integer class from 1 to 20 using the 9 datapoints. Since features have distinct units, scales and directional impacts on a country's grade, a first transformation is run on individual features to transform their

raw values into a *Parameter Impact*.

Before calibrating the *Parameter Impact*, the individual features are normalized. This normalization is not necessarily done through the classical mean-variance normalization. For instance, Rank features, integers values ranging from 1 to the number of available countries, are usually centered through their median instead of their average.

$$\text{Normalized Feature} = (\text{Raw Feature Value} - \text{Median}),$$

$$\text{Or alternatively Normalized Parameter} = \frac{\text{Raw Feature Value} - \text{Average}}{\text{Feature Variance}}.$$

A linear transform of slope β_f is then applied to each *Normalized Feature*:

$$\text{Parameter Impact} = \beta_f \cdot \text{Normalized Feature}.$$

Finally, the individual *Parameter Impact* figures are combined in the final predicted score, with reactivity β_G , then rounded to the nearest integer to obtain the predicted sovereign credit grade:

$$\text{Predicted Score} = \beta_G \cdot \sum_f \text{Parameter Impact}.$$

In practice, it is required of the model that its average / median predicted value has to match the average / median observed grade, *BBB*, 9 (see Table 3.2).

Thus, intercepts can be added to either the final *Predicted Score*, I_G or to each of the *Parameter Impact*, I_f , or to both. *Intercept* values grow proportionately to every feature's standard deviation σ_f .

$$\text{Parameter Impact} = \beta_f \cdot \text{Normalized Feature} + I_f.$$

$$\text{Predicted Score} = \beta_G \cdot \sum_f \text{Parameter Impact} + I_G.$$

The variability of the implementations lies in the methodology used to obtain the β_G , β_f coefficients and the intercepts. Some implementations set β_G to one and individual or global intercepts to zero. Others use the feature specific regressions as normalizations only.

The final coefficients and intercepts $(\beta_G^*, \vec{\beta}_f^*, \vec{I}_f^*)$ can all be obtained at once minimizing the quadratic loss function of the whole model:

$$\min_{\beta_G, \vec{\beta}_f, \vec{I}_f} \sum_{(Country, Year)} \left(\beta_G^*, \vec{\beta}_f^*, I_G(\beta_G^*, \vec{\beta}_f^*, \vec{I}_f^*), \vec{I}_f^* \right) = \left(\text{Grade} - \text{Predicted Score}(\beta_G, \vec{\beta}_f, I_G(\beta_G, \vec{\beta}_f, \vec{I}_f), \vec{I}_f) \right)^2.$$

Where I_G^* is set depending on the other features so that the average (or median) grade

predicted is *BBB*:

$$I_G(\beta_G, \vec{\beta}_f, \vec{I}_f) = 9 - \frac{1}{|\{(Country, Year)\}|} \sum_{(Country, Year)} Predicted\ Score(\beta_G, \vec{\beta}_f, 0, \vec{I}_f). \quad (3.1)$$

In this setting, individual intercepts I_f , if not set to zero, are not used to ensure the average predicted grade is *BBB*. Also, as some features might be correlated, this optimisation might artificially reduce the weight of one feature against another.

Alternatively, the optimisation can be processed in two steps, first finding optimal individual β_f^* and I_f^* . Once the first optimisation is run, the final global coefficient β_G^* and intercept I_G^* are determined so that they minimise the global loss function.

$$\begin{aligned} (\beta_f^*, I_f(\beta_f^*)) &= \min_{\beta_f} \sum_{(Country, Year)} (Grade - Parameter\ Impact(\beta_f, I_f(\beta_f)))^2, \\ I_f(\beta_f) &= 9 - \frac{1}{|\{(Country, Year)\}|} \sum_{(Country, Year)} Parameter\ Impact(\beta_f, 0). \end{aligned}$$

Analogously to I_G^* in the previous optimisation, each I_f^* is defined so that, given β_f , the average / median *Parameter Impact* predicted over the dataset corresponds to *BBB*, the median observed grade.

As outlined, global features are then defined as:

$$\begin{aligned} (\beta_G^*, I_G(\beta_G^*, \beta_f^*, I_f^*)) &= \min_{\beta_G} \sum_{(Country, Year)} (Grade \\ &\quad - Predicted\ Score(\beta_G, I_G^*(\beta_G, \beta_f^*, I_f^*), \beta_f^*, I_f^*))^2. \end{aligned}$$

And I_G^* is determined according to (3.1).

Effectively, the two calibration methods laid out yield similar results⁶. While the former is by definition more accurate, the latter grants a more intuitive explainability of the model, each feature being initially treated independently as a separate predictor.

For both agencies and sell-side institution, results lower than 17 are capped at 17 so that a country can only be classified as *In Default* through a human decision.

Table 3.2 presents the expected quality of a typical industry model constructed as outlined above. We replicated this model, and implemented the first optimisation technique. Results obtained for data from 2017 & 2018, with a total of 244 entries, without separating for a train, validation and test set, are in accordance with Table 3.2. The set of grades used is from Standard & Poor's. Swapping it for Moody's, Fitch's

⁶Redacted.

or proprietary ones does not bring substantial change to the accuracy as the volatility of their grading is similar (Hill et al., 2010).

For this dataset, Sovereign Credit Grades' distribution is presented in *Table 3.2*. As this distribution is stationary over long periods of time (Ratings, 2015), one can use it as a general distribution of Sovereign Credit Grades.

S&P:	AAA	AA+/-	A+/-	BBB+/-	BB+/-	B+/-	CCC+/-	CC & C/D	NG
Notches:	1	2-4	5-7	8-10	11-13	14-16	17-19	20	-1
Dist:	10%	10%	17%	18%	18%	16%	3%	7%	-

Table 3.2: Distribution of sovereign ratings by Standard & Poors, 2019.

Prediction for S&P Grades	Exact Match	Within 1 Notch	Within 2 Notches	Total
Correction Rate	~ 20%	~ 65%	~ 85%	100%

Table 3.3: Legacy model typical prediction quality. Prediction one notch away for the true value can be considered acceptable. *Redacted*.

2 CONSTANT PREDICTOR & MODELLING CHALLENGES

The Linear Regression model presented in the prior *Section 1* can serve as a benchmark for a more advanced model. In this section, by providing some insight in the datasets behaviour we build a benchmark that is easier to build, maintain and more accurate than the legacy linear models.

In particular, we establish the class imbalance of the dataset, its unequal geographic distribution and its relatively small effective size.

Data from public institutions (IMF, World Bank) is generally available and widely used. Since this model is developed in collaboration with Nomura, also accessible are datasets from private vendors such as FactSet or Bloomberg. Proprietary datasets are available too. The dataset's structure is shown below (*see fig. 3.3*). The full list of available indicators is shown in *Appendix 1*. Data vendors report these indicators since 2000 but for the most part, data is reliably procurable after 2007.

We now summarize the main challenges of predicting Sovereign Credit Grades. As mentioned, from this analysis we justify the construction of the "*Constant Predictor*". In parallel, these challenges entail limitations on the maximal quality of potential models.

A - The constructed dataset we work with does not represent every class of countries equally.

In theory there are around 195 countries in the world. Sovereign socio-economic

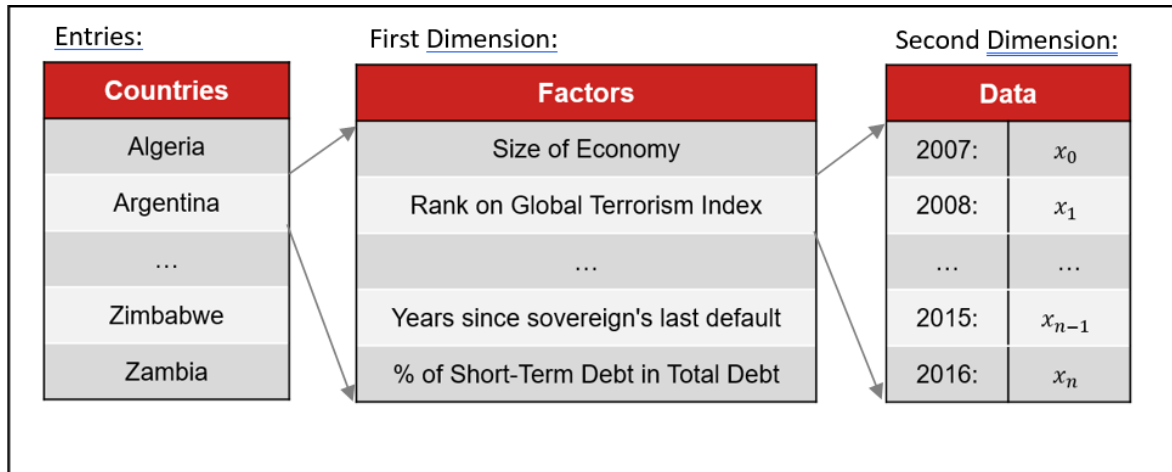


Figure 3.3: Structure of the Available Dataset.

figures have been reported for at least 20 years. Considering every theoretically available *(Country, Year)* pair, the dataset could contain about 4000 entries.

However, no more than 170 countries are graded by credit agencies (Standard's and Poor's has graded 164 countries in total but some are not graded anymore/were graded for a few years only). Indeed, some countries have no sovereign publicly traded bonds and consequently require no sovereign credit grade. Nomura and sell-side institutions usually grade a larger set of actors, as their portfolios often contain exposures to debt that is not publicly traded.

Among the set of features used for each country and that are officially relevant to Credit Agencies' calibration, many are available for less than 20 years. For parts of the world, data is not available prior to 2008. For other countries, data is not reliably reported.

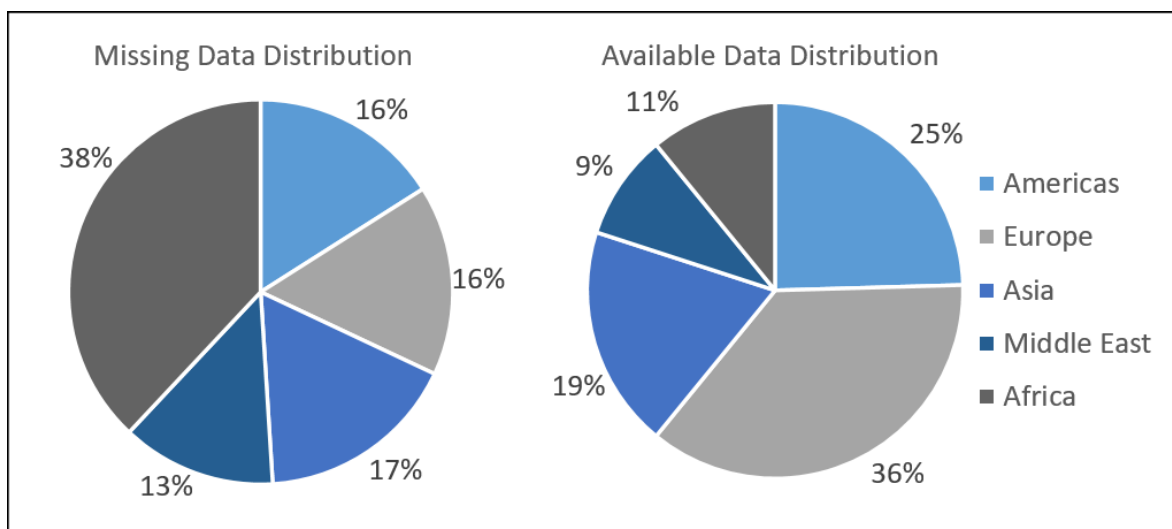


Figure 3.4: Geographical repartition of countries with thoroughly reported data from 2008 onwards. The set of available countries represents disproportionately wealthy western countries. While wealthier countries see less grade updates, banks' sovereign bonds exposure is concentrated on investment grade countries.

As shown in the next point (*see fig. 3.5*), the average behaviour of countries' credit grades differs depending on their level of economic development. While predictors can differentiate countries' development level through provided features, it entails that some effective classes have a very limited size. Indeed, data availability is very unequal. Good performance over developed countries over-represented in the dataset could shadow worse accuracy over the rest of the world. *Fig. 3.4* shows that the imbalance is largest for Africa and Europe. Africa represents 11% of available entries despite representing more than one in four sovereign state. Europe composes 36% of our dataset, but is home to only 22% of nations⁷. This distribution does not take into account singular missing entries: countries counted as missing are not even included with empty time series returned. These sovereign actors are simply not reported upon.

Our aim is to use a large number of indicators at once to predict grades. Thus, to include a *(Country, Year)* entry in the dataset, all chosen indicators need to be available for a given year. This decreases the effective size of the training set further.

In effect, from economists in the Credit Risk Department, we get a set of features that are considered to be the most adequate to explain sovereign credit grades' movements. From this set, we extract 13 features that are used in the random forest predictor presented in *Section 5* (*see Table 3.11*). Combining these 13 features, the number of *(Country, Year)* entries in the dataset is reduced to 404.

Furthermore, an ideal model would include multiple prior years of data in order to capture a country's economic trajectory. Requiring more datapoints per entry, it reduces the dataset's size further. With the previous set of indicators, using a 3 years window of data to predict grades decreases the number of entries by 25%. This point is discussed more thoroughly in *Section 4*.

Finally, the reliability of the sourced data is known not to be perfect. Macroeconomic and social indicators in particular are known to be inaccurate. Recently published figures are not final and are regularly revised ([Glass and Fritsche, 2015](#)). Furthermore, it is often not straightforward to integrate long term historical data into one dataset, as methodologies and accounting practices have changed ([Weir, 1986](#)). As methodologies have variated across time, they also vary between reporting institutions and countries ([Ariyo, 1995](#)). Some countries are widely suspected of reporting wrong economic data, either voluntarily or through incentivizing over/under-valuation of the reported data ([Koch-Weser, 2013](#)).

⁷According to the United Nations Group of Experts on Geographical Names (UNGEGN) ([UNGEGN, 2011](#)).

As a general example regarding US corporate debt specifically, bankruptcy laws in the US have changed substantively across decades. As a result, the frequency as well as the distribution of corporate defaults has changed, independently of the surrounding macroeconomic context (Zywicki and Skeel, 2003).

These data quality and reliability concerns are not covered within the scope of this Chapter. Nomura already has extensive processes in place to ensure the quality of the data it uses. It has engineered proprietary correction and proxying methods to correct potential inadequacies.

- B -** Grade updates by more than 2 notches within a year are extremely rare. Among these occurrences, almost all happen when countries default on their debt. As they default, their rating jumps to 20⁸. Often, the previous grade of defaulting nations was above *CCC*, thus an unusually wide jump is observed. This reflects the fact that grades representing non-defaulted bonds are expectations, while the top of the scale depends on actual status of securities.

Some countries defaulted multiple times across the 20 years encompassed in our study. In this scenario, rating agencies correctly give distinct grades to distinct bonds⁹. As such, some countries carry multiple credit grades: Argentina's bonds are graded (*D*, *D*, *D*, *D*, *CCC*). Only its newest issuances have a non-defaulted grade.

Therefore, we reconstruct full time series of sovereign credit grades by concatenating the series of bonds as follows.

For countries with no history of default, rating agencies report a single sovereign credit grade per year, no further work needs to be carried out.

Nations which have defaulted usually re-list publicly traded bonds a few years after their default event. As the re-listing happens, rating agencies continue reporting the grade on the previously defaulted bonds as 20 or above. The newly emitted bonds, trading publicly, get assigned a new grade, which behaves as a regular sovereign credit grade.

For each year, if available, we use the non-defaulted grade of a country. However, we ensure that each default event appears in time series. Indeed, countries sometimes re-emit bonds in the very year they defaulted. In this case, we report the country's grade as in default (20) for one year, before switching to the more recent non-defaulted bonds' grading.

⁸Defaulted bonds actually have a higher rating, which is capped at 20 in our examination (*see fig. 3.2*).

⁹Actually, all countries typically carry more than one credit grade. As they often have debt denominated in multiple currencies, their domestic currency and the US Dollar, each category owns a different grade. Short-term and long-term debt also carry specific ratings. The generally admitted consensus is to consider the long-term US Dollar credit grade, and the domestic one shall a country's domestic currency be a reserve one.

As states successively default and re-list bonds, sometimes even graded as *B* bonds, our constructed credit grade jumps to and from 20.

We described that almost all the occurrences of grade jumps by more than one notch relate to default events. There are 20 different classes to predict: one per grade, from 1 to 20.

Considering that the size of the dataset is limited by the number of countries with publicly traded debt, some classes can get very small, especially considering the biases outlined in **A**. As they represent emerging countries, less represented in our dataset, classes corresponding to speculative bonds in particular are smaller.

To circumvent the issue, we rely on the fact that financial actors are interested in predicting a country’s grade relative to its current one. That is, we want to predict by how many notches a grade is going to change. Hence, we replace the initial 20 classes containing few instances each by classes representing the grade changes instead.

Since grade updates by more than two notches do not occur in significant numbers, we can reduce the target set to:

- $\{-2, -1, 0, +1, +2\}$.

We can reduced it further to:

- $\{Negative\ Change, No\ Change, Positive\ Change\}$,
- $\{No\ Change, Change\}$,
- $\{No\ Change, 1\ Notch\ Change, > 1\ Notch\ Change\}$.

Each of these four options will yield larger training classes, but which will be unbalanced. As already mentioned, grades on average change every four years only. Thus the “*no-change*” class is multiple times larger than the other ones. This modification also implies that countries at very different stages of development will be included in the same class. As analysts’ standards vary for various types of countries, the classes are now much less homogenous.

Indeed, the arbitrary human component of rating agencies’ decision affects countries differently depending on their level of development (*see fig. 3.5*).

Deciding between the four redefinitions of targets can be motivated by end-user motivation:

- The neural networks developed, lacking explainability use the first redefinition $\{-2, -1, 0, +1, +2\}$.

The aim of neural networks, presented in *Section 3* and *4* is to maximise the global accuracy of the predictor. It is not possible to easily explain the

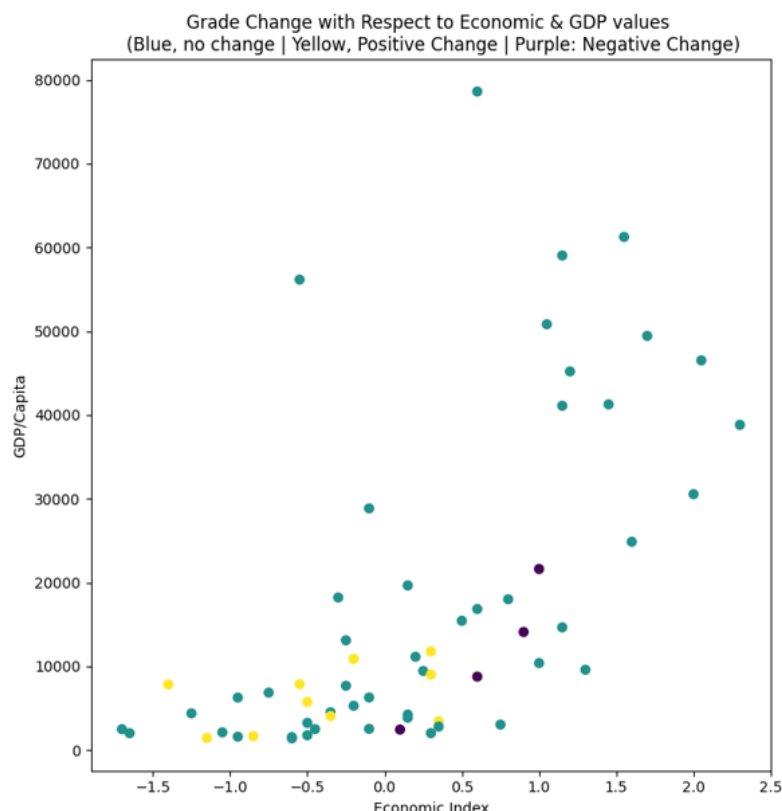


Figure 3.5: Grade changes – Data from 2018 & 2019 only, Countries with data missing for 2018 or 2019 excluded.

We clearly observe that the positive grade changes that occurred are concentrated in low income and developing countries. All high income countries saw their grade unchanged between 2018 and 2019. All countries downgraded also are low-income nations.

drivers of such neural networks' outputs. The neural network model, while initially motivated to be a proof of concept, intends to be embedded within wider Nomura systems.

As no human intervention will occur after the model's prediction is made, it is essential to know automatically in which direction the rating will move.

By no human intervention, we refer to applications that, while being run by Nomura employees, while take the credit grade output of the model directly.

For instance, this happens within risk models with long horizons. These models need to calibrate assets' risk weights for the duration of these long horizons. Thus, for calibrating risk over sovereign bonds with months of holding periods, the best predictors for grades' outlooks are useful.

Similarly, computing haircuts for bonds, automatically knowing accurately in advance the risk-discount to apply, tied to the future credit risk of the securities, is of substantial value.

These applications motivate the aforementioned selection of target classes.

- Presented in the last section of this Chapter, the random forest model uses $\{ \textit{“No-Change”}, \textit{“Change”} \}$ as target classes. Contrasting with the previous model, it is intended to permit credit analysts to focus on a narrowed list of countries. Indeed, typically, 200 countries have to receive a credit report every quarter. The random forest classifier can highlight which bonds have the highest likelihood of having their rating changed. Providing an explanation as to the drivers of the prediction by analysing the features’ respective importances is also practical. It allows for the direction of the prediction to be determined easily despite not being a formal output of the model.

As explained, we want to output a reduced list of nations to focus upon which will contain a proportion of all countries in the *“Change”* class as large as possible. We thus minimize the number of Type-II Errors first (*“Change”* labelled as *“no-change”*) and reduce the weight associated to false positive errors (*“no-change”* classified as *“Change”*).

- C - Credit rating agencies normally schedule grade reviews quarterly, though in singular situations grades can be updated more often. Generally grades are not changed when reviewed. At a yearly level, over >75% of available countries do not see their grade changed. As already outlined, swapping yearly for quarterly data would result in a bigger dataset with a wider class size disparity.

The low rate of grade changes entails a strong imbalance in the target classes defined. In the $\{ \textit{“No-Change”}, \textit{“Change”} \}$ setting, the *“No-Change”* class represents 77% of the total available dataset. It is thus significantly more costly for our models to have a lower prediction rate over the majority *“no-change”* class. This in turn drives the grade change predictions towards 0.

In parallel, also detailed above, when grades change, they change at most by one notch each quarter and very rarely move more than twice a year. From a yearly perspective, this implies that any prediction not within $\{ \textit{last year’s grade} \pm 1 \}$ is normally wrong.

From the last point, one can devise a simple and accurate benchmark for this classification problem. For every $(\textit{Country}, \textit{Year})$ entry, the *“constant”* predictor outputs the previous year’s grade for this country.

This predictor has high accuracy. Its errors’ amplitude also has very low volatility and are easily controlled as more-than-one-notch yearly changes are exceptional. The performance of this predictor should remain on average unchanged in the future as it does not suffer from data shift.

For a financial company, especially if the predictor’s output is to be used embedded in other systems, its reliability and virtual inexistent maintenance duties make it very practical.

Prediction for S&P Grades	Exact Match	Within 1 Notch	Within 2 Notches	Total
Linear Model	$\sim 20\%$	$\sim 65\%$	$\sim 85\%$	100%
Constant Model	77.1%	96.3%	98.6%	100%

Table 3.4: Constant Model prediction quality compared to typical legacy model. Constant Model outperforms significantly linear regression models.

Grades > 17 (*CCC*) are set to 17 so that no names are classified as "*in default*" without manual operation.

Historically Calibrated over the 2007 – 2017 period.

Grade moves of >2 notches were reduced to 2 notches. Grades to predict are in "*Previous Grade $\pm \{0, 1, 2\}$* ".

Redacted.

However this predictor holds no predictive power. Consequently, it holds no value to the bank other than as a benchmark. Indeed, as mentioned, a financial actor would value a model that yields insight on a grade’s direction, or simply on whether a change is imminent. Both as embedded sovereign credit grades predictors and as add-ons supporting Credit Risk analysts, better models than legacy ones that hold predictive power are required by sell-side institutions.

We first present a neural network model, trained on analogous data to the legacy models. Its higher accuracy than legacy implementations confirms that model more complex than regressions are suitable to tackle this predictive task.

3 NO HISTORY NEURAL NET MODEL

As an initial requirement, we want to build a neural network predictor on the same framework as current legacy models. This acts as a proof of concept, underlining the importance of non-linear feature combinations. More importantly, since it uses the same set of features as current models (*see Table 3.1*), it is simple to execute the model replacement: no data pipeline needs to be constructed, and the model’s output is similar to the previous one.

This shallow neural network with relatively few weights outperforms linear regression legacy implementations substantially. It achieves 65% perfect accuracy, and manages to get 80% of predictions with one notch tolerance. It however does not attain the accuracy exhibited by the "*constant predictor*". Thus, while it proves the well-soundness of using more complex machine learning than linear regressions to tackle the problem of predicting sovereign credit grades, a more developed version is required (*see Section 4*).

The set of features used by the internal legacy model is proprietary. In order to be able to present the section’s predictor’s results, another set of features is used. *Table 3.1* shows broadly which category of indicators each feature used by the legacy model is

included in. In this setting, we pick a representative feature for each of these categories. We use the resulting nine features to train our model, shown in *Table 3.5*.

Government Effectiveness Score
Labor Index
Sovereign Debt to Revenue
Economic Complexity Score
Number of past Defaults
Banking Index
GDP per Capita
Government Savings
Exterior Government Debt

Table 3.5: Features chosen to correspond to the nine ones currently in use. Each feature fits in the respective category presented in *Table 3.1*.

The previous section showed that countries not reported by data vendors were distributed unevenly, biased towards developing countries (*see fig. 3.4*). Among the countries reported however, this bias is less strong. That is, among countries which are reported upon by data vendors and internal Nomura systems, there is little systematic bias towards more reliable data on wealthy western nations (*see fig. 3.6*). This implies that the bias towards wealthy countries observed above is not entirely due to data unavailability. Indeed, shall data from developing nations be less available, it would be reflected on the developing countries included in our dataset too. Conversely, should data vendors only report countries for which data is reliably accessible, we would not observe a large proportion of included nations lacking data.

We thus opine that the absence of certain countries in widely used financial datasets is due to a lack of demand, not a lack of availability. Many of the entirely missing countries also are nations with no publicly traded debt. For third-party vendors, creating pipelines to report on dozens of indicators for new countries, especially ones where government reliability and digitization is not cutting edge, is costly. As a consequence, when their financial clients do not require data for these countries in particular, they are not reported upon and are simply omitted from the datasets. In other words, the omission of countries from the financial datasets is partly explained by the lack of exposure by large sell-side financial institutions to securities originating from these nations.

We then split the dataset in training and testing sets. As the dataset’s size is limited, we for now only keep 5% of entries in the testing set. Only intended as a proof of concept, we use a relatively small test-set size for this model. In the following Sections using larger datasets, we increase the size of the testing set to ensure results’ reliability.

The network is given as a tenth feature sovereign credit grades from the prior year.

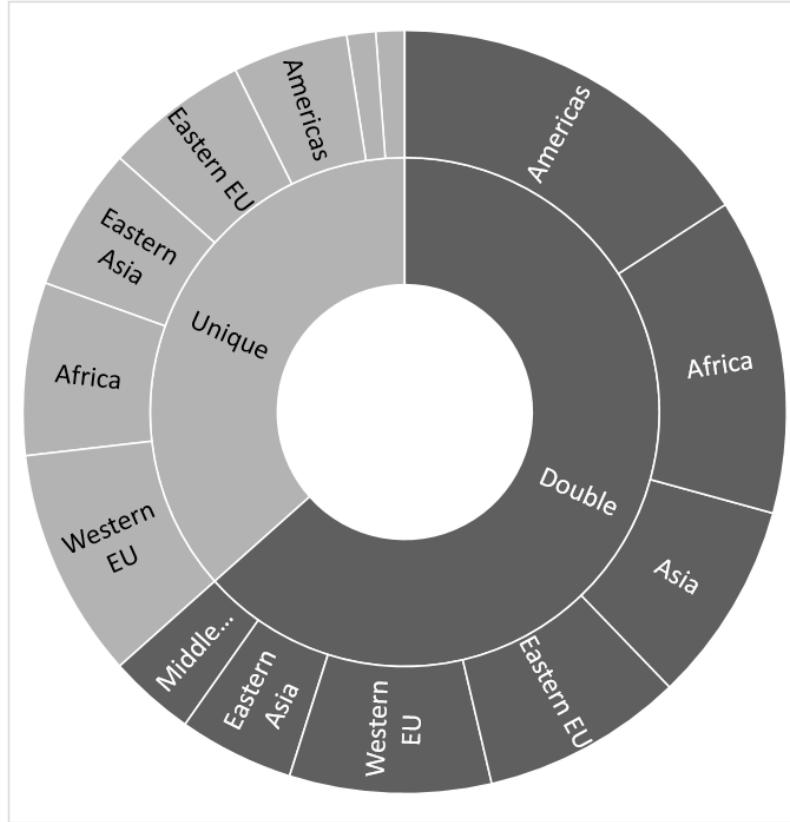


Figure 3.6: Data distribution within the 2018 & 2019 subset. While most countries available are so 2 years in a row (“Double”), many are not (“Unique”). Across 20 years, the relative presence of countries for which reliable data is available is increased.

The relative weight of developing and wealthy nations in the each of the two categories, “Unique” and “Double”, does not significantly differ.

E.g., it has access to any country’s grade in 2017 to predict its grade in 2018.¹⁰

Let us assume a given country has an entry in our dataset for both 2017 and 2018, and that its 2018 entry lies in the training set while its 2017 one is allocated to the testing set. While testing the model, it implies that the network has already seen the 2017 grade for the chosen country as part of its 2018 entry. Naturally, it is not obvious, given the relatively small size of the dataset, that the network will manage to exploit this leakage (Wang and Ruf, 2022).

As a consequence, the testing test is constructed so that no country appears both in the training and testing sets (see Section 4). In practice, leakage from entries to other entries is not exploited by the model in this setting. Whether we construct the testing set as described or not, or whether we only include each country once in the training set, the model behaves in a similar fashion. Such leakage however appears as an issue in the models with a larger scope presented in the next Section 4.

We tested multiple classical feed-forward neural networks of varying depth. We first

¹⁰As mentioned in Section 1, instances of Linear Models are used to set sovereign credit grades (Moody’s, 2016; Wilde, 1997). Contrary to setups designed to predict grades’ evolution, models calibrating credit grades are not given former grades as inputs.

find that deep neural networks, encompassing a large number of parameters, do not perform well and their training usually does not converge well. Indeed, the small size of our dataset restricts the number of parameters that the model can carry. This analysis is performed as no clear consensus exists yet in the finance field over which kind of Neural Network to use (Israel et al., 2020; Wang and Ruf, 2022).

Table 3.6 below summarizes the quality of the fit of various shallow networks.

Optimal networks have few hidden layers, ideally two, with a smaller amount of nodes in each. We settled on hidden layers of width 100 and 50 neurons respectively. This network already predicts Sovereign Credit Grades correctly with thrice the accuracy of a Linear Regression model. While its errors cannot be completely controlled within 2 or 3 notches of the grade to predict, it still manages 65% perfect accuracy and 80% within one notch accuracy.

Hidden Layers Structure	[100]	[1000]	[100, 50]	[100, 100]	[100, 100, 50]
Exact Match	45%	55%	65%	55%	40%
Within 1 Notch	80%	85%	90%	90%	90%
Within 2 Notches	90%	100%	95%	95%	100%
3+ Notches away	100 %				

Table 3.6: Quality of the out-of-sample fit of neural networks over the restrained dataset.

Each entry consists of 1 year of data only for one country, and each model needs to predict the Credit Grade for *(Country, Year)* pair.

The best performing networks' architectures in this Section as well as the subsequent ones in this Chapter perform only slightly better in-sample than out-of-sample. We assume this difference is driven by idiosyncratic sovereign events, which, in the testing set, are not labelled correctly.

Contrary to legacy models presented earlier, this design is able to take into account non-linear impacts of features on credit grades, as well as interplay between multiple features. While not all problems in Financial Economics benefit from implementing complex modern statistical models (Ruf and Wang, 2022), well chosen ones can gain from even small neural networks utilization (Bianchi et al., 2019). Indeed, even a straightforward approach such as this section's outperform classical models.

The next section empirically exhibits that incorporating more features to the model substantially improves accuracy. Also, observing a country's economic trajectory is necessary for the model to outperform the *"constant predictor"* benchmark. Such models use inputs of dimensions multiple times larger than this section's. In this case, linear regressions are theoretically ill-suited as the number of entries nears the number of features.

4 MULTIPLE YEARS NEURAL NET MODEL

Extending the set of used features, and multiple years of data per entry, we build a neural network predictor outperforming the *Constant Predictor*. Its eventual goal is to be embedded within other systems, with no human intervention. Its goal is thus solely

to maximise overall accuracy. The predictor presented predicts 80% of grades exactly, and 96.6% of grades within a notch of the true one.

The previous models do not incorporate any time structure in the inputs provided to the models. The large majority of indicators themselves do not capture information on the evolution of a country's financial and economic situation.

Consider two countries with the same ratio of foreign denominated sovereign debt to foreign reserves. The first country's ratio has been increasing sharply over the last years while the other's has remained stable over the same period. Assuming that these different dynamics do not impact any other indicator, the models presented above will not be able to differentiate these two countries despite their intuitively very distinct situations.

As a telling example, in 2020, Sri Lanka's foreign currency reserves were reported to be worth upwards of 8 *Billion* US Dollars (George et al., 2022; Sharma et al., 2022). This number remained relatively stable through the prior decade. Due to a combination of factors, reduced foreign income through tourism, increased price of vital imports such as food and fuel (sponsored by the state through a price pegging mechanism), Sri Lanka's current account was reduced (Blanchard and Milesi-Ferretti, 2010; Chinn et al., 2014; Obstfeld and Rogoff, 2005; Shaikh, 2016). As a result, by 2021, the state had halved its foreign currency reserves. By 2022, the foreign reserves had halved again, and Sri Lanka defaulted on 7 *Billion* US Dollars of foreign denominated sovereign debt. The worth of its entire 51 *Billion* US Dollars sovereign foreign denominated debt was affected. Even in 2021 however, the nation's foreign reserves represented 5% of its GDP. Indeed, according to World Bank data, in 2021, Sri Lanka's GDP was 88.2 *Billion* US Dollars¹¹.

In parallel, Luxembourg's foreign reserves in 2021 stabilised at 3 *Billion* US Dollars, 33% less than Sri Lanka's at the time. Its Gross Domestic Product was also similar to Sri Lanka's, 85.58 *Billion* US Dollars in 2021, according to the World Bank. Its sovereign external debt was half of Sri Lanka's, totalling 24 *Billion* US Dollars. Fitch's outlook for Luxembourg's sovereign credit grade was AAA, stable (Ratings, 2022).

In short, the three indicators considered, GDP, external debt and foreign reserves are unable to differentiate the two countries, on opposite ends of the sovereign credit grading scale.

Considering multiple years of data allows the model to easily differentiate the two nations presented as trends in the three viewed features are clearly distinguishable. Using other additional features is also an option. The following model implements both, using more features and multiple years of data for each entry.¹²

¹¹ Accessible at <https://data.worldbank.org>. Last accessed: April 2024

¹² We attempted to fit convolutional neural networks on our dataset. While obtaining robust con-

Multiplying the dimension of the input space naturally increases the size of the network. It is then harder for it to converge during the training. Yet, each entry containing multiple years can increase the size of our dataset. To this end, the new entries are constructed using the following process.

1 BUILDING THE DATASET

Every entry in the training and testing sets now has the following format:

Input:

$$\begin{aligned} \text{Feature 1} : & [Value(i_0), Value(i_0 + S), \dots, Value(i_0 + (WL - 1) \cdot S)], \\ & \dots \\ \text{Feature } M : & [Value(i_0), Value(i_0 + S), \dots, Value(i_0 + (WL - 1) \cdot S)], \end{aligned}$$

Output:

$$\text{Credit Grade}(i_0 + (WL - 1) \cdot S).$$

With the following definitions, represented in *fig. 3.7*:

- WL is the Window Length/Kernel Size, the number of datapoints utilised per indicator. Since time series' frequency is annual, WL controls for the number of years that are used to predict a country's grade. For instance, with $WL = 3$, the network uses 3 years of data, e.g. 2017 - 2019, to predict each country's sovereign credit grade in 2019.
- S is the stride, the frequency at which entries are fed to the network. For ($S = WL$), there is no overlap and no datapoint is included in two distinct entries. For $S = 1$, every (*Country*, *Year*) pair available in the dataset corresponds to one entry. Every yearly sovereign credit grade is an output to be predicted by the model.

Comparing the model's results with ($S = WL$) and ($S = 1$) permits easy detection of leakage in the ($S = 1$) case. Indeed, everything else equal, if the model's performance decreases considerably by solely removing the overlap, it indicates that the model's accuracy for ($S = 1$) is driven by the model's ability to calibrate on the leakage.

As detailed below, using a standard randomized train-test split actually allows

vergence of regular neural networks is not automatic, managing to stabilize the behaviour of *CNNs* has proven not yet feasible. Their structure would have ensured the correct temporal treatment of the input data (Wang and Ruf, 2022).

for the model to calibrate on the leakage, which we observed swapping ($S = 1$) for ($S = 3$) or ($S = 4$).

- M is the number of indicators used. In the previous *Section* 1 and 3, we had $M = 9$. *Appendix* 1 presents the large number of features available. It is not possible to use all nor a majority of reported features at once. Some have very low availability rates. So, there is a trade-off between the number of features and the size of the training set.

In a redacted process, by clustering features, we settle on ~ 20 features.

- i_0 is the starting year for the entry, and $(i_0 + (WL - 1) \cdot S)$ the last one passed as input.

We use $i_0 = 2000$, our dataset ends in 2021.

Table 3.7 presents the size of the resulting training set for typical (WL, S) values, aggregated from 2000 to 2021.

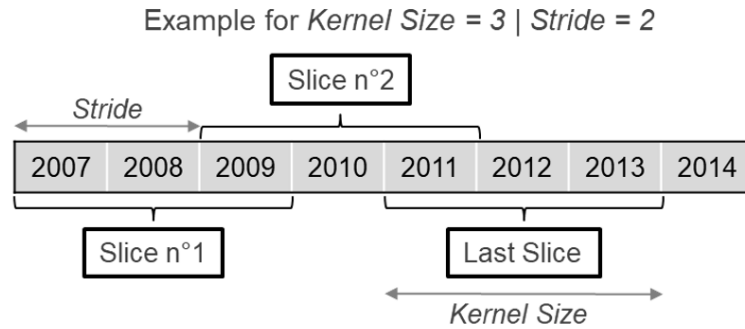


Figure 3.7: Time Series Data Slicing Process.

Kernel Size	1	2	3	4	5	6	7	8	9
Total # Entries	3322	3171	3020	2869	2718	2567	2416	2265	2114
Missing Grade Target	793	642	575	511	450	396	351	308	269
Usable	1003	934	851	770	689	610	534	463	395

Table 3.7: Impact of varying the kernel size WL on number of available entries in the dataset resulting from the described slicing process.

Highlighted is the value of WL used in production.

Max Fill (see below) is set to three, and stride S to one.

Before the data is sliced in the process described above, we reconstruct some missing entries to allow for more data to be used at once. Indeed, even standard institutional data sources such as the World Bank do not reliably provide data for countries that are usually reported on.

For credit grades themselves, our prediction target, we do not artificially recreate any values. However, we concatenate the different sovereign credit grades datasets at a

country level. That is, while we focus on the S&P grades globally, for nations rated exclusively by Nomura or Moody’s, these grades are provided as targets to the model. Filling individual years for which one agency’s grade is missing with another’s is not implemented. That is, despite having grades originating from different entities for different countries, each country’s grades only originate from an unique agency. Building a single country’s credit grade history from multiple sources would artificially add non-existent grade changes. Indeed, while S&P, Moody’s or Nomura grades are sometimes distinct, they usually move together in the same direction. Typically, the distinction lies in one agency rating a country consistently one notch above / below another agency’s rating. Thus, for a given country, replacing S&P’s credit grade by another agency’s when the former becomes unavailable would often yield a downgrade / upgrade, despite none of the individual time series exhibiting a grade change. To avoid this, we solely add entire sovereign credit grade time series missing from S&P’s dataset. In particular, public agencies generally only rate nations with publicly traded debt. Nomura however is exposed to non-marketable sovereign bonds but also to non-sovereign assets which risk is driven by the sovereign credit grade of the country the are tied to. These states are thus graded by Nomura, but not by public credit rating actors.

For the features of our model, the data reconstruction process is separated in three analogous operations. They are described thoroughly in *Appendix 2*. Essentially, each $(Country, Feature)$ is treated independently. The hyper-parameter $(Max\ Fill)$ controls for the maximal number of consecutive missing entries that we fill back for each time series. For instance, if $(Max\ Fill)$ is set to two, we only fill back missing values in time series less than two years away from at least one available entry.

Given $(Max\ Fill)$, each missing entry to be re-filled is calibrated as the weighted average of 3 linear trends, a future short-term one, a past short-term one and a global long-term trend.

For a given year T , the weight attached to the past short-term trend grows as the inverse of the number of consecutive missing entries in the time series prior to year T . Similarly, the weight attached to the future short-term trend grows as the inverse of the number of consecutive missing entries in the time series after year T .

The long-term trend’s weight is set so that the three weights sum to one.

As mentioned, the exact list of features used by the neural network is redacted. To assess the impact of outlined back-filling process, we exhibit in *Table 3.8* its impact on the training set built random forest predictor. This dataset is comprised of the thirteen features shown in *Table 3.11*. Setting $(Max\ Fill)$ to three doubles the effective number of entries in the training set. For low values of $(Max\ Fill)$, the number of entries added for each $(Max\ Fill)$ increment is large. Many “holes”, time series missing few entries in between available dates, are reconstructed. As $(Max\ Fill)$ increases further,

each (*Max Fill*) increment adds ~ 80 entries. These entries refer to adding data to the oldest point of the time series. For (*Max Fill*) larger than three, the back-filling process essentially recreates data before any was available for each given country and feature. As this is not the target of the back-filling, intended to provide a solution to the data's unreliability, and not its limited range, we fix (*Max Fill*) to three. *Table 3.9* shows the final distribution of grade changes in the engineered dataset.

<i>Max Fill</i>	0	1	2	3	4	5	6	7	8	9
Usable Entries	466	618	770	851	935	1014	1090	1160	1228	1295

Table 3.8: For varying maximal filling size, number of available entries in the resulting dataset. The largest jump in the number of available entries is swapping (*Max Fill*) from zero to one, from the large number of sporadically unavailable features' entries.

Highlighted: value of (*Max Fill*) used in production.

Slicing features - *Kernel Size: 3, Stride: 1 – Total Entries, with NaNs: 3020.*

Grade Change	<i>WL</i> = 3		<i>WL</i> = 4	
	Total	Usable	Total	Usable
0	1884	652	1820	589
± 1	438	162	420	89
± 2	73	22	71	18
± 3	15	4	14	3
± 4	11	2	7	2
± 5	18	8	18	8
± 6	7	1	3	1
± 7	2	0	2	0
± 8	3	0	2	0
± 9	1	0	1	0
NaN	575	0	511	0

Table 3.9: Grade changes distribution of the whole reconstructed dataset. We observe the imbalance of the classes, the no change class representing 76% of the final dataset with both (*WL* = 3) (652 of 851 entries) and (*WL* = 4) (589 of 770 entries).

In grade changes strictly larger than 1 notch, there is significant imbalance between upgrades and downgrades both in the original dataset and in the processed one. Negative change represent 70% > 1 notch updates (for (*WL* = 3), 91 out of 130 in the uncleaned dataset and 26 out of 37 for the final one). This imbalance is in part due to the imbalance between jump-to-defaults and relistings. It is also caused by more frequent consecutive quarterly downgrade updates in stressed situations, than the opposite.

Left: *Kernel Size: 3 – Stride: 1 – Max Fill: 3* & **Right:** *Kernel Size: 4 – Stride: 1 – Max Fill: 3*

Finally, as noted above, without careful separation of the training and testing sets, the network carries *look ahead* bias and artificially high accuracies are observed. As mentioned in the previous section, using multiple years at once introduces further data leakage.

The described dataset structure means that without consideration, the network has seen in the training set a credit grade as an input that it might have to predict as an

output in another entry. Even in the absence of potential leakage, taking into account the temporal structure in processing is not straightforward (Hakami, 2024; Wang and Ruf, 2022).

In our setting, the issue is not only that the network has seen a grade it needs to predict in a previous train example. The model is also able to differentiate each distinct $(Country, Year)$ entry. Indeed, having 20 features for each $(Country, Year)$ entry, no two entries share identical sets of indicators, and many carry a distinctive feature. Since the model is also given as input the history of credit grades of the relevant country, the network in practice develops a way to separate all $(Country, Year, Grade)$ entries. It is however strictly unable to carry predictions on data unseen before, solely relying on already having seen each grade it is required to predict in its training set.

Initially, we randomly separated the entries in training and testing sets. We then randomly shuffled the entries allocated to the training set and let the network calibrate its weights on these entries.¹³ Observing the results of this calibration process on the separated test set, the model has nearly 100% accuracy (where the non perfect accuracy is driven by instances of predictions for 2021 assigned to the test set. As 2021 is the last available year in our dataset, the network has not yet seen the grade to predict in another entry). However, we found if we first isolate some arbitrarily chosen countries from the training phase, testing the model after the same training on the rest of the data, its performance is very low, even regularly outputting impossible values. The same substantial drop in accuracy is observed, though less significant, if we isolate the last year of data from the training entirely, instead of individual nations.

To ensure we do not miss such instances of overfitting / data leakage, we split the training and test sets so that each country appears exclusively in one of the two. Alternatively, we can keep in the training set the last year of data available. The latter implementation allows for the replication of real model use once in production: the model needs to predict grades having calibrated its features on previous years of data. It is however more challenging, since regular supra-national events alter the way credit agencies grade nations. Covid, as an example saw many sovereign actors experience situations which would have triggered a credit grade update should the crisis have impacted an unique actor. Thus, we are formally not testing the model on a set representative of the data it has seen in training. Similarly, assigning separate countries in the testing set formally assesses the model’s performance on assets distinct from the assets included in the training set, making the problem harder.

These processes however, while permitting the user to spot the overfit, do not provide a solution the *look ahead* from training.

¹³We have implemented multiple batch sizes, from very low ones the highest our available computing power allowed us. No remarkable change in performance was observed (Keskar et al., 2016).

To this end, we analogously stratify the train set separately to ensure that it has never seen a grade to predict earlier in training. We stratify either by nation first, then by year, or the reverse, yielding the same predictive power presented in *Table 3.10*.

2 MODEL RESULTS

In the results presented below, for each entry, we use 4 years of data (e.g. 2012-2015) to predict the last year’s grade (e.g. 2015). The results obtained are significantly better than the results of the previous models. Being able to observe a country’s financial and economic history is indeed helpful towards sovereign credit grade prediction. Setting the kernel size to 3 or 4 yields qualitatively similar results, shorter and longer horizons hinder performance. We opine that longer horizon, multiplying the dimension of the feature set makes the calibration less straightforward. Meanwhile it reduces the effective number of entries, and consequently reduces the accuracy substantially. As in the previous section’s model, two datapoints per indicator is not enough to capture a country’s trajectory.

Prediction for S&P Grades	Exact Match	Within 1 Notch	Within 2 Notches	Total
Linear Model	$\sim 20\%$	$\sim 65\%$	$\sim 85\%$	100%
Constant Model	77.1%	96.3%	98.6%	100%
Multi-Year Neural Network	80%	96.6%	100%	100%

Table 3.10: Multi-Year Neural Network model results outperform the results from the constant model, meeting its benchmark target.

Grades > 17 (*CCC*) are set to 17 so that no names are classified as “*in default*” without manual operation.

Historically Calibrated over the 2007 – 2017 period.

While the models do not replicate the behaviour of the “*constant predictor*”, their accuracy over the minority class is low.

All presented multi-year neural networks use a ReLU activation function, a quadratic loss function (implementation of Huber and Linear losses did not impact performance), and 10 training epochs.

Redacted.

The multi-year *NN* implementation beats all other previous models and meets the “*constant*” model benchmark, using the same two-layered structure as the “*no-history*” model (*see Table 3.10*). It does not replicate the constant predictor either.

However, it is not an ideal predictor. Firstly, its convergence is not perfectly stable: calibrating the network shuffling the training set in various acceptable orders does not yield the same results, as it is common to observe nets’ training not converging and predicting impossible results.

Furthermore, its accuracy on the “*no-change*” class beats its total accuracy. This is expected as the “*no-change*” class is largely more frequent than the other ones: a large increase in accuracy in the “*change*” class does not compensate for the associated

small decrease in accuracy in the large class. Multiple attempts to solve this issue were unsuccessful. Enforcing strict stratification so that the model initially sees “*change*” entries at the same rate as “*no-change*”, is not enough. Swapping target classes from $\{1, \dots, 20\}$ to $\{0, 1\}$ to the presented $\{-2, -1, 0, 1, 2\}$ also does not modify the imbalance in resulting quality.

As such, we opine that while the imbalance in the size of the two classes is a large driver of the difference in prediction quality, it is not the sole one. The stratification process just outlined would have helped toward reducing the performance gap shall size imbalance had been the only factor.

We argue that despite the “*no-change*” class being preponderant, the “*change*” class contains an outsized diversity of scenarios. A large majority of instances of the “*no-change*” class in practice represent (*Country, Year*) pairs during which no significant financial nor economic event happened. By contrast, every entry in the “*change*” class represents a more convoluted national scenario, which justified an upgrade, or more frequently a downgrade (see Table 3.9).

The aim of this predictor was solely accuracy, which is satisfied as this implementation trumps all previous ones, both internal and in published research. For a model used in a systematic fashion without human review, overall performance is the most important metric. Yet, for Credit Risk analysts, who manually review the output of the model, high accuracy on each class, and especially over the “*change*” class, is central. To satisfy this requirement, we now present a random forest classifier, built to minimize the Type-II Error, i.e. missing a grade change.

5 RANDOM FORESTS

The random forest predictor presented in this section is intended to be used by Credit Risk analysts when countries are periodically reviewed. The model is designed as a tool to reduce the pool of nations over which human intervention has to focus. As described in the introduction, it is crucial for the bank to flag grades which are going to change. It is substantially more costly to miss a grade being modified than to mistakenly label a (*Country, Year*) pair as a “*change*”. In parallel, contrasting with the neural networks approaches presented earlier, random forests allow for easy analysis of the drivers of a classification. This is also useful to manual reviewers as they can quickly focus on the indicators that are expected to drive a change in credit grade.¹⁴ As we will describe in more details in this part, we manage to construct a predictor 80% accurate on the “*change*” class, which remains 92% accurate over “*no-change*” class. While naturally not perfect predictors, the quality of these results is promising.¹⁵

¹⁴The analysis of the features driving the changes is redacted.

¹⁵As the goal of the neural networks predictors (see Section 3, 4) was to maximise overall accuracy only, we focused on their performance across the two classes.

Since this model is to be used by Credit Risk, this Department prescribed the list of indicators we input in the model. We however remove a few indicators from the list shared as their availability is too low (*see Table 5*). This list is composed with two purposes. It is assembled so that it contains the features flagged by credit rating agencies as essential in their rating process. It also includes indicators that are considered the most relevant to actual grade updates by Credit Risk. This second category is used by Nomura in the review of its own internal grades, and we presume that they are used by public rating agencies too.

Age Dependency Ratio (%)
Broad Money to Reserves
Economic Complexity Score
External Debt (% of GDP)
External Debt (% of Reserves)
General Government Debt (% of Revenue)
General Government Interest (% of Revenue)
Government Effectiveness
Inflation (Year-on-Year)
Local Currency Share of Debt
Net Inward FDI (% of GDP)
Rule of Law Index
Savings (% of GDP)

Table 3.11: Features chosen by Credit Risk as the most relevant to predict Sovereign Credit Grades.

As described, from the list of provided features, we removed the following five as their availability rate is too low to be efficiently used in calibration (<25% of countries are reported):

- Capital Adequacy Ratio,
- Local Currency Share of Debt,
- Non-Performing Loans (NPLs),
- Real Effective Exchange Rate (%) (REER),
- Short-Term External Debt (% of Reserves).

Conversely, to the given set of features, we add the following two, created from the values of *Inflation (Year-on-Year)*, available at a monthly frequency.

- Consumer Price Inflation (CPI) Volatility,
- Consumer Price Inflation (CPI) Rolling change.

Table 3.12 confirms that this list of provided indicators indeed is strongly correlated to time series of sovereign credit grades. Naturally, as non-linear interplay between

various indicators is tantamount to the value of our model, correlations only give partial information. *Table 3.12* particularly highlights that, as expected, indicators with the highest absolute correlation to sovereign credit grades describe the financial structure of the country in regard to its debt: *External Debt*, *Servicing Costs*, *General Government Reserves*, *Non-Performing Loans*. Yet, social structural indicators such as *Rule of Law* are also correlated to sovereign credit grades.

We set *WL* to three for the rest of the Chapter, yielding in total 851 entries in our reconstructed dataset.

Feature:	Corr. to Sov. Credit Grades	
	Original Set	Reconstructed
Age Dependency Ratio (%)	-0.02	-0.0
Broad Money to Reserves	-0.03	-0.02
Capital Adequacy Ratio	-0.22	-0.22
CPI Inflation - Change	-0.03	0.01
CPI Inflation - Volatility	-0.0	-0.04
Economic Complexity Score	0.07	0.07
External Debt (% of GDP)	-0.28	-0.27
External Debt (% of Reserves)	-0.25	-0.26
General Government Debt (% of Revenue)	-0.42	-0.43
General Government Interest (% of Revenue)	-0.12	-0.11
Government Effectiveness	0.21	0.23
Inflation (% year-on-year)	0.02	0.02
Local currency share of debt	0	0.04
Net Inward FDI (% of GDP)	0.13	0.14
Non-Performing Loans	-0.23	-0.24
Real Effective Exchange Rate (%)	0.05	0.04
Rule of Law	0.14	0.15
Savings (% of GDP)	0.16	0.16
Short-term Ext Debt (% of Reserves)	-0.04	-0.07

Table 3.12: Average correlation of features' time series to time series of sovereign credit grades. For each feature, we compute individually the Pearson correlation (Similar results are found using Spearman rank correlation instead) between each country's credit grade evolution and corresponding feature values. For each feature, results presented are averaged across available countries.

Our back-filling process does not alter the correlation structure significantly. *Max Fill* was set to three. Also, while some correlations exhibit near independence, some are highly (anti)-correlated with the target.

Since we are to implement a random forest classifier, as we changed the set of indicators, we also change the structure of the features given as inputs to the model.¹⁶ For each indicator, we replace the *WL* consecutive years of data by their average increment (trend), average squared increment (volatility) and last available value across the *WL* years in the window. Since *WL* is set to three, this process does not modify the

¹⁶Each node in each decision tree from the random forest is only able to split entries depending on one unique feature being larger or not than a calibrated cutoff value. As such, inputting *WL* years of data per indicator is not sufficient for the model to calibrate on, e.g., the trend or volatility of the indicator, and is thus detrimental to its performance.

dimension of the input space. Though, for larger values of WL , this procedure reduces the dimension of each entry.

Table 3.13 outlines the optimal number of trees in the forest, and Table 3.14 presents the optimal values of other hyper-parameters.

Trees in Random Forest	Accuracy per Class	
	“no-change”	“change”
50	86.5%	<50%
100	81.6%	54.4%
200	82.88%	53.76%
500	85.12%	53.28%
1000	83.84%	52.48%

Table 3.13: Average Random Forest performance per class for varying number of individual decision trees in the forest.

To obtain these figures, we calibrated random forests over grids of acceptable other hyper-parameters, in particular (*Minimum Sample Split*, *Share of Features per Tree*, *Share of Samples per Tree*). Our grid is voluntarily not exhaustive. We constraint tuples of hyper-parameters in the grid to be near optimal. Thus, we do not use sub-optimal classifiers to assess the adequacy of each potential value for the number of trees in the forest.

We then average the results over all tuples of hyper-parameters in the grid to obtain these figures.

As presented below, the average performance, even with bad values of hyper-parameters excluded, is remarkably lower than the optimal accuracy.

Table 3.14 presents the accuracy of the random forest predictor. Minimizing the Type-II Error (see Appendix 3), the classifier attains 80% accuracy over the “change” class, while maintaining over 90% accuracy over the “no-change” class.

Minimization Target	Equal Weights	Type-II Error	Type-I Error
Weight “no-change”	1	1	5
Weight “change”	1	5	1
<i>Share of Samples per Tree</i>	0.5	0.7	0.7
<i>Share of Features per Tree</i>	0.7	0.7	0.7
Results			
Accuracy over “change”	88%	92%	96%
Accuracy over “no-change”	76%	80%	48%

Table 3.14: By assigning unequal weights to the two kinds of errors, we can shift the model's performance. *Minimum Sample Split* is set to eight, smaller values introduces overfit: the train accuracy increases but *OOB* and test accuracies are not improved.

A clear optimal of 0.7 arises for *Share of Features per Tree*. Values of *Share of Samples per Tree* is less clear above 0.2 result in correct accuracy. Its optimal is raised from 0.5 for *Equal Weights* to 0.7 for unequal weighting.

Minimizing the Type-I Error, letting *Share of Features per Tree* increase to 0.75, the Type-I Error rate drops to near 0. The Type-II Error rate then is 40%.

6 CONCLUSION

In this Chapter, we presented three models attempting to predict the evolution of Sovereign Credit Grades. Each model is intended to serve a distinct purpose within the systems and functioning of Nomura International Securities.

We highlighted the need for such predictors: the impact of sovereign credit grades changes on banks' positions is widespread and momentous. Indeed, sovereign bonds are a central driver of risk in Prime Brokerage, Market-Making but also Wholesale Banking in its entirety. Sovereign bonds' risk levels are in large part driven by their underlying credit grade. Consequently, better prediction of nations' credit grades, even offering only slight improvement over current models, is of high value to all major financial institutions.

Our first model, a neural network using only one year of financial, social and economic data to infer credit grades, uses the same inputs as legacy models. It largely outperforms current implementations. Its development was carried out to demonstrate that significant improvements in the prediction of sovereign credit grades can be made. Its results emphasize the need for non-linear and modern statistical methods in forecasting sovereign credit grades. As it uses the same data pipeline as existing model, their replacement by this model would already improve the adequacy of risk limits and haircuts applied by financial institutions.

The other two classifiers presented attempt to attain a level of accuracy rendering their use in production by the bank worthwhile. The first, a multi-year multi-layered perceptron, aims at overall performance only. Its accuracy surpasses not only all previous published implementations but is, to our knowledge, the first model to exceed the predictive ability of the “*constant predictor*”, which simply always predicts that a grade change will not happen. Implementing this model, error rates compared to the simple “*constant predictor*” benchmark are reduced by 25%. Also, as for the “*constant predictor*”, when our predictor's outlook is wrong, the magnitude of the mistake is small. Even without any further work on this model, using it in pricing and risk engines, replacing the use of the “*constant predictor*” or legacy models would entail a substantial increase in efficiency.

Our last classifier, a random forest, attempts to minimize the Type-II Error: mistakenly labelling an entry as “*no-change*” and thus missing a grade change. Indeed, the bank's is asymmetrically exposed to mislabelling of grades' outlooks. It is more costly to miss a grade change than to assume one will change, despite it not being altered in the end. Solving this issue is more challenging.

Firstly, given an arbitrarily chosen (*Country*, *Year*) pair, the probability that its grade is identical to its previous one is thrice higher than it is of being different. Separating

the *(Country, Year)* entries in our dataset between “*change*” and “*no-change*”, the classes derived are heavily imbalanced. The “*no-change*” class contains 77% of entries. We thus have only relatively few instances of sovereign credit grades changing. As a result, on average, making a mistake over the “*no-change*” class is three times more costly than on the “*change*” class. Yet, we argued that the class imbalance is only part of the challenge of reducing the Type-II Error. Indeed, careful stratification ensuring the two classes are weighted in the same fashion yielded qualitatively similar results. As one class is more populated, its diversity is also smaller. A myriad of scenarios entail a credit grade melioration, and more frequently, deterioration. Many distressed financial and economic situations only occurred once. By contrast, the “*no-change*” class is composed of numerous instances of countries where volatility of all the given features is low. We argue that it is inherently harder to precisely anticipate grade modifications due to the relatively higher potential number of such scenarios compared to the limited numbers of actual instances included in our historical datasets.

Despite the arguments outlined, the random forest classifier performs relatively well on both classes, managing to correctly identify at least 80% of the members of each label. It will soon officially be used by Credit Risk analysts as a support system in their periodic sovereign credit grades outlook reports.

This Chapter presented advancements in modelling sovereign credit grades. As the arrangement of the manuscript indicates, the largest share of the work and its main challenges lays in in sourcing, pre-processing and carefully structuring our dataset prior to training any statistical engine over it. While more work could be carried out in these directions, the proper modelling part could also be improved. The two types of models used, neural networks and random forests, are standard approaches in modern statistics. More modern methodologies appear applicable to the prediction task at hand, with short time series and a limited number of features. Signatures in particular are a promising way to capture intricate interplay between time series (Kidger et al., 2019; Lyons, 2014; Salvi et al., 2021). For an introduction on signature kernels, see (Chevyrev and Kormilitzin, 2016; Lyons and McLeod, 2022; Lyons and Xu, 2014). Alternatively, randomized signatures could yield a performance boost (Teichman et al., 2021). In parallel, while we focused on sovereign credit grades in this Chapter, our analysis could be extended to corporate names. Indeed, sovereign grades often act as a ceiling for corporate grades: more accurate prediction of sovereign credit grades can help directly in the modelling of their corporate counterparts. Furthermore, the models described in this Chapter could be adapted to be able to run on corporate data, thus providing a novel way of predicting corporate credit grades, hopefully with an analogous boost in accuracy as the one presented in this Chapter.

A APPENDIX

1 AVAILABLE INDICATORS

Government Effectiveness Rank
Labour Market Efficiency (Rank)
Government Debt to Revenue
Economic Complexity Score
Default History
Soundness of Bank (Rank)
Per Capita GDP
Savings as % of GDP
External Debt to GDP
5-year Real GDP Growth Average
Broad Money to Total Reserves
Corruption Control (Rank)
Credit Growth over GDP Growth
Current Account as % of GDP
Capital Adequacy Ratio (%)
Monetary Union Member (Binary)
Own Currency Should be Reserve (Binary)
History of Capital/Exchange Control (Binary)
Debt Majority in own float Currency. (Binary)
Fiscal Balance as a share of GDP
Government Effectiveness Estimation
Government External Debt to total External Debt
Government External Debt to Reserves
Human Development Index (Rank)
Inflation Level (%)
Inflation Volatility
Political Stability & No Violence (Rank)
Number of sectors with more than 10% of GDP
Number of Sectors with more than 20% of GDP
Number of Export Items with more than 20% of total
Property Rights (Rank)
Per Capita Income Level (US\$)

Table A.15: Complete List of Available Features concatenated from various data vendors. **Part I.**
Redacted.

Affordability of Financial Services (Rank)
Business Cost of Crime (Rank)
Business Cost of Terrorism (Rank)
Availability of Financial Services (Rank)
Economic Complexity (Rank)
Foreign Market Size (Rank)
Higher Education & Training (Rank)
Labour Market Efficiency (Rank)
Legal Rights Index (Rank)
Private Institutions (Rank)
Quality of Infrastructure (Rank)
Reg. of Securities Exchanges (Rank)
Public Institutions (Rank)
Technological Readiness (Rank)
Undue Influence (Rank)
World Press Freedom Index (Rank)
Regulatory Quality (Rank)
Rule of Law (Estimate)
Rule of Law (Rank)
Short Term Debt as % of total reserves
Short Term Ext. Debt (US\$) Short Term Debt in Government Debt
Size of Economy (US\$)
Tax Revenues as % of GDP
Total Population
Transparency of Policy Making
Unemployment Rate
Working Age Population (Aged 15-64)
Years Since Last Sovereign Default
Credit Grade Diff. Between Sovereign
Debt Denominated in Local Currency and Foreign Currency

Table A.16: Complete List of Available Features concatenated from various data vendors. **Part II.** *Redacted.*

2 DATA RECONSTRUCTION

Let us denote $V := \{V_t, t \in [2000, 2021]\}$ the time series corresponding to an indicator's reported values for an individual country, with yearly frequency.

A missing entry in V for year t is denoted $V_t = \emptyset$.

Firstly, if $\forall t \in [2000, 2021], V_t = \emptyset$, then the reconstructed time series is also empty: for each country, entirely missing features are left empty. In the following, we thus assume that:

$$\exists t \in [2000, 2021], \text{ s.t. } V_t \neq \emptyset.$$

For $t \in [2000, 2021]$, let us define T^+ and T^- as functions of t :

$$T^+(t) = \begin{cases} \infty, & \text{if } V_s = \emptyset, \forall s \in [t, 2021], \\ \min \{s \in [t, 2021], V_s \neq \emptyset\}, & \text{else.} \end{cases}$$

$$T^-(t) = \begin{cases} -\infty, & \text{if } V_s = \emptyset, \forall s \in [2000, t], \\ \max \{s \in [2000, t], V_s \neq \emptyset\}, & \text{else.} \end{cases}$$

Given $t \in [2000, 2021]$, let us now assume $V_t = \emptyset$.

Given $(Max\ Fill) \in \mathbb{N}^+$, let us first determine whether V_t is to be back-filled or not.

Let $\hat{V} := \{\hat{V}_t, t \in [2000, 2021]\}$ be the reconstructed time series.

$\forall t \in [2000, 2021]$ we have:

If $Max\ Fill \geq \min(|T^+(t) - t|, |t - T^-(t)|) > 0$, the value of $\hat{V}_t$ is reconstructed.

Else, $\hat{V}_t = V_t$.

As above, let us now assume that we have $t \in [2000, 2021]$ such that $V_t = \emptyset$, and that V_t is to be back-filled under the provided $(Max\ Fill)$.

- If $T^+(t) < \infty$ and $T^-(t) > -\infty$:

$$\text{We define } \gamma_L = \frac{\mathbb{1}_{\{T^-(t)=t-1\}}}{2 + \mathbb{1}_{\{T^+(t)=t+1\}}} \frac{T^+(t) - t}{T^+(t) - T^-(t)},$$

$$\text{and } \gamma_R = \frac{\mathbb{1}_{\{T^+(t)=t+1\}}}{2 + \mathbb{1}_{\{T^-(t)=t-1\}}} \frac{t - T^-(t)}{T^+(t) - T^-(t)}.$$

Then:

$$\hat{V}_t = \frac{V_{T^-(t)}(T^+(t) - t) + V_{T^+(t)}(t - T^-(t))}{T^+(t) - T^-(t)} + \gamma_L \cdot (V_{T^-(t)} - V_{T^-(t)-1}) + \gamma_R \cdot (V_{T^+(t)} - V_{T^+(t)+1}).$$

The first term is a linear interpolation between the two observations closest to the year being reconstructed. γ_L and γ_R terms respectively represent the impacts of future and past short-term trends on the current value. When $t = T^-$ (*resp.* $t = T^+$), V 's tendency prior to V_{T^-} (*after* V_{T^+}) is into account as $(V_{T^-(t)} - V_{T^-(t)-1})$ (*resp.* $(V_{T^+(t)} - V_{T^+(t)+1})$).

- Else, if $T^+(t) < \infty$:

$$\hat{V}_t = V_{T^+(t)} + (V_{T^+(t)} - \bar{V}) \cdot \frac{T^+(t) - t}{\max\{s \in [2000, 2021], V_s \neq \emptyset\} - T^+(t)} + \frac{V_{T^+(t)} - \frac{V_{T^+(t)} + V_{T^+(t)+1} + V_{T^+(t)+2}}{3}}{T^+(t) - t}.$$

- Else, when $T^-(t) > -\infty$:

$$\hat{V}_t = V_{T^-(t)} + (V_{T^-(t)} - \bar{V}) \cdot \frac{t - T^-(t)}{T^-(t) - \min\{s \in [2000, 2021], V_s \neq \emptyset\}} + \frac{V_{T^-(t)} - \frac{V_{T^-(t)} + V_{T^-(t)-1} + V_{T^-(t)-2}}{3}}{t - T^-(t)}.$$

In the two latter cases, where $T^+(t) = \infty$ or $T^-(t) = -\infty$, the reconstructed values are computed starting from the last originally available value in the time series, $T^+(t)$ and $T^-(t)$ respectively. That is for instance where $T^+(t) = \infty$, we start by back-filling $(T^+(t) - 1)$, and only then $(T^+(t) - 2)$ if (*Max Fill* ≥ 2). This ensures that the terms of the rolling average $V_{T^+(t)}$, $V_{T^+(t)+1}$ and $V_{T^+(t)+2}$ ($V_{T^-(t)}$, $V_{T^-(t)-1}$ and $V_{T^-(t)-2}$ respectively) are available at each step of the back-filling process.

In edge cases, when $(T^+(t) + 1)$ or $(T^-(t) - 1)$ are not well-defined for instance, their associated weight γ is set to zero and the long-term trend term γ_G is accordingly increased.

3 CLASS WEIGHTS IN RANDOM FORESTS

The goal of the Random Forests implementations is to ensure that good overall performance also translates to satisfactory accuracy over the minority “*change*” class. To maximise the performance over the “*no-change*” class and minimize the Type-II Error,

we increase the weight of the minority class to five.¹⁷ In the previous results, we used an equal weighting: mislabels over each class carried the same cost.

In practice, to increase a class' weight, we modify the function I assessing the nodes' purity.

Let us consider a node N in a decision tree. The node's population is split in two classes $\{\text{"change"}, \text{"no - change"}\}$. Their respective ratios in the node's population are $p_N(\text{"change"})$ and $p_N(\text{"no - change"})$. Since there are only two classes, we have:

$$p_N(\text{"no - change"}) = 1 - p_N(\text{"change"}).$$

From this, we define the node's entropy impurity as follows:

$$I(N) = - \sum_{\chi \in \{\text{"change"}, \text{"no-change"}\}} \omega(\chi) \cdot (p_N(\chi) \ln(p_N(\chi))).$$

When splitting a node in children nodes, the separation is determined to minimize the aggregate entropy of the resulting children nodes.

By setting $\omega(\text{"change"})$ to five instead of one, we incentivize the construction of individual decision trees to classify correctly the $(Country, Year)$ pairs labelled as *"change"*.

¹⁷We are thus overweighting the *"no-change"* class above its balanced weight. Its balanced weight is two (or four, if the weight of the *"change"* class is normalized to one).

Indeed: $\frac{Dataset\ Size}{\{\# \text{"change"}\} \cdot \{\# of\ classes\}} = \frac{851}{(0.24 \cdot 851) \cdot 2} \approx 2.$

REFERENCES

- Afonso, A. (2003). Understanding the determinants of government debt ratings: Evidence for the two leading agencies. *Journal of Economics and Finance*, 27(1):56–74.
- Afonso, A., Furceri, D., and Gomes, P. (2012). Sovereign credit ratings and financial markets linkages: Application to european data. *Journal of International Money and Finance*, 31(3):606–638.
- Afza, T. and Mirza, H. H. (2011). Institutional shareholdings and corporate dividend policy in pakistan. *African Journal of Business Management*, 5(22):8941–8951.
- Agirman, E., Sercemeli, M., and Ozcan, M. (2013). Shadow banking: An overview.
- Aharon, D. Y. and Ali, S. (2024). A high-frequency data dive into svb collapse. *Finance Research Letters*, 59(104823).
- Aharony, J. and Swary, I. (1980). Quarterly dividend and earnings announcements and stockholders’ returns: An empirical analysis. *The Journal of finance*, 35(1):1–12.
- Akbas, F., Boehmer, E., Jiang, C., and Koch, P. D. (2022). Overnight returns, daytime reversals, and future stock returns. *Journal of Financial Economics*, 145(3):850–875.
- Allen, F., Bernardo, A. E., and Welch, I. (2000). A theory of dividends based on tax clienteles. *Journal of Finance*, 55:2499–2536.
- Allen, F. and Michaely, R. (2003). Payout policy. In *Handbook of the Economics of Finance*, volume 1, Part 1, chapter 07, pages 337–429. 1 edition.
- Alletzhauser, A. (1990). *The House of Nomura: The Inside Story of the Legendary Japanese Financial Dynasty*. Bloomsbury Publishing.
- Almeida, H., Fos, V., and Kronlund, M. (2016). The real effects of share repurchases. *Journal of Financial Economics*, 119(1):168–185.
- Almgren, R. and Chriss, N. A. (1997). Optimal liquidation. *Journal of Economic Perspectives*.
- Alonso Olmedo, J., Anguren Martín, R., Gamoneda Roca, M., and Pérez Rodríguez, P. (2021). Archegos and greensill: collapse, reactions and common features. *Financial Stability Review / Banco de España (Autumn 2021)*, pages 47–62.

- Alquhaif, A., Al-Gamrh, B., and Abdul Latif, R. (2020). An overview of share buybacks: A descriptive case from malaysia. *Journal of Behavioral and Experimental Finance*, 28:100415.
- Altman, E. I. (1998). The importance and subtlety of credit rating migration. *Journal of Banking and Finance*, 22:1231–1247.
- Altman, E. I. and Kishore, V. M. (1996). Almost everything you wanted to know about recoveries on defaulted bonds. *Financial Analysts Journal*, 52.
- Ambarish, R., John, K., and Williams, J. (1987). Efficient signalling with dividends and investments. *The Journal of Finance*, 42(2):321–343.
- Anand, A., Baesens, B., and Vanpee, R. (2023). Sovereign credit risk modeling using machine learning: a novel approach to sovereign credit risk incorporating private sector and sustainability risks. *Journal of Credit Risk*, 19.
- Andres, C., Doumet, M., Fernau, E., and Theissen, E. (2015). The lintner model revisited: Dividends versus total payouts. *Journal of Banking & Finance*, 55:56–69. Global Governance and Financial Stability.
- Aramonte, S. (2020). Mind the buybacks, beware of the leverage. *BIS Quarterly Review*.
- Aristotle (BC). *The Politics*. Penguin Classics.
- Ariyo, A. (1995). Reliability of macroeconomic data: An evidence from nigeria’s debt data series. *African Development Review - Revue Africaine de Developpement*.
- Asness, C. S., Hazelkorn, T. M., and Richardson, S. A. (2017). Buyback derangement syndrome. *Journal of Portfolio Management*, *Forthcoming*.
- Asquith, P. and Mullins Jr, D. W. (1983). The impact of initiating dividend payments on shareholders’ wealth. *Journal of business*, pages 77–96.
- Asquith, P. and Mullins Jr, D. W. (1986). Signalling with dividends, stock repurchases, and equity issues. *Financial management*, pages 27–44.
- Atanasov, V. (2005). How much value can blockholders tunnel? evidence from the bulgarian mass privatization auctions. *Journal of Financial Economics*, 76(1):191–234.
- Atsushi, I. and Lutz, K. (2004). Bagging time series models.
- Augustin, P., Subrahmanyam, M. G., Tang, D. Y., and Wang, S. Q. (2016). Credit default swaps: Past, present, and future. *Annual Review of Financial Economics*, 8:175–196.

- Avazzadeh, E. (2015). The effect of corporate governance components on dividend and financing policies. *Journal of Management and Accounting Studies*, 3(02):13–20.
- Baber, W. R., Kang, S., and Kumar, K. R. (1999). The explanatory power of earnings levels vs. earnings changes in the context of executive compensation. *The Accounting Review*, 74(4):459–472.
- Baek, J.-S., Kang, J.-K., and Park, K. S. (2004). Corporate governance and firm value: Evidence from the korean financial crisis. *Journal of Financial economics*, 71(2):265–313.
- Bagwell, L. S. and Shoven, J. B. (1989). Cash distributions to shareholders. *Journal of Economic Perspectives*, 3(3):129–140.
- Baird, G., Dodd, J., and Middleton, L. (2020). A growth adjusted price-earnings ratio.
- Bajaj, M. and Vijh, A. M. (1990). Dividend clienteles and the information content of dividend changes. *Journal of Financial Economics*, 26(2):193–219.
- Baker, M. and Wurgler, J. (2004). A catering theory of dividends. *Journal of Finance*, 59:1125–1165.
- Balachandran, B., Krishnamurti, C., Theobald, M., and Vidanapathirana, B. (2012). Dividend reductions, the timing of dividend payments and information content. *Journal of Corporate Finance*, 18(5):1232–1247.
- Bank for International Settlements (2013). Treatment of sovereign risk in the basel capital framework. *BIS Quarterly Review*.
- Bank of England (2022a). Cp16/22 – implementation of the basel 3.1 standards: Credit risk – standardised approach, chapter 3. Technical report, Bank of England, Prudential Regulation.
- Bank of England (2022b). Dp3/22 – operational resilience: Critical third parties to the uk financial.
- Banker, R. D., Das, S., and Datar, S. M. (1993). Complementarity of prior accounting information: The case of stock dividend announcements. *Accounting Review*, pages 28–47.
- Banyi, M. L., Dyl, E. A., and Kahle, K. M. (2008). Errors in estimating share repurchases. *Journal of Corporate Finance*, 14(4):460–474.
- Bar-Yosef, S. and Venezia, I. (1991). Earnings information and the determination of dividend policy. *Journal of Economics and Business*, 43(3):197–214.

- Basel Committee for Banking Supervision (2009). Guidelines for computing capital for incremental risk in the trading book.
- Basel Committee for Banking Supervision (2019a). Calculation of rwa for market risk. credit valuation adjustment framework.
- Basel Committee for Banking Supervision (2019b). Minimum capital requirements for market risk.
- Basel Committee for Banking Supervision (2022a). Calculation of rwa for market risk. internal models approach: General provisions.
- Basel Committee for Banking Supervision (2022b). Calculation of rwa for market risk. internal models approach: Internal requirements.
- Basel Committee for Banking Supervision (2023). Otc derivatives statistics at end-june 2023.
- Basel Committee on Banking Supervision (2013). Regulatory consistency assessment programme (rcap) - analysis of risk-weighted assets for credit risk in the banking book.
- Basel Committee on Banking Supervision (2019). The basel framework. Technical report, Basel Committee on Banking Supervision.
- Bauer, L. and Bhattacharyya, N. (2006). Rethinking lintner: An alternative dynamic model of dividends. *Second Singapore International Conference on Finance 2008*.
- Benartzi, S., Michaely, R., and Thaler, R. (1997a). Do changes in dividends signal the future or the past? *The Journal of finance*, 52(3):1007–1034.
- Benartzi, S., Michaely, R., and Thaler, R. (1997b). Do changes in dividends signal the future or the past? *The Journal of Finance*, 52(3):1007–1034.
- Berle, A. A. and Means, G. C. (1932). The modern corporation and private property.
- Best, P. (2000). *Implementing Value-at-Risk*. John Wiley and Sons.
- Bhattacharyya, S. (1979). Imperfect information, dividend policy, and "the bird in the hand" fallacy. *The Bell Journal of Economics*, 10(1):259–270.
- Biagini, F., Gonon, L., Mazzon, A., and Meyer-Brandis, T. (2022). Detecting asset price bubbles using deep learning.
- Bianchi, D., Büchner, M., and Tamoni, A. (2019). What matters when? time-varying sparsity in expected returns. *Time-Varying Sparsity in Expected Returns (August 17, 2019)*. WBS Finance Group Research Paper.

- Biau, G. (2012). Analysis of a random forests model. *Journal of Machine Learning Research*, 13:1063–1095.
- Bierman Jr, H. and West, R. (1966). The acquisition of common stock by the corporate issuer. *Journal of Finance*, 21:687–696.
- Blanchard, O. J. and Milesi-Ferretti, G. M. (2010). Global imbalances: in midstream? *CEPR Discussion Paper*, (DP7693).
- Bluhm, C., Overbeck, L., and Wagner, C. (2010). *Introduction to Credit Risk Modelling (Second Edition)*, Chap. 2.4. Chapman and Hall.
- Boissel, C. and Matray, A. (2019). Higher dividend taxes, no problem! evidence from taxing entrepreneurs in france. *SSRN Electronic Journal*.
- Bouveret, A. and Haferkorn, M. (2023). Leverage and derivatives: The case of archegos. *Journal of Securities Operations and Custody*, 15(3):238–250.
- Bradley, M. and Wakeman, L. M. (1983). The wealth effects of targeted share repurchases. *Journal of Financial Economics*, 11(1-4):301–328.
- Brav, A., Harvey, C. R., Graham, J. R., and Michaely, R. (2005). Payout policy in the 21st century. *Tuck Contemporary Corporate Finance Issues III Conference Paper*.
- Breiman, L. (2013). Bagging predictors. *Machine Learning*, 24:123–140.
- Brennan, M. J. (1970). Taxes, market valuation and corporate financial policy. *National Tax Journal*, 23(4):417–427.
- Brennan, M. J. and Thakor, A. V. (1990). Shareholder preferences and dividend policy. *The Journal of Finance*, 45:993–1018.
- Brigham, E. F. (1964). The profitability of a firm’s purchase of its own common stock. *California Management Review*, 7:69 – 76.
- Brunton, S. L., Proctor, J. L., and Kutz, J. N. (2016). Discovering governing equations from data by sparse identification of nonlinear dynamical systems. *Proceedings of the National Academy of Sciences*, 113(15):3932–3937.
- Buehler, H., Gonon, L., Teichmann, J., and Wood, B. (2019a). Deep hedging. *Journal of Quantitative Finance*, 19:1271–1291.
- Buehler, H., Gonon, L., Teichmann, J., Wood, B., Mohan, B., and Kochems, J. (2019b). Deep hedging. *Swiss Finance Institute Research Paper*.

- Burgstahler, D. and Dichev, I. (1997). Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics*, 24(1):99–126. Properties of Accounting Earnings.
- Byrne, J. and O’Connor, T. (2017). Creditor rights, culture and dividend payout policy. *Journal of multinational financial management*, 39:60–77.
- C. Mitchell Conover, G. R. J. and Simpson, M. W. (2016). What difference do dividends make? *Financial Analysts Journal*, 72(6):28–40.
- Campello, M. and Graham, J. R. (2013). Do stock prices influence corporate decisions? evidence from the technology bubble. *Journal of Financial Economics*, 107(1):89–110.
- Cannata, F., Casellina, S., and Guidi, G. (2012). Inside the labyrinth of basel risk-weighted assets: How not to get lost. *Banca d’Italia Occasional Paper*, 132.
- Cantor, R. and Packer, F. (1996). Determinants and impact of sovereign credit ratings. *Economic Policy Review*, 2(2).
- Cantor, R. and Varma, P. (2004). Determinants of recovery rates on defaulted bonds and loans for north american corporate issuers: 1983-2003. *Journal of Fixed Income*.
- Carpenter, J. N. (1998). The exercise and valuation of executive stock options. *Journal of Financial Economics*, 48(2):127–158.
- Carroll, T. J. (1995). The information content of quarterly dividend changes. *Journal of Accounting, Auditing and Finance*, 10.
- Chang, B. and Dutta, S. (2012). Dividends and corporate governance: Canadian evidence. *IUP Journal of applied finance*, 18(4).
- Chatterjee, S. and Simonoff, J. S. (2012). *Handbook of Regression Analysis*. John Wiley and Sons.
- Cheikh, N., Hmiden, O., Zaied, Y., and al. (2021). Do sovereign credit ratings matter for corporate credit ratings? *Annals of Operations Research*, 297:77—114.
- Chen, A. and Obizhaeva, O. (2022). Stock buyback motivations and consequences: A literature review. In *CFA Institute Research Foundation Books*. S&P Global Market Intelligence.
- Chetty, R. and Saez, E. (2005). Dividend taxes and corporate behavior: Evidence from the 2003 dividend tax cut. *The Quarterly Journal of Economics*, 120(3):791–833.

- Chevyrev, I. and Kormilitzin, A. (2016). A primer on the signature method in machine learning.
- Chinn, M. D., Eichengreen, B., and Ito, H. (2014). A forensic analysis of global imbalances. *Oxford Economic Papers*, 66:465–490.
- Christensen, R. (2019). *Advanced Linear Modeling, Statistical Learning and Dependent Data*. Springer.
- Claessens, S., Djankov, S., and Lang, L. H. (2000). The separation of ownership and control in east asian corporations. *Journal of financial Economics*, 58(1-2):81–112.
- Coggan, P. (2009). *The Money Machine: How the City Works*. Penguin Books.
- Comment, R. and Jarrell, G. A. (1991). The relative signalling power of dutch-auction and fixed-price self-tender offers and open-market share repurchases. *The Journal of Finance*, 46:1243–1271.
- Consler, J., Lepak, G. M., and Havranek, S. F. (2011). Earnings per share versus cash flow per share as predictor of dividends per share. *Managerial Finance*, 37(5):482–488.
- Constantinides, G. M. and Grundy, B. D. (1989). Optimal investment with stock repurchase and financing as signals. *The Review of Financial Studies*, 2(4):445–465.
- Cronqvist, H., Heyman, F., Nilsson, M., Svaleryd, H., and Vlachos, J. (2009). Do entrenched managers pay their workers more? *the Journal of Finance*, 64(1):309–339.
- Cronqvist, H. and Nilsson, M. (2003). Agency costs of controlling minority shareholders. *Journal of Financial and Quantitative analysis*, 38(4):695–719.
- Crouhy, M., Galai, D., and Mark, R. (2000). A comparative analysis of current credit risk models. *Journal of Banking and Finance*, 24:59–117.
- Crum, W. L. (1935). Compensation of corporate executives – the 1928-1932 period. *Harvard Business Review*, pages 321–333.
- Cudd, M., Duggal, R., and Sarkar, S. (1996). Share repurchase motives and stock market reaction. *Quarterly Journal of Business and Economics*, 35(2):66–76.
- Cutler, A., Cutler, D. R., and Stevens, J. R. (2011). Random forests. pages 157–176.
- Dann, L. Y., Masulis, R. W., and Mayers, D. (1991). Repurchase tender offers and earnings information. *Journal of Accounting and Economics*, 14(3):217–251.

- De Moor, L., Luitel, P., Sercu, P., and Vanpee, R. (2018). Subjectivity in sovereign credit ratings. *Journal of Banking and Finance*, 88:366–392.
- DeAngelo, H. (2023). The attack on share buybacks. *European Financial Management*, 29(2):389–398.
- DeAngelo, H., DeAngelo, L., and Skinner, D. J. (1992). Dividends and losses. *The Journal of Finance*, 47(5):1837–1863.
- DeAngelo, H., DeAngelo, L., and Skinner, D. J. (2004). Are dividends disappearing? dividend concentration and the consolidation of earnings. *Journal of Financial Economics*, 72:425–456.
- DeAngelo, H. and Masulis, R. W. (1980). Leverage and dividend irrelevancy under corporate and personal taxation. *The Journal of Finance*, 35(2):453–464.
- Dedman, E., Kungwal, T., and Stark, A. W. (2014). The information content of regular dividends and share buybacks for market value and future earnings in the uk.
- Delianedis, G. and Geske, R. L. (2003). Credit risk and risk neutral default probabilities: Information about rating migrations and defaults. *S&P Global – Market Intelligence*.
- Dempsey, S. J., Laber, G., and Rozeff, M. S. (1993). Dividend policies in practice: Is there an industry effect? *Quarterly Journal of Business and Economics*, 32(4):3–13.
- Denis, D. J. (1990). Defensive changes in corporate payout policy: Share repurchases and special dividends.
- Denis, D. J., Denis, D. K., and Sarin, A. (1994). The information content of dividend changes: Cash flow signaling, overinvestment, and dividend clienteles. *Journal of Financial and Quantitative Analysis*, 29(4):567–587.
- Dillen, J. v. and Poitras, Geoffrey and Majithia, A. (2006). Isaac le maire and the early trading in dutch east india company shares. In *Pioneers of financial economics: volume 1*. Edward Elgar Publishing.
- Dionne, G. and Ouederni, K. (2011). Corporate risk management and dividend signaling theory. *Finance Research Letters*, 8(4):188–195.
- Dittmar, A., Mahrt-Smith, J., and Servaes, H. (2003). International corporate governance and corporate cash holdings. *Journal of Financial and Quantitative analysis*, 38(1):111–133.
- Dittmar, A. K. (2000). Why do firms repurchase stock. *The Journal of Business*, 73(3):331–355.

- Doyran, M. (2014). The argentine dilemma: “vulture funds” and the risks posed to developing economies. *Class, Race and Corporate Power*, 2(7).
- Duan, J.-C., Gauthier, G., and Simonato, J.-G. (2004). On the equivalence of the kmv and maximum likelihood methods for structural credit risk models.
- Duffie, D. and Pan, J. (1997). An overview of value-at-risk.
- Dyl, E. A. and Weigand, R. A. (1998). The information content of dividend initiations: Additional evidence. *Financial Management*, 27(3):27–35.
- Décamps, J. and Villeneuve, S. (2007). Optimal dividend policy and growth option. *Finance Stoch*, 11:3–27.
- Easterbrook, F. H. (1984). Two agency-cost explanations of dividends. *The American economic review*, 74(4):650–659.
- Ekmekcioglu, M., Kaya, T., and Tokmakcioglu, K. (2023). Predicting sovereign credit ratings using machine learning algorithms. In Calisir, F. and Durucu, M., editors, *Industrial Engineering in the Covid-19 Era*, pages 52–61. Springer Nature Switzerland.
- Elgers, P. T. and Murray, D. (1985). Financial characteristics related to managements’ stock split and stock dividend decisions. *Journal of Business Finance & Accounting*, 12(4):543–551.
- Elgouacem, A. and Zago, R. (2020). Share buybacks, monetary policy and the cost of debt. *Banque de France, Eurosystème: Working Paper Series no. 773*.
- Elton, E. and Gruber, M. (1970). Marginal stockholder tax rates and the clientele effect. *The Review of Economics and Statistics*, 52(1):68–74.
- Epperlein, E. and Herskovits, J. (2024). Infrequent mtm reduces neither value-at-risk nor backtesting exceptions. *Risk.net - Cutting Edge*.
- Erdem, O. and Varli, Y. (2014). Understanding the sovereign credit ratings of emerging markets. *Emerging Markets Review*, 20:42–57.
- European Banking Authority (2012). Final report on the implementation of capital plans following the eba’s 2011 recommendation on the creation of temporary capital buffers to restore market confidence. Technical report, EBA.
- European Banking Authority (2014). Results of 2014 eu-wide stress test. Technical report, EBA.

- European Central Bank Longaric, P. A., Grodzicki, M., Kaufmann, C., Pietsch, A., Serrano Ascandoni, P., Storz, M., and Telesca, E. (2023). Sovereign bond markets and financial stability: examining the risk to absorption capacity. *Financial Stability Review*.
- Fama, E. F. (1980). Agency problems and the theory of the firm. *Journal of political economy*, 88(2):288–307.
- Fama, E. F. and Babiak, H. (1968). Dividend policy: An empirical analysis. *Journal of the American statistical Association*, 63(324):1132–1161.
- Fama, E. F. and Jensen, M. C. (1983). Separation of ownership and control. *The journal of law and Economics*, 26(2):301–325.
- Farrar, D. E. and Selwyn, L. L. (1967). Taxes, corporate financial policy and return to investors. *National Tax Journal*, 20(4):444–454.
- Ferguson, N. and Prebble, S. (2008). *The Ascent of Money: A Financial History of the World*. Tantor Media; Penguin.
- Financial Stability Board (2022). Liquidity in core government bond markets.
- Franks, J. R. and Mayer, C. (2002). Governance as a source of managerial discipline. *National Bank of Belgium Working Paper*, (31).
- Gao, X. and Kronlund, M. (2020). Does equity-based compensation cause firms to manage earnings per share?
- Garcia, R. and Arroyo, J. M. (2012). Comparing risk-weighted assets: The importance of the supervisory validation processes.
- George, H., George, A. S., and Baskar, T. (2022). Sri lanka’s economic crisis: A brief overview. *Partners Universal International Research Journal*, 1.
- Ghadhab, I. (2023). Bonding, signaling theory and dividend policy: Evidence from multinational firms. *Journal of Asset Management*, 24:69–83.
- Ginglinger, E. and L’her, J.-F. (2006). Ownership structure and open market stock repurchases in france. *The European Journal of Finance*, 12(1):77–94.
- Glass, K. and Fritsche, U. (2015). Real-time macroeconomic data and uncertainty. DEP (Socioeconomics) Discussion Papers - Macroeconomics and Finance Series 6/2014R, Hamburg.
- Gogas, P. and Papdimitriou, T. (2021). *Computational Economics, Machine Learning in Economics and Finance*, volume 57.

- Goldstein, R., Ju, N., and Leland, H. (2001). An ebit-based model of dynamic capital structure. *The Journal of Business*, 74(4):483–512.
- Gomes, J. F. and Schmid, L. (2019). Equilibrium asset pricing with leverage and default. *Journal of Finance*, Forthcoming.
- Gonedes, N. J. (1978). Corporate signaling, external accounting, and capital market equilibrium: Evidence on dividends, income, and extraordinary items. *Journal of Accounting Research*, 16(1):26–79.
- Goodhart, C. A. E. (2011). *The Basel Committee on Banking Supervision : a history of the early years, 1974-1997*. Cambridge University Press.
- Graeme, G. (2020). The dark side of share buybacks.
- Grant, C. (2024). Buybacks are back: Corporate america is on a spending spree. Technical report, Wall Street Journal.
- Griffin, D. and De Paoli, L. (2024). Nomura, mizuho face losses on all blue fund’s failed trades. Technical report, Bloomberg Markets, UK.
- Grullon, G. and Ikenberry, D. L. (2000). What do we know about stock repurchases? *Journal of Applied Corporate Finance*, 13:31–51.
- Guay, W. and Harford, J. (2000). The cash-flow permanence and information content of dividend increases versus repurchases. *Journal of Financial economics*, 57(3):385–415.
- Gugler, K. and Yurtoglu, B. B. (2003). Corporate governance and dividend pay-out policy in germany. *European economic review*, 47(4):731–758.
- Guney, Y., Ozkan, A., and Ozkan, N. (2007). International evidence on the non-linear impact of leverage on corporate cash holdings. *Journal of Multinational financial management*, 17(1):45–60.
- Gupta, J. and Wagner, F. W. (2018). The announcement effect of open market share buybacks: The case of or european firms. *International Journal of Economics and Finance*, 10.
- Hakami, A. (2024). Strategies for overcoming data scarcity, imbalance, and feature selection challenges in machine learning models for predictive maintenance. *Scientific Reports*, 14(9065).
- Hammer, R. M. (1975). The taxation of income from corporate shareholders: Review of present systems in canada, france, germany, japan and the u.k. *National Tax Journal*, 28.

- Han, J., Arnulf, J., and Weinan, E. (2018). Solving high-dimensional partial differential equations using deep learning. *PNAS*.
- Harada, K. and Nguyen, P. (2011). Ownership concentration and dividend policy in japan. *Managerial Finance*, 37(4):362–379.
- Harris, T. and Ramsay, I. (1995). An empirical investigation of australian share buy-backs. *Australian Journal of Corporate Law*, 4.
- Hauf, P. and Posth, J.-A. (2023). Silicon valley bank - (why) did regulation and risk management fail to uncover substantial risks? *SSRN*.
- HBR (1926). Stock dividends – capital or income. *Harvard Business Review*, pages 102–115.
- HBR (1927a). Conflicting interests in a company’s purchase of its own securities. *Harvard Business Review*, pages 481–487.
- HBR (1927b). The dividend limit. *Harvard Business Review*, pages 501–507.
- Henault, B. (2023). France’s macron wants companies to share more profits with workers. *Reuters*.
- Hill, P., Brooks, R., and Faff, R. (2010). Variations in sovereign credit quality assessments across rating agencies. *Journal of Banking and Finance*, 34(6):1327–1343.
- Holt, W. A. and Morris, E. L. (1934). Some aspects of reacquired stock. *Harvard Business Review*, pages 505–510.
- Houston, J. F., Lin, C., and Xie, W. (2018). Shareholder protection and the cost of capital. *The Journal of Law and Economics*, 61(4):677–710.
- Hovakimian, A., Opler, T., and Titman, S. (2002). The capital structure choice: New evidence for a dynamic tradeoff model. *Journal of Applied Corporate Finance*, 15:24–30.
- Howe, K. M., He, J., and Kao, G. W. (1992). One-time cash flow announcements and free cash-flow theory: Share repurchases and special dividends. *The Journal of Finance*, 47(5):1963–1975.
- Hribar, P., Jenkins, N. T., and Johnson, W. B. (2006). Stock repurchases as an earnings management device. *Journal of Accounting and Economics*, 41(1):3–27.
- Hull, P., Kolesar, M., and Walters, C. (2022). Labor by design: contributions of david card, joshua angrist, and guido imbens. *The Scandinavian Journal of Economics*, 124:603–645.

- Hunting, M. and Paulsen, J. (2013). Optimal dividend policies with transaction costs for a class of jump-diffusion processes. *Finance Stoch*, 17:73–106.
- Hwang, L.-S., Kim, H., Park, K., and Park, R. S. (2013). Corporate governance and payout policy: Evidence from korean business groups. *Pacific-Basin Finance Journal*, 24:179–198.
- Ikenberry, D., Lakonishok, J., and Vermaelen, T. (1995). Market underreaction to open market share repurchases. *Journal of financial economics*, 39(2-3):181–208.
- Internal Revenue Service (2022a). Reg-115710-22 - excise tax on repurchase of corporate stock. Technical report, Department of the Treasury, U.S. Federal Government.
- Internal Revenue Service (2022b). Reg-118499-23 - excise tax on repurchase of corporate stock – procedure and administration. Technical report, Department of the Treasury, U.S. Federal Government.
- ISDA and for Financial Markets in Europe, A. (2024). Isda response to eba on internal models benchmarking.
- Israel, R., Kelly, B. T., and Moskowitz, T. J. (2020). Can machines ‘learn’ finance? *Journal of Investment Management*.
- Jackson Jr., R. J. (2018). Stock buybacks and corporate cashouts. Technical report, U.S. Securities and Exchange Commission.
- Jagannathan, M., Stephens, C. P., and Weisbach, M. S. (2000). Financial flexibility and the choice between dividends and stock repurchases. *Journal of financial Economics*, 57(3):355–384.
- Jarrow, R. A., David, L., and Yu, F. (2005). Default risk and diversification: Theory and empirical implications. *Mathematical Finance*, 15.
- Jensen, M. C. and Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics*, 3(4):305–360.
- John, K. and Williams, J. (1985). Dividends, dilution, and taxes: A signalling equilibrium. *The Journal of Finance*, 40(4):1053–1070.
- Johnson, S., La Porta, R., Lopez-de Silanes, F., and Shleifer, A. (2000). Tunneling. *American economic review*, 90(2):22–27.
- Jongh, M. d. (2011). Shareholder activism at the dutch east india company 1622-1625. In *Origins of Shareholder Advocacy*. Palgrave Macmillan.

- Julio, B. and Ikenberry, D. L. (2004). Reappearing dividends. *Journal of applied corporate finance*, 16(4):89–100.
- Kamenis, S. D. (2014). Vulture funds and the sovereign debt market: Lessons from argentina and greece. *Crisis Observatory Research Paper*, (13).
- Kandpal, V. and Kavidayal, P. C. (2015). A study of dividend policy and its effect on market value of shares of selected banks in india. *IOSR Journal of Business and Management*, 17(1):41–44.
- Keskar, N. S., Mudigere, D., Nocedal, J., Smelyanskiy, M., and Tang, P. T. P. (2016). On large-batch training for deep learning: Generalization gap and sharp minima. *CoRR*, abs/1609.04836.
- Kidger, P., Bonnier, P., Perez Arribas, I., Salvi, C., and Lyons, T. (2019). Deep signature transforms. In Wallach, H., Larochelle, H., Beygelzimer, A., d’Alch’e Buc, F., Fox, E., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, volume 32. Curran Associates, Inc.
- Kim, J., Schremper, R., and Varaiya, N. P. (2024). Open market repurchase regulations: A cross-country examination. *Corporate Finance Review*, 9:29–38.
- Klug, A. (1993). The german buybacks, 1932-1939: A cure for overhang? *Princeton Studies in International Finance*, No75.
- Koch-Weser, I. N. (2013). The reliability of china’s economic data: An analysis of national output. Technical report, U.S.-China Economic and Security Review Commission (USCC).
- Koot, G. (1995). The voc, the dutch east india company, 1602-1799. *History Department, University of Massachusetts Dartmouth*.
- KPMG (2023). Pra basel 3.1 policy statement - ps 17/23. near-final rules for market, cva, counterparty credit and operational risk.
- KPMG (2024). Kpmg report: Initial analysis of stock repurchase excise tax proposed regulations. Technical report, KPMG.
- Kumar, P. (1988). Shareholder-manager conflict and the information content of dividends. *The Review of Financial Studies*, 1(2):111–136.
- Lambert, R. A., Lanen, W. N., and Larcker, D. F. (1989). Executive stock option plans and corporate dividend policy. *Journal of financial and quantitative analysis*, 24(4):409–425.

- Lambrecht, B. M. and Myers, S. C. (2017). The dynamics of investment, payout and debt. *The Review of Financial Studies*, 30(11):3759–3800.
- Lanotte, M., Manzelli, G., Rinaldi, A. M., Taboga, M., and Tommasino, P. (2016). Easier said than done? reforming the prudential treatment of banks’ sovereign exposures. *Banca d’Italia Occasional Paper*, (326).
- Lazonick, W. (2014). Profits without prosperity. *Harvard Business Review*.
- Lazonick, W. (2015). Buybacks: From basics to politics. Technical report.
- Le Leslé, V. and Avramova, S. Y. (2012). Revisiting risk-weighted assets. *International Monetary Fund (IMF) Working Paper*, 12/90.
- Lee, Y.-G., Jung, S.-C., and Thornton, J. H. (2005). Long-term stock performance after open-market repurchases in korea. *Global Finance Journal*, 16(2):191–209.
- Lefebvre, M.-N., Lehmann, E., Sicsic, M., and Zanoutene, E. (2021). Faut-il mettre au bareme les dividendes ?. *Revue francaise d’economie*, 26:57–98.
- Lei, Z. and Zhang, C. (2016). Leveraged buybacks. *Journal of Corporate Finance*, 39:242–262.
- Leland, H. E. (1994). Corporate debt value, bond covenants, and optimal capital structure. *The Journal of Finance*, 49(4):1213–1252.
- Leland, H. E. and Toft, K. B. (1996). Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *The Journal of Finance*, 51(3):987–1019.
- Lins, K. V. (2003). Equity ownership and firm value in emerging markets. *Journal of Financial and Quantitative Analysis*, 38(1):159–184.
- Lintner, J. (1956). Distribution of incomes of corporations among dividends, retained earnings and taxes.
- Lintner, J. (1962). Dividends, earnings, leverage, stock prices and the supply of capital to corporations. *The Review of Economics and Statistics*, 44(3):243–269.
- Lintner, J. (1964). Optimal dividends and corporate growth under uncertainty. *The Quarterly Journal of Economics*, 78(1):49–95.
- Lipson, M. L., Maquieira, C. P., and Megginson, W. (1998). Dividend initiations and earnings surprises. *Financial Management*, 27(3):36–45.

- Litzenberger, R. H. and Ramaswamy, K. (1982). The effects of dividends on common stock prices tax effects or information effects? *The Journal of Finance*, 37(2):429–443.
- Lv, J., Ye, Y., and Luo, R. (2023). The power of minority shareholders: evidence from voting on the related party transaction proposals in china. *Review of Accounting and Finance*, 22(2):157–176.
- Lyons, T. (2014). Rough paths, signatures and the modelling of functions on streams.
- Lyons, T. and McLeod, A. D. (2022). Signature methods in machine learning.
- Lyons, T. and Xu, W. (2014). Inverting the signature of a path. *Journal of the European Mathematical Society*, 20.
- Løkka, A. and Zervos, M. (2008). Optimal dividend and issuance of equity policies in the presence of proportional costs. *Insurance: Mathematics and Economics*, 42(3):954–961.
- Machuga, S. M., Pfeiffer, R. J., and Verma, K. (2002). Economic value added, future accounting earnings, and financial analysts’ earnings per share forecasts. *Review of Quantitative Finance and Accounting*, 18:59–73.
- Masulis, R. W. (1980). Stock repurchase by tender offer: An analysis of the causes of common stock price changes. *Journal of Finance*, 35:305–319.
- Masulis, R. W. and Trueman, B. (1988). Corporate investment and dividend decisions under differential personal taxation. *The Journal of Financial and Quantitative Analysis*, 23(4):369–385.
- May, G. O. (1922). The taxation of capital gains. *Harvard Business Review*, pages 11–18.
- McNichols, M. and Dravid, A. (1990). Stock dividends, stock splits, and signaling. *the Journal of finance*, 45(3):857–879.
- Mellouli, Y. and Khiari, W. (2015). The development of the proper-stock repurchases in france: Role of the stock-option plan. *International Review of Management and Business Research*, 4(2):402.
- Metz, A. and Cantor, R. (2006). Moody’s credit rating prediction model. Technical report, Moody’s Investors Service.
- Miller, M. H. and Rock, K. (1985). Dividend policy under asymmetric information. *The Journal of Finance*, 40(4):1031–1051.

- Miller, M. H. and Scholes, M. S. (1982). Dividends and taxes: Some empirical evidence. *Journal of Political Economy*, 90(6):1118–1141.
- Ministry of Finance, J. G. (2023a). Japanese government bonds, january 2023 newsletter.
- Ministry of Finance, J. G. (2023b). List of auctions participants (as of december 27, 2023).
- Ministry of Finance, Japanese Government (2023). Announcement of liquidity enhancement auction on december 20,2023.
- Mishra, A. (2005). An empirical analysis of share buybacks in india. *ICFAI Journal of Applied Finance*, 11:5–24.
- Mizruchi, M. S. (1983a). Relations among large american corporations, 1904-1974. *Social Science History*, 7(2):165–182.
- Mizruchi, M. S. (1983b). Who controls whom? an examination of the relation between management and boards of directors in large american corporations. *Academy of management Review*, 8(3):426–435.
- Modigliani, F. and Miller, M. H. (1958). The cost of capital, corporation finance and the theory of investment. *The American Economic Review*, 48(3):261–297.
- Modigliani, F. and Miller, M. H. (1961). Dividend policy, growth, and the valuation of shares. *The Journal of Business*, 34(4):411–433.
- Modigliani, F. and Miller, M. H. (1963). Corporate income taxes and the cost of capital: A correction. *The American Economic Review*, 53(3):433–443.
- Moody’s Investors Service (2016). Moody’s rating symbols and definitions. Technical report.
- Myers, S. C. and Majluf, N. S. (1984). Corporate financing and investment decisions when firms have information that investors do not have. *Journal of financial economics*, 13(2):187–221.
- Nellis, S., Cherney, M. A., and Malik, Y. (2024). Apple unveils record \$110 billion buyback as results beat low expectations. Technical report, Reuters.
- Nishikawa, T., Prevost, A. K., and Rao, R. P. (2011). Bond market reaction to stock repurchases: Is there a wealth transfer effect? *Journal of Financial Research*, 34(3):503–522.

- Nissim, D. and Ziv, A. (2002). Dividend changes and future profitability. *The Journal of Finance*, 56:2111–2133.
- Obstfeld, M. and Rogoff, K. S. (2005). Global current account imbalances and exchange rate adjustments. *Brookings papers on economic activity*, 36:67–146.
- Ofer, A. R. and Thakor, A. V. (1987). A theory of stock price responses to alternative corporate cash disbursement methods: Stock repurchases and dividends. *The Journal of Finance*, 42(2):365–394.
- Oral, M., Kettani, O., Cosset, J.-C., and Daouas, M. (1992). An estimation model for country risk rating. *International Journal of Forecasting*, 8(4):583–593.
- Overes, B. H. and van der Wel, M. (2023). Modelling sovereign credit ratings: Evaluating the accuracy and driving factors using machine learning techniques. *Computational Economics*, 61:1273–1303.
- Ozbayoglu, A. M., Gudelek, M. U., and Sezer, O. B. (2020). Deep learning for financial applications : A survey. *Applied Soft Computing*, 93:106384.
- Ozturk, H., Namli, E., and Erdal, H. I. (2016). Modelling sovereign credit ratings: The accuracy of models in a heterogeneous sample. *Economic Modelling*, 54:469–478.
- Partnoy, F., Levich, R., Majnoni, G., and Reinhart, C. M. (2002). *The paradox of credit ratings: Ratings, rating agencies and the global financial system*. Kluwer Academic Press.
- Petram, L. (2014). The world’s first stock exchange. Columbia University Press.
- Pettit, R. R. (1972). Dividend announcements, security performance, and capital market efficiency. *The Journal of finance*, 27(5):993–1007.
- Pinkowitz, L., Stulz, R., and Williamson, R. (2006). Does the contribution of corporate cash holdings and dividends to firm value depend on governance? a cross-country analysis. *The Journal of Finance*, 61:2725–2751.
- Pinter, G. (2023). An anatomy of the 2022 gilt market crisis. *Bank of England Working Paper*, (1019).
- Poor’s, S. . (2011). Sovereign government rating methodology and assumptions. Technical report.
- Proctor, L. (1997). The barings collapse: A regulatory failure, or a failure of supervision. *Brooklyn Journal of International Law*, 22(3).

- Prudential Regulation Authority (2023). Ps17/23 – implementation of the basel 3.1 standards near-final part 1. near-final part 1 policy statement 17/23.
- PwC (2016). Basel iv: Revised internal models approach for market risk. new challenges for banks.
- PwC (2020). Risks not in var (rniv) and their role in market risk management.
- PwC (2023). Worldwide tax summaries (2017 - 2023), capital gains tax (cgt) rates.
- Ratings, F. (2022). Rating action commentary fitch affirms luxembourg at 'aaa'; outlook stable. Technical report, Fitch Ratings.
- Ratings, S. . P. G. (2015). Sovereign ratings history since 1975. Technical report.
- Reusens, P. and Croux, C. (2016). Sovereign credit rating determinants: The impact of the european debt crisis. Technical Report 1425, Faculty of Economics and Business, KU Leuven.
- Reuters (2011). The day the u.s. (technically) defaulted. Technical report.
- Ross, S. A. (1977). The determination of financial structure: the incentive-signalling approach. *The bell journal of economics*, pages 23–40.
- Rowland, P. (2004). Determinants of spread, credit ratings and creditworthiness for emerging market sovereign debt: A follow-up study using pooled data analysis. Technical report, Banco de la Republica de Colombia.
- Rozeff, M. S. (1982). Growth, beta and agency costs as determinants of dividend payout ratios. *The Journal of Financial Research*, 5:249–259.
- Ruf, J. and Wang, W. (2020). Neural networks for option pricing and hedging: a literature review. *The Journal of Computational Finance*.
- Ruf, J. and Wang, W. (2022). Hedging with linear regressions and neural networks. *Journal of Business & Economic Statistics*, 40(4):1442–1454.
- Russell, S. and Norvig, P. (2010). *Artificial intelligence: A modern approach*. Prentice-Hall, 3 edition.
- Salvi, C., Cass, T., Foster, J., Lyons, T., and Yang, W. (2021). The signature kernel is the solution of a goursat pde. *SIAM Journal on Mathematics of Data Science*, 3(3):873–899.
- Schaede, U. (1989). Forwards and futures in tokugawa-period japan: A new perspective on the dojima rice market. *Journal of Banking and Finance*, 13:487–513.

- Shaikh, A. (2016). *Capitalism: Competition, conflict, crises*. Oxford University Press.
- Shalev-Shwartz, S. and Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. Cambridge University Press.
- Sharma, D., Anawade, P., Sahu, A., and Sharma, M. (2022). The economic crisis faced by island nation-sri lanka: An empirical study. *Journal of Contemporary Issues in Business and Government*, 28:920–927.
- Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *The Journal of Finance*, 19:425–442.
- Shefrin, H. M. and Statman, M. (1984). Explaining investor preference for cash dividends. *Journal of Financial Economics*, 13(2):253–282.
- Shleifer, A. and Vishny, R. W. (1997). A survey of corporate governance. *The journal of finance*, 52(2):737–783.
- Shleifer, A., Vishny, R. W., Porta, R., and Lopez-de Silanes, F. (2000). Investor protection and corporate governance. *Journal of financial economics*, 58(1-2):3–27.
- Shoven, J. B. and Simon, L. B. (1987). Share repurchases and acquisitions: An analysis of which firms participate. *National Bureau of Economic Research*.
- Silva, D., do Rego, T. G., and Frascaroli, B. (2019). Sovereign risk ratings’ country classification using machine learning.
- Sinha, S. (1991). Share repurchase as a takeover defense. *Journal of Financial and Quantitative Analysis*, 26(2):233–244.
- Smith, A. (1776). *The wealth of nations*. Cannan Edition.
- Sohnke, M. B., Conrad, J., Lee, J., and Subrahmanyam, M. G. (2022). Credit default swaps around the world. *The Review of Financial Studies*, 35(5):2464–2524.
- Sorkin, A. R., Karaian, J., Gandel, S., de la Merced, M. J., Hirsch, L., and Livni, E. (2022). Biden renews pushback against stock buybacks. *The New York Times*.
- Stephens, C. P. and Weisbach, M. S. (1998). Actual share reacquisitions in open-market repurchase programs. *The Journal of finance*, 53(1):313–333.
- Stulz, R. M. (2010). Credit default swaps and the credit crisis. *The Journal of Economic Perspectives*, 24(1):73–92.
- Sy, A. N. (2009). The systemic regulation of credit rating agencies and rated markets. *World Economics Data Papers*, 10(4):69–108.

- Tabassum, M. Y. (2024). Why do banks use credit default swaps (cds)? a systematic review. *Journal of Economic Surveys*, 38(1):201–231.
- Teichman, J., Ortega, J.-P., Grigoryeva, L., Gonon, L., and Cuchiero, C. (2021). Expressive power of randomized signature.
- Thanatawee, Y. (2013a). Actual share repurchases and liquidity: Evidence from thailand. *International Research Journal of Finance and Economics*, (104):42–50.
- Thanatawee, Y. (2013b). Ownership structure and dividend policy: Evidence from thailand. *International Journal of Economics and Finance*, 5(1):121–132.
- Thirumagal, P. G. and Vasantha, S. (2018). A research paper on impact of dividend payout on shareholders wealth in indian industries.
- Thirumalvalavan, P. and Sunitha, K. (2006). Share price behaviour around buy back and dividend announcements in india. *Indian Institute of Capital Markets 9th Capital Markets Conference Paper*.
- Tse, A. S. L. (2018). Dividend policy and capital structure of a defaultable firm.
- UNGEGN (2011). Ungegn list of country names. Technical report, United Nations Group of Experts on Geographical Names (UNGEGN).
- Utzig, S. (2010). The financial crisis and the regulation of credit rating agencies: A european banking perspective. Technical Report 188, Asian Development Bank Institute.
- Vafeas, N. and Joy, O. M. (1995). Open market share repurchases and the free cash flow hypothesis g35. *Economics letters*, 48(3-4):405–410.
- Vermaelen, T. (1981). Common stock repurchases and market signalling: An empirical study. *Journal of financial economics*, 9(2):139–183.
- Vermaelen, T. (1984). Repurchase tender offers, signaling, and managerial incentives. *Journal of Financial and Quantitative Analysis*, 19(2):163–181.
- Vij, M. (2005). The determinants of country risk analysis. *Journal of Management Research*, 5(1):20–31.
- Vithessonthi, C. (2008). Stock price performance of initial public offerings: The thai experience. *Journal of International Finance and Economics*, 8(2):30–43.
- Vogt, S. C. (1994). The cash flow/investment relationship: evidence from us manufacturing firms. *Financial management*, pages 3–20.

- Wada, K. (2005). Stock repurchases in japan. *Working Paper, Keio University, Japan*.
- Wang, H. B., Nguyen, C., and Rafi, N. A. (2021). The effectiveness of price-stabilizing share buybacks: Evidence from listed firms in vietnam. *The North American Journal of Economics and Finance*, 57:101436.
- Wang, W. and Ruf, J. (2022). A note on spurious model selection. *Quantitative Finance*, 22(10):1797–1800.
- Wansley, J. W., Lane, W. R., and Sarkar, S. (1989). Managements’ view on share repurchase and tender offer premiums. *Financial Management*, 18:97–110.
- Watts, R. (1973). The information content of dividends. *The Journal of Business*, 46:191–211.
- Weir, D. R. (1986). The reliability of historical macroeconomic data for comparing cyclical stability. *The Journal of Economic History*, 46:353–365.
- Weisbenner, S. J. (2000). Corporate share repurchases in the 1990s: What role do stock options play? *Finance and Economics Discussion Series (FEDS)*.
- Werner, R. (2015). *Princes of the Yen: Japan’s Central Bankers and the Transformation of the Economy*. Routledge.
- West, M. D. (2000). Private ordering at the world’s first futures exchange. *Michigan Law Review*, 98(8).
- Weston, J. F. and Siu, J. A. (2003). Changing motives for share repurchases. *The Finance*.
- Wilde, T. (1997). Creditrisk+ a credit risk management framework. *Credit Suisse Financial Products*.
- Williamson, O. E. (1984). The economics of governance: Framework and implications. *Zeitschrift für die gesamte Staatswissenschaft / Journal of Institutional and Theoretical Economics*, 140(1):195–223.
- Yim, J. and Mitchell, H. (2005). Comparison of country risk models: Hybrid neural networks, logit models, discriminant analysis and cluster techniques. *Expert Systems with Applications*, 28(1):137–148.
- Yoon, P. S. and Starks, L. T. (2015). Signaling, Investment Opportunities, and Dividend Announcements. *The Review of Financial Studies*, 8(4):995–1018.

- Yun, D. and Hai-ming, D. (2011). Mechanism model of the influence on board structure from the control of single blockholder. In *2011 International Conference on Management Science & Engineering 18th Annual Conference Proceedings*, pages 67–72. IEEE.
- Zingales, L. (1994). The value of the voting right: A study of the milan stock exchange experience. *The Review of Financial Studies*, 7(1):125–148.
- Zywicki, T. J. and Skeel, D. A. (2003). The past, present, and future of bankruptcy law in america. *Michigan Law Review*, 101(6):2016–2036.

