

Forecasting Consumer Price Inflation with Generalised Network Autoregressive Processes

Guy P. Nason^a, Henry Antonio Palasciano^{a,*}

^a*Department of Mathematics, Imperial College London, 180 Queen's Gate, London, SW7 2AZ, United Kingdom*

Abstract

This article forecasts CPI inflation in the United Kingdom using modified Generalised Network Autoregressive (GNAR) Processes, a class of autoregressive models that operate on time series associated with graphs. We fit GNAR models to a large set of random networks generated according to the Erdős–Rényi–Gilbert model and select the best-performing networks each month to compute out-of-sample forecasts. Our models significantly outperform traditional benchmarks across all horizons, especially at medium- to long-term forecast horizons, which is of great importance to policymakers charged with interest rate setting. They are also able to deliver more accurate predictions than the Bank of England's four to six month inflation rate forecasts published in their quarterly Monetary Policy Reports. In line with existing literature on inflation forecasting using autoregressive methods, we find that low-order GNAR models tend to be the most effective. By analysing the structure of the best-performing graphs, we are also able to identify the key CPI components that influence inflation rates during different periods for the best forecasters.

Keywords: inflation, forecasting, statistical learning, graphs, autoregressive processes, CPI disaggregated series.

1. Introduction

Forecasting inflation is crucial for economic policymakers, such as the Bank of England, and plays a significant role in the broader economy by aiding businesses and firms in their strategic planning. Inflation forecasts are fundamental to shaping monetary policy, with the primary objectives of maintaining price stability and fostering economic growth, while also serving as tools to communicate the rationale behind policy decisions to financial institutions and the general public (Bernanke and Woodford, 1997; Batini and Haldane, 1999; Binder and Sekkel, 2024; Bernanke, 2024). This has been highlighted by the UK's (and the global) recent struggle with high inflation, which presented considerable challenges for policymakers.

*Corresponding author

Email addresses: `g.nason@imperial.ac.uk` (Guy P. Nason), `henry.palasciano17@imperial.ac.uk` (Henry Antonio Palasciano)

Following a sustained period of elevated inflation, the rate has only now begun to approach the Bank of England’s target, highlighting the need for robust and reliable forecasting models to enable better *earlier* decision-making. This need is further emphasised by the fact that inflation forecasting accuracy has deteriorated in recent years, both among central banks and external forecasters (Binder and Sekkel, 2024; Bernanke, 2024).

For many years, univariate time series forecasting has been a fundamental approach for predicting inflation. Faust and Wright (2013) provide a comprehensive review of various forecasting techniques, including autoregressive models, direct forecasting methods, random walks, and the Phillips curve. The authors find that the most effective approaches tend to be simple methods requiring few or no parameters. The effectiveness of the Phillips curve, a method which essentially extends an autoregressive model by incorporating the unemployment rate as an additional predictor, has been the subject of much debate (Stock and Watson, 1999, 2009; Atkeson and Ohanian, 2001; Faust and Wright, 2013). For example, Atkeson and Ohanian (2001) show that the Phillips curve struggles to outperform even simple naive forecasts and advocate for the use of a random walk of order four instead. On the other hand, Stock and Watson (2009) support the Phillips curve over more complex multivariate forecasting techniques, although its performance relative to other univariate methods is inconsistent and highly dependent on prevailing macroeconomic conditions. Furthermore, Stock and Watson (2007) argue that univariate models with time varying parameters are the best models for forecasting inflation. In general, much of the literature seems to suggest that simple univariate time series are remarkably difficult to surpass when it comes to forecasting inflation (Atkeson and Ohanian, 2001; Stock and Watson, 2003, 2007; Faust and Wright, 2013).

There have been several attempts to improve upon linear univariate time series inflation rate forecasts using more advanced techniques. For instance, Joseph et al. (2024) explore a range of machine learning and shrinkage methods to forecast UK inflation, using the individual item series as predictors. They find that only shrinkage methods are able to outperform the benchmarks, and even then, only during specific sub-periods. However, Barkan et al. (2023); Boaretto and Medeiros (2023); Medeiros et al. (2021) have outperformed classical benchmarks with machine learning methods, including recurrent neural networks and long short-term memory models, when applied to the disaggregate item series in the US and Brazil. In contrast, Almosova and Andresen (2023) find that such non-linear forecasting techniques do not yield significant improvements in performance over autoregressive models when applied to the US inflation rate time series alone. Makridakis et al. (2018) argue that standard statistical techniques are more accurate than machine learning methods when applied to a large range of economic and business data series, although such a conclusion was reached on average, and in certain areas, machine learning methods might still be better suited. Alternative data sources have also been explored for the purpose of forecasting inflation. For example, Aparicio and Bertolotto (2020) demonstrated success using online prices to forecast inflation rates across multiple countries, exploiting the fact that such information is available at a higher frequency. More recently, Eugster and Uhl (2024) applied sentiment analysis to a large corpus of news articles, producing US inflation forecasts that outperform a naive random walk model. Hence, although improvements in performance can

be obtained, univariate time series models remain highly robust inflation rate forecasters, especially when one takes into account the simplicity of such methods.

Generally speaking, the best kinds of forecasts are those that involve the use of surveys or a degree of expert judgement (Ang et al., 2007; Faust and Wright, 2013; Binder and Sekkel, 2024), such as the Bloomberg, U.S. and European Central Bank Surveys of Professional Forecasters and the Bank of England Monetary Policy Reports and Survey of External Forecasters (Boero et al., 2008; Binder and Sekkel, 2024; Clements et al., 2023; Clements, 2024). These methods integrate expert insights and blend forecasts from a wide array of models. For example, the Bank of England’s forecasting model typically makes use of a suite of at least forty models, including its central model, COMPASS, when generating a forecast (Burgess et al., 2013; Bernanke, 2024). Further, a single forecast involves several hundred runs of the COMPASS algorithm and at least five Monetary Policy Committee meetings, where judgemental adjustments are made to both the forecasts and the models used. Given the complexity and thoroughness of such processes, it is extremely challenging for other methods to compete. For example, Domit et al. (2019) find that a Bayesian VAR is unable to outperform the Bank of England’s inflation rate forecasts. This is why simple univariate time series models are often used as benchmarks — not only are they straightforward, but also among the most effective when relying solely on historical data.

Graphs are popular tools that have been extensively used in many fields, such as Bayesian statistics (Heckerman, 2008; Koski and Noble, 2009), time series analysis (Knight et al., 2016; Zhu et al., 2017; Knight et al., 2020; Kang et al., 2021; Lucchese et al., 2023), feature extraction and classification (Peach et al., 2019, 2021), deep learning (Scarselli et al., 2008; Wu et al., 2019; Zhou et al., 2020a), causal inference (Pearl, 2009a,b; Yao et al., 2021), and neuroscience (Bullmore and Sporns, 2009; Zhou et al., 2020b), to name but a few. Excellent and comprehensive introductions to statistical network analyses can be found in Kolaczyk (2009) and Kolaczyk and Csárdi (2020).

This article uses graphs constructed on the Consumer Price Index (CPI) component time series (treated as vertices) to forecast CPI inflation in the United Kingdom. More specifically, given a configuration of edges, we fit the Generalised Network Autoregressive (GNAR) Processes introduced by Knight et al. (2020) to compute forecasts for horizons up to twelve months ahead. GNAR models are an extension of autoregressive models which use information from neighbouring nodes (vertices that represent individual time series) as additional predictors. We generate a set of random graphs according to the Erdős–Rényi–Gilbert model (Gilbert, 1959) and then select the ones which perform best at each time step, which are then used to obtain our inflation rate forecasts. We find that simple GNAR models of order one or two are able to consistently outperform the classic univariate benchmarks. A welcome and surprising finding is that our forecasts are also able to beat the Bank of England’s four- to six-month ahead forecasts, further underlining the strength of our approach. Moreover, by analysing the structure of the best performing networks, we are able to identify the set of specific components which most influence the inflation rate forecast each month.

Significant research has been conducted on studying and forecasting inflation using disaggregated item series. For instance, Stock and Watson (2020) examine the components of inflation in the US with a focus on cyclical variation, while Hubrich (2005); Hendry and

Hubrich (2011); Chu et al. (2018) explore the effectiveness of using aggregate data for forecasting in the US, the Euro area, and the UK, respectively. As previously mentioned, studies by Medeiros et al. (2021); Barkan et al. (2023); Boaretto and Medeiros (2023); Joseph et al. (2024) have also attempted to forecast inflation using various shrinkage and machine learning techniques, yielding mixed results. The literature generally indicates that individual component series offer valuable insights for forecasting inflation, which networks and GNAR processes are particularly well-suited to exploit.

This article is organised as follows. Section 2 introduces GNAR processes. Section 3 provides a brief review of CPI in the UK and describes the data used in this work. Section 4 describes our approach in more detail. Section 5 compares our approach to various benchmark models, as well as to the Bank of England forecasts. Finally, Section 6 analyses the structure of the most successful networks in order to determine the most influential components.

2. Generalised Network Autoregressive Processes

Generalised Network Autoregressive (GNAR) Processes (Knight et al., 2016; Leeming, 2019; Knight et al., 2020) describe the behaviour of multivariate time series linked according to an associated graph. They have found successful application in economics (Nason and Wei, 2022; Mohammad and Khochiani, 2022), medicine (Nason et al., 2023), finance (Lorenzo-Valdés, 2024) and computational politics (Nason et al., 2024), to name a few. For many multivariate time series it has been shown, by several of the previous papers mentioned, for example, that GNAR models provide very parsimonious models, that are extremely rapid to compute and store and exhibit excellent forecasting properties. We now provide a brief review of these processes, borrowing notation from Knight et al. (2020).

A graph $\mathcal{G} = (\mathcal{K}, \mathcal{E})$ consists of a set of nodes $\mathcal{K}$ and a set of edges $\mathcal{E}$. Edges between nodes i and j can be either directed, denoted by $i \rightsquigarrow j$, or undirected, denoted by $i \leftrightarrow j$. For brevity, we define everything for undirected graphs; the definitions for directed graphs are analogous by replacing $\leftrightarrow$ with $\rightsquigarrow$.

The edge set of a network is defined as $\mathcal{E} = \{(i, j) : i \leftrightarrow j, i, j \in \mathcal{K}\}$, while the neighbour set of a set of nodes $A \subset \mathcal{K}$ is defined as $\mathcal{N}(A) = \{j \in \mathcal{K} : i \leftrightarrow j \text{ for some } i \in A\}$. The stage r neighbour set can then be defined as ‘neighbours of the $(r-1)$ stage neighbours, not including any previous neighbour already seen’, which, written mathematically is

$$\mathcal{N}^{(r)}(A) = \mathcal{N}\{\mathcal{N}^{(r-1)}(A)\} / [\{\cup_{q=1}^{r-1} \mathcal{N}^{(q)}(A)\} \cup A], \quad (1)$$

representing the set of nodes whose shortest path to A is of length r edges. For simplicity, we write $\mathcal{N}_i^{(r)} = \mathcal{N}^{(r)}(\{i\})$ when referring to the stage- r neighbour set of a single node. Finally, the connection weight between nodes i and j , given that $j \in \mathcal{N}_i^{(r)}$, is defined as

$$\omega_{i,j} = |\mathcal{N}_i^{(r)}|^{-1}, \quad (2)$$

where $|A|$ denotes the cardinality of the set A . If the network edges are weighted, the stage- r connection weights are determined by the combined weights of the shortest path between

node i and its stage- r neighbour, normalised so that the sum of all stage- r connection weights to that node equals one.

We can now define GNAR processes as follows.

Definition 1. (Generalised Network Autoregressive Process). (Knight et al., 2016, 2020). Let $\mathcal{G} = (\mathcal{K}, \mathcal{E})$ be a graph and $\mathbf{X}_t = (X_{1,t}, \dots, X_{N,t})^T$ be an N -dimensional vector time series, i.e. $X_{i,t}$ corresponds to the time series at the i -th node/vertex on the graph $\mathcal{G}$ for $i = 1, \dots, N$. Then, the generalised autoregressive process of order $(p, \mathbf{s}) \in \mathbb{N} \times \mathbb{N}_0^p$ for $\mathbf{X}_t$, denoted $\text{GNAR}(p, \mathbf{s})$, is given by

$$X_{i,t} = \sum_{j=1}^p \left(\alpha_{i,j} X_{i,t-j} + \sum_{r=1}^{s_j} \beta_{i,j,r} \sum_{q \in \mathcal{N}_{i,t}^{(r)}} \omega_{i,q}^{(t)} X_{q,t-j} \right) + u_{i,t}, \quad (3)$$

for $i \in \mathcal{K}$ and $t \in \{1, \dots, T\} = \mathcal{T}$ and where p is the maximum time lag, $\mathbf{s} = (s_1, \dots, s_p)$ and $s_j \in \mathbb{N}_0$ the maximum stage of neighbour dependence for lag j , with $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$, $\mathcal{N}_{i,t}^{(r)}$ is the r -th stage neighbour set of node i at time t and $\omega_{i,q}^{(t)} \in [0, 1]$ is the connection weight between node i and q at time t . The $\{u_{i,t}\}_{i \in \mathcal{K}, t \in \mathcal{T}}$ is assumed to be an independent and identically distributed set of random variables each with zero mean and variance of σ_i^2 . The $\alpha_{i,j} \in \mathbb{R}$ represent the ‘standard’ autoregressive coefficients for $i \in \mathcal{K}$ and $j = 1, \dots, p$, whereas the $\beta_{i,j,r} \in \mathbb{R}$ capture the influence of r -th stage neighbours at lag j on node i . Here $r = 1, \dots, s_j$, where $s_j \geq 0$ are integers for $j = 1, \dots, p$.

In the standard GNAR model formulation due to Knight et al. (2020), the parameters $\beta_{i,j,r}$ are global, meaning that $\beta_{i,j,r} = \beta_{j,r}$. Hence, we refer to the model in which the $\beta_{i,j,r}$ are allowed to vary from node to node as a local- $\alpha\beta$ $\text{GNAR}(p, \mathbf{s})$ process, distinguishing it from the original model. Additionally, Knight et al. (2020) define a global- α $\text{GNAR}(p, \mathbf{s})$ model where the parameters $\alpha_{i,j}$ are shared among nodes, or more formally $\alpha_{i,j} = \alpha_j$.

3. Consumer Price Index Data

The Consumer Price Index (CPI) is a key indicator used to measure inflation, which tracks changes in the cost of a basket of approximately 700 goods and services consumed by households in the United Kingdom. This basket is carefully designed to represent typical consumer spending patterns, with the weight assigned to each item reflecting its relative importance in the average household’s expenditure. Both the composition of the basket and the weights assigned to each item are reviewed and adjusted annually to ensure they remain relevant to current trends in the economy. Each month, the Office for National Statistics (ONS) gathers price data from a wide range of locations across the country, including physical retail outlets, online stores and service providers. These prices are first combined to produce item-level indices, which are then systematically aggregated to create broader indices that track prices at the class, group, and division levels. Finally, these division-level indices are combined to produce the overall CPI, which is used to calculate the inflation rate (Tucker and Gooding, 2017; Consumer Prices Team, 2024). An excellent introduction

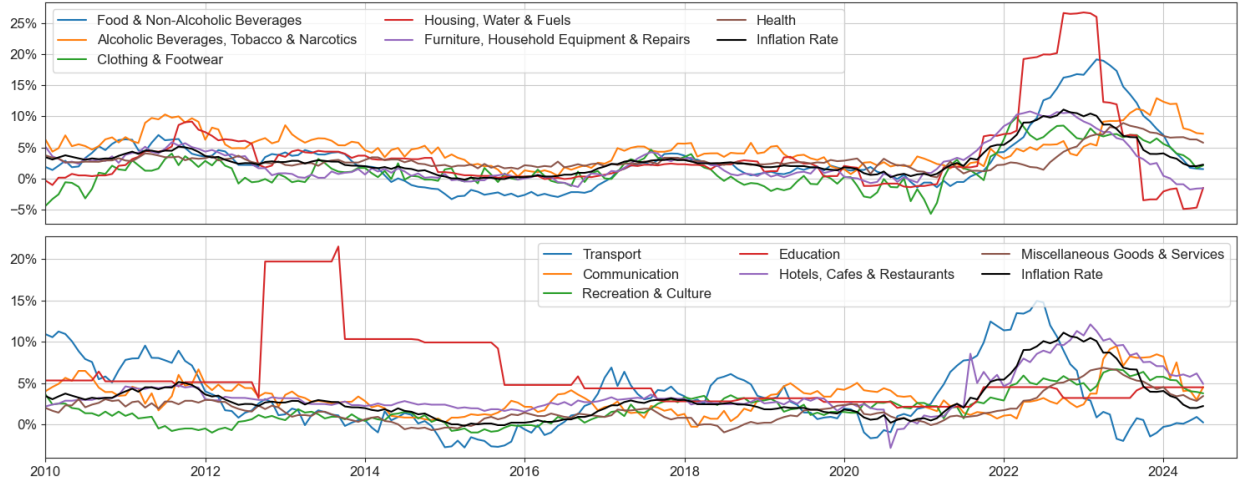


Figure 1: Year-on-year percentage change time series of the Consumer Price Index divisions.

to price indices and CPI can be found in [Ralph et al. \(2015\)](#). The twelve divisions are food and non-alcoholic beverages, alcoholic beverages and tobacco, clothing and footwear, housing, water, electricity, gas and other fuels, furniture, household equipment and maintenance, health, transport, communication, recreation and culture, education, restaurants and hotels and miscellaneous goods and services.

The data used here consists of monthly observations from January 1988 to the present, sourced from the Office for National Statistics (ONS) ([Consumer Price Inflation Team, 2024](#)). As of the latest release, the twelve main divisions are further broken down into 40 groups, 71 classes, and 192 item-level series. Our study considers time series up to and including the class level, excluding any series consisting of annual observations. In total, we consider $N = 114$ individual time series, including the overall CPI. These series are transformed into year-on-year percentage changes to mitigate for trends and seasonality, and to help stabilise the variance.

For visualisation purposes, Figure 1 presents the year-on-year percentage changes for the twelve CPI division-level series and the inflation rate. Here, we clearly see that the Education division is of a yearly nature, and in fact its sub-components are excluded from our analysis. Disaggregated CPI component series can contain valuable information, revealing dynamics that may not be immediately visible in the aggregated CPI. For example, from 2020 onwards, the Transport division time series appears to anticipate movements in the overall inflation rate, illustrating how certain sub-components might act as early signals. For a more in-depth review of the CPI disaggregated item series, including descriptive statistics, see Section 2 of [Joseph et al. \(2024\)](#).

4. Modelling and Forecasting Methodology

4.1. Network Construction

Our approach involves selecting the best networks each month from a set of possible networks. With $N = 114$ nodes (separate time series) in the dataset, this corresponds to a possible 2^{6441} distinct networks, making it computationally infeasible to iterate through all of them to see how well each one can be used to forecast. Instead, we propose selecting an ‘optimal’ (or very good) network(s) from a set of G undirected random networks generated according to the Erdős–Rényi–Gilbert model (Gilbert, 1959), where each edge exists with probability π . The next two propositions provide some insight into the behaviour of the neighbour sets as a function of π .

Proposition 1. *Let $(\mathcal{G}, \mathcal{E})$ be a graph generated according to the Erdős–Rényi–Gilbert model (Gilbert, 1959), where edges exist with probability π and $|\mathcal{G}| = N$. Then, for any $i \in \mathcal{G}$,*

$$\mathbb{P}\left(|\mathcal{N}_i^{(r)}| = n_r\right) = \sum_{\mathbf{n} \in \Lambda_{r-1}(N)} \prod_{s=1}^r \mathbb{P}(Z_{n_1:n_{s-1}}(N) = n_s), \quad (4)$$

for $n_r \in \{0, 1, \dots, N - r\}$, where

$$\Lambda_{r-1}(N) = \left\{ \mathbf{n} = (n_1, \dots, n_{r-1}) \in \mathbb{N}^{r-1} : \sum_{s=0}^{r-1} n_s \leq N \right\} \quad (5)$$

is the set of all possible combinations of neighbour set sizes up to stage $r - 1$,

$$Z_{n_1:n_{s-1}}(N) \sim \text{Bin}\left(N - \sum_{k=0}^{s-1} n_k, 1 - (1 - \pi)^{n_{s-1}}\right) \quad (6)$$

and $n_0 = 1$.

Proposition 2. *Let $(\mathcal{G}, \mathcal{E})$ be a graph generated according to the Erdős–Rényi–Gilbert model (Gilbert, 1959), where edges exist with probability π and $|\mathcal{G}| = N$. Then, for any distinct $i, j \in \mathcal{G}$,*

$$\pi_r = \mathbb{P}\left(j \in \mathcal{N}_i^{(r)}\right) \quad (7)$$

$$= \left(1 - \sum_{s=1}^{r-1} \pi_s\right) \left\{ 1 - \sum_{n=0}^{N-r} (1 - \pi)^n \sum_{\mathbf{n} \in \Lambda_{r-2}(N-1)} \prod_{s=1}^{r-1} \mathbb{P}(Z_{n_1:n_{s-1}}(N-1) = n) \right\} \quad (8)$$

where $\pi_1 = \pi$, and $\Lambda_{r-2}(N-1)$ and $Z_{n_1:n_{s-1}}(N-1)$ are as defined in Proposition 1.

Remark 1. For the $r = 2$ case the result in Proposition 2 simplifies to

$$\pi_2 = (1 - \pi) \left\{ 1 - (1 - \pi^2)^{N-2} \right\}, \quad (9)$$

since there is only a single binomial to consider, namely $Z_{n_1}(N-1) \sim \text{Bin}(N-2, \pi)$.

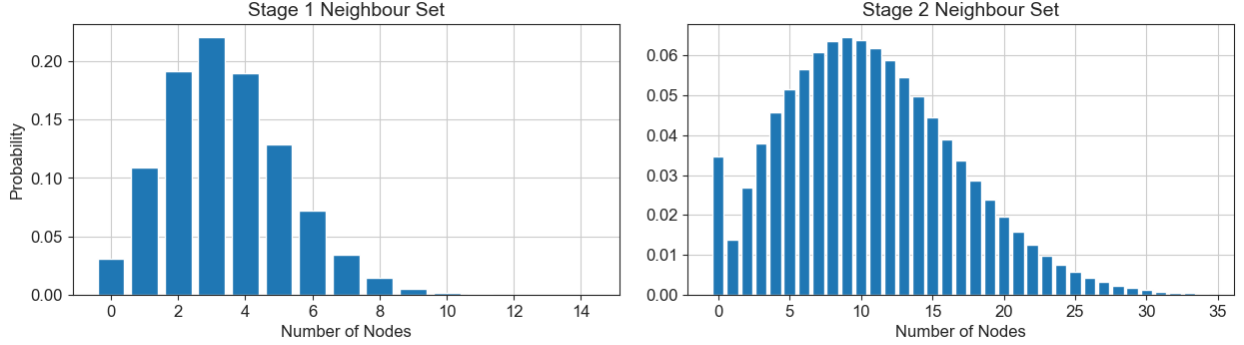


Figure 2: Probability distribution of stage-1 neighbour set size when $N = 114$ and $\pi = 0.03$.

Figure 2 shows the distribution of the stage-1 and stage-2 neighbour set sizes for graphs sampled according to the Erdős–Rényi–Gilbert model (Gilbert, 1959) with $\pi = 0.03$. This setting provides a good range of smaller stage-1 neighbour set sizes, ensuring that only the most influential components are selected as connections to the CPI node each month, as described next. The stage-2 neighbour set allows for exploring larger sizes without the risk of excessive growth, as averaging too many components would result in a time series that mirrors the CPI time series itself. The increased probability of a stage-2 neighbour set size of zero is due to the fact that a stage-1 neighbour set of size zero automatically leads to the same for stage-2. Empirically, we find that generating $G = 10000$ random graphs is sufficient for identifying some networks that perform very well (and larger values did not result in much further improvement). We also tried edge probabilities of $\pi = 0.05$ and $\pi = 0.07$, but there were no discernible performance improvement. The lower chosen value of $\pi = 0.03$ also means fewer neighbours on average, which makes it easier for our algorithm to discern more dominant nodes (time series).

Throughout this article, the networks we consider are constructed from unweighted edges, these being sampled from a Bernoulli distribution with probability of success π . Hence, the GNAR($p, \mathbf{s}$) process (3) can be written as

$$X_{i,t} = \sum_{j=1}^p \left(\alpha_{i,j} X_{i,t-j} + \sum_{r=1}^{s_j} \beta_{i,j,r} \overline{\mathcal{N}}_{i,t-j}^{(r)} \right) + u_{i,t}, \quad (10)$$

where $\overline{\mathcal{N}}_{i,t-j}^{(r)}$ denotes the average over the stage r neighbour set of node i at time $t - j$. Having fitted the model, we can then obtain the one-month ahead forecast as

$$\hat{X}_{i,t+1|t} = \sum_{j=1}^p \left(\hat{\alpha}_{i,j} X_{i,t-j+1} + \sum_{r=1}^{s_j} \hat{\beta}_{i,j,r} \overline{\mathcal{N}}_{i,t-j+1}^{(r)} \right), \quad (11)$$

iterating this procedure in order to compute the h -step ahead forecast $\hat{X}_{i,t+h|t}$.

In science it is important to discuss ideas that did *not* work so well as well as those that did. We briefly mention an early idea that we tried that did not succeed. The construction

of the Consumer Price Index itself closely resembles a graph: individual item series can be progressively combined into larger components in a tree-like network, using basket weights as connection weights, culminating in the final CPI product. This structure suggests a natural fit for applying GNAR processes to the graph, with the aim of generating inflation rate forecasts. However, the inherent properties of GNAR models pose challenges in this context. Specifically, the CPI neighbour sums in such a network then often become redundant relative to the CPI time series itself, leading to parameter estimation instability. Additionally, this method inherently includes edges between CPI components, which may not always contribute meaningfully to forecasting accuracy.

4.2. Model Fitting and Selection

We consider GNAR models of orders $p = 1$ and $p = 2$. Empirically, we find that going past $p = 2$ does not yield improvements in performance, which agrees with the wider literature on autoregressive processes for forecasting inflation. We consider neighbour stage dependence, $\mathbf{s}$, up to and including stage two, again finding little benefit in including higher stages. Stage three neighbour sets can become quite large in size and hence, when included in the model, correspond to an average of a large number of diverse nodes, information that should already be contained in the CPI node. In addition, we keep the neighbour stage fixed at all lags. The models we consider are therefore global- α , standard and local- $\alpha\beta$ GNAR(1, **1**), GNAR(1, **2**), GNAR(2, **1**) and GNAR(2, **2**) processes, where $\mathbf{s}$ denotes a vector of length p with all entries equal to s .

Having generated $G = 10000$ undirected graphs, we fit the GNAR models on a rolling basis using $n_{train} = 240$ observations, equivalent to 20 years of data, and use these to compute an out-of-sample one-month ahead inflation forecast. The models are then ranked each month based on the root mean squared error (RMSE) of the previous $n_{val} = 30$ out-of-sample one-step ahead inflation rate forecasts. Formally, the RMSE for the k -th GNAR model at time t is given by

$$\text{RMSE}_t^{(k)} = \left\{ n_{val}^{-1} \sum_{i=0}^{n_{val}-1} \left(X_{t-i} - \hat{X}_{t-i|t-i-1}^{(k)} \right)^2 \right\}^{1/2}, \quad (12)$$

where $\hat{X}_{t|t-1}^{(k)}$ denotes the forecasts of the k -th GNAR model. The best performing models are then re-fitted with the data up to time t and used to compute inflation rate forecasts for times $t + 1$ up to $t + 12$.

In selecting the best network, we consider the one-step ahead RMSE at (only) the node corresponding to the CPI time series as a metric for performance. Hence, we only fit the GNAR model to that node, speeding the computation up significantly. Once a graph, or set of graphs, is chosen for a given month, we then fit the entire model and forecast CPI inflation multiple months ahead. A potential drawback is that, by selecting networks like so, edges between nodes excluding the CPI one are essentially chosen at random. However, since these other nodes only affect CPI through second order effects and, even so, these effects consist of averages over sets of nodes, this randomness should have a minimal impact on the effectiveness of our method, and may even be beneficial in preventing our models

from over-fitting. Moreover, we will also compute forecasts by averaging the inflation rate predictions of several GNAR processes, further diminishing the importance of the specific network structure outside that of the CPI node, while allowing information to flow between the disaggregated item series.

The process of trying each random graph can be carried out quite independently for each network. Hence, our method would be simple to parallelise. However, the results below, fitting a new model each month from 2009, and then selecting the best network, can already be carried out in twenty minutes on a regular single CPU machine/processing core.

For comparison, we also consider an algorithm which constructs a directed network based on the one-month ahead RMSE at all nodes. To do so, we select the best graph for each node separately, and then construct a single directed network by placing edges from the stage one neighbours of a given node to that node itself. Formally, let $\mathcal{G}_t^i$ be the network with the lowest RMSE at node i at time t and let $\mathcal{G}_t$ denote the directed graph we are constructing. Then, for each CPI component i , $j \rightsquigarrow i$ in $\mathcal{G}_t$ only if $j \in \mathcal{N}(i)$ in $\mathcal{G}_t^i$. By using directed edges, each node is only directly affected by its own stage one neighbours. Unfortunately, this approach only works for constructing graphs for GNAR processes of neighbour stage dependence of at most one. However, as we will see in Section 5, this approach offers little to no benefit, especially given its significantly higher computational complexity.

The basic idea of using random networks in GNAR modelling and prediction was introduced by Leeming (2019); Knight et al. (2020), but our work has tailored the idea for CPI forecasting and there are several key differences. The original Leeming (2019) work was additionally interested in finding ‘the’ ‘correct’ network, which may or may not exist (and in many real cases probably does not). Here, we only use different random networks as a means to an end and we do not believe any one network is the ‘correct’ network, or at least not for any sustained period. Having said that, we do note the components that enter the best set of networks to get an idea of which CPI components are important for forecasting, see Section 6. In contrast with Leeming (2019), our method also does not use connection weights — the averages in (10) are ordinary and not weighted averages, in spite of having some idea of what connection weights might be by considering functions of the CPI basket weights and sales volumes. A key innovation is that prediction accuracy in Leeming (2019) was an aggregate prediction error taken over all time series, whereas here we focus intensely on the single CPI variable (as mentioned in the previous but one paragraph), which confers considerable speed and prediction accuracy advantages. The other major innovation is that we consider both local- α and local- $\alpha\beta$ models.

4.3. Benchmark Models

As suggested in the introduction and highly-respected articles such as Joseph et al. (2024), we use random walks and autoregressive processes as benchmark models, which have been consistently found to be difficult to beat and among the best forecasting techniques that rely on the historical data alone.

Random Walk (RW) Models: (Atkeson and Ohanian, 2001; Stock and Watson, 2007; Faust and Wright, 2013). The forecast of a RW of order n , which we denote $\text{RW}(n)$ is defined

Horizon	1	3	6	9	12
AR(1)	0.43	0.92	1.67	2.41	3.08
AR(2)	0.42	0.89	1.60	2.29	2.88
RW(1)	0.43	0.90	1.57	2.20	2.71
RW(4)	0.73	1.21	1.87	2.43	2.88

Table 1: RMSEs for benchmark model CPI rate forecasts from 2009 onwards against forecast horizon (months).

as the average of the past n observations

$$\hat{y}_{t+h|t} = \frac{1}{n} \sum_{i=0}^{n-1} y_{t-i}, \quad (13)$$

for all horizons h . We consider both a naive RW(1) forecast, as well as RW forecasts which are constructed as averages of the last four monthly inflation observations.

Autoregressive (AR) Processes: (Brockwell and Davis, 1991; Stock and Watson, 2003; Faust and Wright, 2013) An AR process of order p , which we denote $AR(p)$, is defined as

$$y_t = \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} + u_t, \quad (14)$$

where $\phi_i \in \mathbb{R}$ are the coefficients of the model in the zone of stationarity with $\phi_p \neq 0$ and the ‘noise’ $\{u_t\}$ is a set of zero mean constant variance independent and identically random variables. We consider both AR(1) and AR(2) models, fitting these each month using the previous 20 years of data.

Table 1 presents the RMSEs for our benchmark models across different horizons. The AR(2) and RW(1) models perform the best, with the RW(1) model providing more accurate inflation forecasts at longer horizons.

Several related works use the Phillips curve as a benchmark as well. Broadly speaking, the Phillips curve can be viewed as equivalent to an $AR(p)$ process with an additional measure of economic activity as an extra predictor, often the unemployment rate (Stock and Watson, 1999; Atkeson and Ohanian, 2001; Stock and Watson, 2009; Faust and Wright, 2013). However, the literature seems to suggest that the Phillips curve does not provide significant improvements over other univariate time series models, and, in fact, Atkeson and Ohanian (2001) show that a naive RW forecaster generally outperforms the Phillips curve. Other benchmarks often also include the direct forecasting method (Faust and Wright, 2013), which have been shown to be outperformed by the iterative forecasts of an autoregressive model (Marcellino et al., 2006), and hence we do not include these here either.

5. Results

This section presents the results for the various GNAR models discussed earlier, including a comparison of our forecasts with those released by the Bank of England. We generate 100

sets of $G = 10000$ random networks according to the Erdős–Rényi–Gilbert model (Gilbert, 1959) and compute inflation rate forecasts based on each of the methods described previously. This approach allows us to estimate the average RMSE across the 100 sets of random graphs, as well as the standard deviation in the performance. The results are detailed in the tables throughout this section.

5.1. Best Performing Graph

We first present results obtained by selecting the best network at each step and using it to forecast CPI inflation for the next twelve months. Table 2 displays the average relative-to-benchmark RMSE over 100 runs for different GNAR processes. We can see that all three classes of models (global- α , standard and local- $\alpha\beta$) outperform *all* of the benchmarks. Overall, the standard GNAR(2, 1) model seems to be the best overall, with the local- $\alpha\beta$ models also faring quite well, at the cost of a higher number of parameters. The global- α GNAR models are able to compete with the benchmarks, with the GNAR(1, 1) being able to beat all of them most comfortably. This becomes more impressive once one considers the number of parameters involved: after having removed the mean from each time series, the global- α GNAR(1, 1) model only has two parameters in total.

By contrast a standard vector autoregressive model of order one, VAR(1), contains $114^2 = 12996$ possible parameters. Our CPI data consists of twenty years of monthly data for $N = 114$ time series, which equates to 27360 observations in total. The raw data size to parameter ratio for GNAR is dramatically better than for standard VAR, which goes some way to explaining its attractive predictive abilities; GNAR models are rarely in danger of overfitting. Of course, the parameter explosion in VAR models is well known and so modellers often use sparse or restricted VAR models, which cut down the number of parameters fitted, but even here GNAR often outpredicts them, see Nason and Wei (2022); Nason et al. (2023, 2024), for example. Indeed, we also tried VAR models (not reported) but they were not competitive.

5.2. Model Averaging

We now consider averaging the predictions of the GNAR processes fitted to the top n networks. Tables 3 and 4 compare our forecasts for $n = 2$ and $n = 5$ against those of the benchmark models. Our results improve as n increases, while displaying the same patterns as for the $n = 1$ case discussed in the previous sub-section. For instance, with $n = 5$, our forecasts consistently outperform all benchmarks, except for the GNAR(2, 2), which is only marginally surpassed by a first-order random walk. Moreover, as n increases, the RMSE standard deviations decrease, indicating that averaging predictions across multiple networks yields more robust forecasts.

Figure 3 compares the forecasts for the $n = 5$ case to the AR(2) benchmark models from 2018 onwards. This period was marked by high inflation rates, reaching 11% in October 2022 before falling back to the Bank of England’s target of 2% in the second quarter of 2024. As we can see from the plots, the GNAR models are significantly faster than the benchmarks at noticing that inflation began stabilising in 2023. This is a real and clear advantage of our

Class	GNAR Model	Number of Parameters	Forecast Horizon (months)				
			1	3	6	9	12
global- α	GNAR(1,1)	2	0.93 $\pm$ 0.04	0.88 $\pm$ 0.05	0.82 $\pm$ 0.04	0.78 $\pm$ 0.03	0.74 $\pm$ 0.02
	GNAR(1,2)	3	0.93 $\pm$ 0.03	0.89 $\pm$ 0.05	0.86 $\pm$ 0.05	0.84 $\pm$ 0.04	0.81 $\pm$ 0.04
	GNAR(2,1)	4	0.90 $\pm$ 0.03	0.84 $\pm$ 0.03	0.80 $\pm$ 0.03	0.79 $\pm$ 0.03	0.78 $\pm$ 0.04
	GNAR(2,2)	6	0.92 $\pm$ 0.06	0.89 $\pm$ 0.06	0.86 $\pm$ 0.06	0.86 $\pm$ 0.06	0.85 $\pm$ 0.06
standard (local- α)	GNAR(1,1)	115	0.93 $\pm$ 0.04	0.85 $\pm$ 0.05	0.79 $\pm$ 0.05	0.78 $\pm$ 0.05	0.77 $\pm$ 0.05
	GNAR(1,2)	116	0.91 $\pm$ 0.03	0.85 $\pm$ 0.04	0.80 $\pm$ 0.05	0.79 $\pm$ 0.05	0.79 $\pm$ 0.05
	GNAR(2,1)	230	0.89 $\pm$ 0.03	0.79 $\pm$ 0.02	0.72 $\pm$ 0.03	0.72 $\pm$ 0.04	0.71 $\pm$ 0.04
	GNAR(2,2)	232	0.90 $\pm$ 0.06	0.82 $\pm$ 0.04	0.78 $\pm$ 0.05	0.78 $\pm$ 0.08	0.77 $\pm$ 0.06
local- $\alpha\beta$	GNAR(1,1)	228	0.89 $\pm$ 0.02	0.79 $\pm$ 0.02	0.72 $\pm$ 0.03	0.71 $\pm$ 0.03	0.70 $\pm$ 0.03
	GNAR(1,2)	342	0.86 $\pm$ 0.02	0.78 $\pm$ 0.03	0.72 $\pm$ 0.03	0.71 $\pm$ 0.04	0.71 $\pm$ 0.04
	GNAR(2,1)	456	0.83 $\pm$ 0.01	0.80 $\pm$ 0.02	0.76 $\pm$ 0.03	0.75 $\pm$ 0.03	0.74 $\pm$ 0.04
	GNAR(2,2)	684	0.83 $\pm$ 0.02	0.80 $\pm$ 0.03	0.76 $\pm$ 0.03	0.76 $\pm$ 0.04	0.76 $\pm$ 0.05

Table 2: Relative to benchmark RMSEs for inflation rate forecasts computed using a GNAR model fit to the best performing network each month. RMSEs are computed by averaging the results from 100 runs of the algorithm, with ± 1 standard deviation in a smaller font size. All RMSEs are relative to the AR(2) benchmark.

methods. For example, this can clearly be seen in the global- α and standard GNAR(2,1) forecasts.

5.3. Directed Graphs

We now move from undirected graph to directed ones. Recall that to construct these, we firstly fit a complete GNAR process to each undirected network at every time step. We then create a new directed graph by combining the networks with the lowest RMSEs for each node. Specifically, we add directed edges from the stage one neighbour set to the node itself, using the neighbour set from the network with the lowest RMSE for that particular node. The results, summarised in Table 5, closely resemble those obtained for the undirected networks discussed in Section 5.1. This similarity suggests that constraining the network’s structure offers little, if any, advantage. In fact, allowing the network structure to remain random may help prevent overfitting.

5.4. Bank of England Forecasting

The Bank of England employs a forecasting platform consisting of a central model, COMPASS, which is a New Keynesian general equilibrium model, a suite of at least 40 other models, including ARMA and survey models for example, as well as judgemental insights from the Monetary Policy Committee (Burgess et al., 2013; Bernanke, 2024). Each forecasting round begins from the previous inflation rate forecasts and builds on these incrementally, combining these with new data as it becomes available. There are at least five meetings between the Monetary Policy Committee and the Bank staff, during which

GNAR		Forecast Horizon (months)				
Model Class	Model	1	3	6	9	12
global- α	GNAR(1,1)	0.91 $\pm$ 0.02	0.85 $\pm$ 0.03	0.80 $\pm$ 0.03	0.77 $\pm$ 0.02	0.74 $\pm$ 0.02
	GNAR(1,2)	0.91 $\pm$ 0.02	0.87 $\pm$ 0.03	0.84 $\pm$ 0.03	0.82 $\pm$ 0.03	0.80 $\pm$ 0.03
	GNAR(2,1)	0.89 $\pm$ 0.03	0.83 $\pm$ 0.02	0.79 $\pm$ 0.02	0.78 $\pm$ 0.02	0.78 $\pm$ 0.02
	GNAR(2,2)	0.89 $\pm$ 0.03	0.87 $\pm$ 0.04	0.85 $\pm$ 0.04	0.85 $\pm$ 0.04	0.85 $\pm$ 0.04
standard (local- α)	GNAR(1,1)	0.92 $\pm$ 0.03	0.84 $\pm$ 0.03	0.78 $\pm$ 0.04	0.77 $\pm$ 0.04	0.76 $\pm$ 0.04
	GNAR(1,2)	0.89 $\pm$ 0.02	0.83 $\pm$ 0.03	0.79 $\pm$ 0.04	0.79 $\pm$ 0.04	0.78 $\pm$ 0.04
	GNAR(2,1)	0.88 $\pm$ 0.02	0.78 $\pm$ 0.02	0.72 $\pm$ 0.02	0.71 $\pm$ 0.03	0.70 $\pm$ 0.03
	GNAR(2,2)	0.88 $\pm$ 0.04	0.80 $\pm$ 0.03	0.76 $\pm$ 0.04	0.77 $\pm$ 0.05	0.77 $\pm$ 0.04
local- $\alpha\beta$	GNAR(1,1)	0.88 $\pm$ 0.02	0.78 $\pm$ 0.02	0.71 $\pm$ 0.02	0.70 $\pm$ 0.02	0.69 $\pm$ 0.02
	GNAR(1,2)	0.84 $\pm$ 0.01	0.76 $\pm$ 0.02	0.70 $\pm$ 0.02	0.70 $\pm$ 0.02	0.70 $\pm$ 0.03
	GNAR(2,1)	0.82 $\pm$ 0.01	0.79 $\pm$ 0.01	0.75 $\pm$ 0.02	0.74 $\pm$ 0.02	0.73 $\pm$ 0.03
	GNAR(2,2)	0.81 $\pm$ 0.01	0.78 $\pm$ 0.02	0.74 $\pm$ 0.03	0.75 $\pm$ 0.03	0.75 $\pm$ 0.04

Table 3: RMSEs for inflation rate forecasts computed by averaging predictions from GNAR models fit to the $n = 2$ best performing networks each month. RMSEs are computed by averaging the results from 100 runs of the algorithm, with ± 1 standard deviation in a smaller font. All RMSEs are relative to the AR(2) benchmark.

GNAR		Forecast Horizon (months)				
Model Class	Model	1	3	6	9	12
global- α	GNAR(1,1)	0.90 $\pm$ 0.01	0.84 $\pm$ 0.02	0.79 $\pm$ 0.02	0.76 $\pm$ 0.01	0.73 $\pm$ 0.01
	GNAR(1,2)	0.90 $\pm$ 0.01	0.86 $\pm$ 0.02	0.83 $\pm$ 0.02	0.82 $\pm$ 0.02	0.79 $\pm$ 0.02
	GNAR(2,1)	0.88 $\pm$ 0.02	0.82 $\pm$ 0.02	0.78 $\pm$ 0.02	0.78 $\pm$ 0.01	0.77 $\pm$ 0.02
	GNAR(2,2)	0.88 $\pm$ 0.03	0.86 $\pm$ 0.02	0.84 $\pm$ 0.02	0.84 $\pm$ 0.02	0.84 $\pm$ 0.03
standard (local- α)	GNAR(1,1)	0.90 $\pm$ 0.02	0.82 $\pm$ 0.02	0.76 $\pm$ 0.03	0.76 $\pm$ 0.03	0.76 $\pm$ 0.03
	GNAR(1,2)	0.89 $\pm$ 0.01	0.82 $\pm$ 0.02	0.78 $\pm$ 0.03	0.78 $\pm$ 0.03	0.78 $\pm$ 0.03
	GNAR(2,1)	0.87 $\pm$ 0.01	0.78 $\pm$ 0.01	0.71 $\pm$ 0.01	0.71 $\pm$ 0.02	0.70 $\pm$ 0.02
	GNAR(2,2)	0.86 $\pm$ 0.02	0.79 $\pm$ 0.02	0.76 $\pm$ 0.03	0.77 $\pm$ 0.03	0.77 $\pm$ 0.03
local- $\alpha\beta$	GNAR(1,1)	0.88 $\pm$ 0.01	0.77 $\pm$ 0.01	0.70 $\pm$ 0.01	0.69 $\pm$ 0.01	0.68 $\pm$ 0.02
	GNAR(1,2)	0.84 $\pm$ 0.01	0.75 $\pm$ 0.01	0.69 $\pm$ 0.02	0.69 $\pm$ 0.02	0.69 $\pm$ 0.02
	GNAR(2,1)	0.82 $\pm$ 0.01	0.79 $\pm$ 0.01	0.74 $\pm$ 0.01	0.74 $\pm$ 0.01	0.72 $\pm$ 0.02
	GNAR(2,2)	0.80 $\pm$ 0.01	0.77 $\pm$ 0.01	0.74 $\pm$ 0.02	0.74 $\pm$ 0.02	0.74 $\pm$ 0.02

Table 4: RMSEs for inflation rate forecasts computed by averaging predictions from GNAR models fit to the $n = 5$ best performing networks each month. RMSEs are computed by averaging the results from 100 runs of the algorithm, with ± 1 standard deviation in a smaller font. All RMSEs are relative to the AR(2) benchmark.

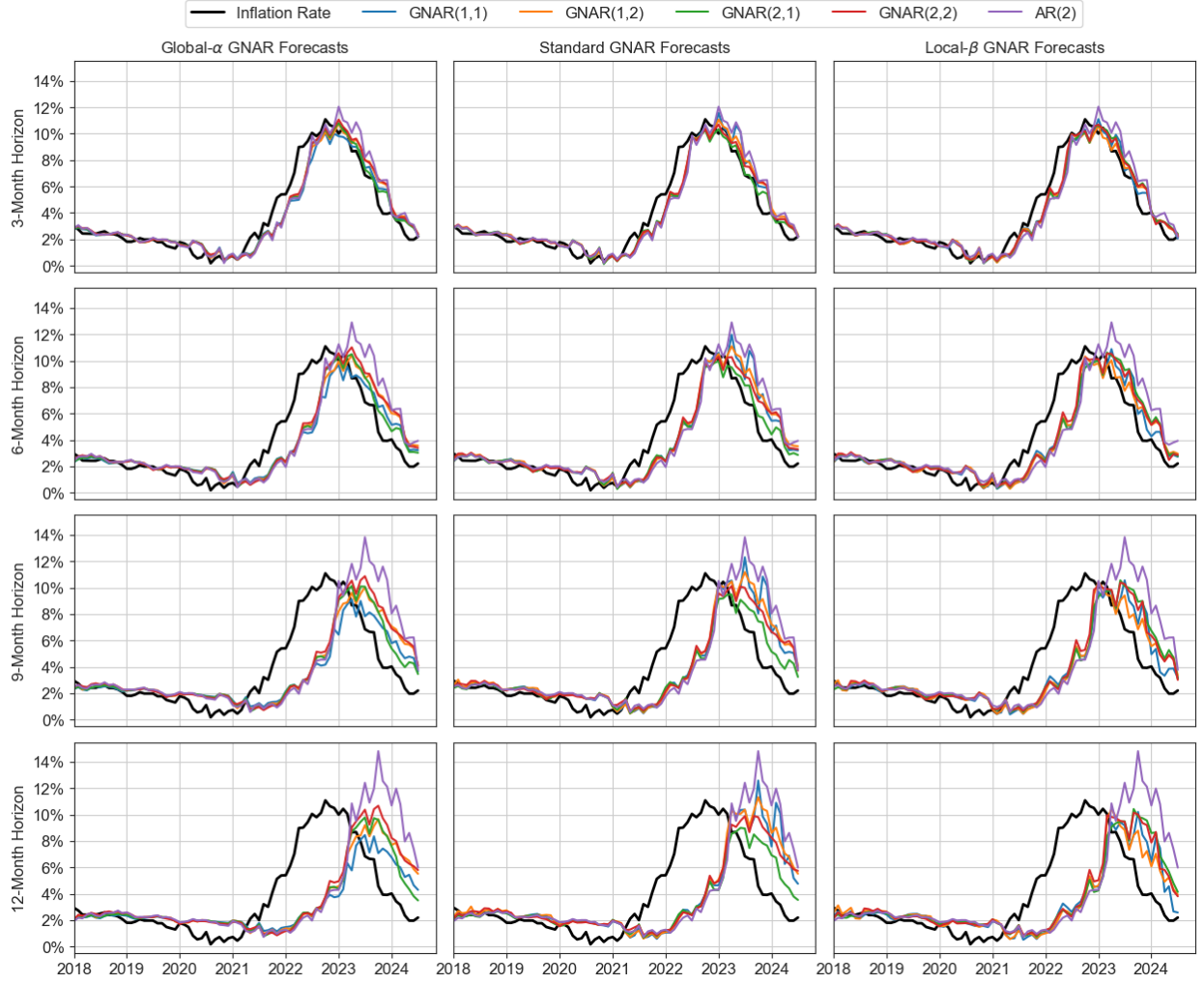


Figure 3: A comparison of inflation forecasts computed by averaging the predictions of GNAR models fit to the $n = 5$ best networks each month against the AR(2) benchmark model.

GNAR		Forecast Horizon (months)				
Model Class	Model	1	3	6	9	12
global- α	GNAR(1,1)	0.98 $\pm$ 0.05	0.91 $\pm$ 0.06	0.86 $\pm$ 0.04	0.84 $\pm$ 0.03	0.81 $\pm$ 0.02
	GNAR(2,1)	0.93 $\pm$ 0.02	0.88 $\pm$ 0.04	0.85 $\pm$ 0.04	0.86 $\pm$ 0.04	0.88 $\pm$ 0.04
standard (local- α)	GNAR(1,1)	0.95 $\pm$ 0.02	0.89 $\pm$ 0.03	0.86 $\pm$ 0.04	0.87 $\pm$ 0.5	0.88 $\pm$ 0.05
	GNAR(2,1)	0.90 $\pm$ 0.02	0.81 $\pm$ 0.02	0.77 $\pm$ 0.03	0.78 $\pm$ 0.03	0.80 $\pm$ 0.03
local- $\alpha\beta$	GNAR(1,1)	0.93 $\pm$ 0.02	0.82 $\pm$ 0.02	0.77 $\pm$ 0.03	0.78 $\pm$ 0.03	0.78 $\pm$ 0.03
	GNAR(2,1)	0.86 $\pm$ 0.01	0.82 $\pm$ 0.02	0.80 $\pm$ 0.03	0.81 $\pm$ 0.03	0.82 $\pm$ 0.03

Table 5: RMSEs for inflation rate forecasts computed by fitting a GNAR process to a directed network constructed monthly. RMSEs are derived from averaging the results of 100 runs of the algorithm, with ± 1 standard deviation shown in a smaller font. All RMSEs are relative to the AR(2) benchmark.

judgemental adjustments are made to the forecasts (Burgess et al., 2013; Bernanke, 2024). Moreover, to produce a single set of forecasts each quarter, the COMPASS algorithm is run hundreds of times in order to incorporate insights from the Monetary Policy Committee and explore different scenarios (Burgess et al., 2013; Bernanke, 2024). In recent years, central bank forecasting accuracy has declined, particularly during the sharp rise in inflation following the COVID-19 pandemic, which both central banks and external forecasters struggled to predict, see Bernanke (2024). Further, Bernanke (2024) notes that “the forecast errors made by the Bank and those made by external forecasters are barely distinguishable”, whereas the results below show that our errors are demonstrably better.

We compare our forecasts with those of the Bank of England, published every quarter in the Bank of England’s Monetary Policy Reports (see Monetary Policy Committee (2024) for the latest one). Until recently, the Bank of England released inflation rate predictions for up to three months ahead; this horizon has now been extended to six months. Given that the Bank publishes forecasts on a quarterly basis, our comparisons are limited to the corresponding months.

Table 6 shows results for all models. The Bank of England strongly outperforms our models for horizons of one, two and three months. However, some GNAR models notably surpass the Bank’s forecast for horizons of four, five and six months, with the standard GNAR(2, 2) model emerging as the overall best performer. This is especially noteworthy given the sophistication of the Bank of England’s forecasting procedure, which integrates multiple models, extensive data analysis, and expert judgment. The Bank’s approach involves complex econometric models and regular adjustments by the Monetary Policy Committee to account for the latest economic developments, making it a formidable benchmark. That our GNAR model can outperform such a comprehensive and refined forecasting system over longer horizons highlights the strength and potential of our approach, particularly in capturing trends and dynamics that may become more pronounced over extended periods. Moreover, since the Bank of England’s forecasts are only available from the end of 2019 onwards, Table 6 also highlights how our GNAR models have outperformed the benchmarks

Model	Forecast Horizon (months)					
	1	2	3	4	5	6
Bank of England	0.17	0.34	0.51	0.96	1.05	1.03
AR(1)	0.93	1.18	1.41	2.48	2.81	3.08
AR(2)	0.89	1.14	1.36	2.27	2.55	2.75
RW(1)	0.93	1.16	1.36	2.30	2.58	2.71
RW(4)	1.30	1.58	1.79	2.78	3.07	3.18
GNAR(1,1)	0.71 $\pm$ 0.03	0.88 $\pm$ 0.03	0.97 $\pm$ 0.05	1.06 $\pm$ 0.11	1.21 $\pm$ 0.13	1.18 $\pm$ 0.13
GNAR(1,2)	0.74 $\pm$ 0.03	0.92 $\pm$ 0.03	1.03 $\pm$ 0.05	1.66 $\pm$ 0.11	1.94 $\pm$ 0.14	2.07 $\pm$ 0.16
GNAR(2,1)	0.68 $\pm$ 0.03	0.86 $\pm$ 0.03	0.94 $\pm$ 0.04	1.24 $\pm$ 0.09	1.42 $\pm$ 0.12	1.55 $\pm$ 0.15
GNAR(2,2)	0.70 $\pm$ 0.05	0.92 $\pm$ 0.03	1.03 $\pm$ 0.05	1.69 $\pm$ 0.12	2.01 $\pm$ 0.15	2.22 $\pm$ 0.18
GNAR(1,1)	0.72 $\pm$ 0.03	0.87 $\pm$ 0.03	0.96 $\pm$ 0.05	1.14 $\pm$ 0.26	1.34 $\pm$ 0.30	1.35 $\pm$ 0.35
GNAR(1,2)	0.73 $\pm$ 0.03	0.89 $\pm$ 0.04	0.96 $\pm$ 0.05	1.34 $\pm$ 0.23	1.60 $\pm$ 0.27	1.68 $\pm$ 0.29
GNAR(2,1)	0.67 $\pm$ 0.02	0.80 $\pm$ 0.02	0.85 $\pm$ 0.03	0.69 $\pm$ 0.13	0.83 $\pm$ 0.13	0.79 $\pm$ 0.15
GNAR(2,2)	0.67 $\pm$ 0.04	0.83 $\pm$ 0.03	0.89 $\pm$ 0.05	1.14 $\pm$ 0.22	1.34 $\pm$ 0.28	1.44 $\pm$ 0.27
GNAR(1,1)	0.69 $\pm$ 0.02	0.78 $\pm$ 0.02	0.85 $\pm$ 0.03	0.73 $\pm$ 0.17	0.77 $\pm$ 0.19	0.68 $\pm$ 0.19
GNAR(1,2)	0.66 $\pm$ 0.01	0.75 $\pm$ 0.02	0.77 $\pm$ 0.02	0.74 $\pm$ 0.09	0.84 $\pm$ 0.11	0.80 $\pm$ 0.10
GNAR(2,1)	0.60 $\pm$ 0.01	0.76 $\pm$ 0.01	0.82 $\pm$ 0.02	1.03 $\pm$ 0.08	1.11 $\pm$ 0.12	1.15 $\pm$ 0.13
GNAR(2,2)	0.59 $\pm$ 0.01	0.74 $\pm$ 0.02	0.79 $\pm$ 0.03	1.10 $\pm$ 0.09	1.20 $\pm$ 0.12	1.26 $\pm$ 0.15

Table 6: Comparison of GNAR and Bank of England inflation rate forecasts on the dates they are available. GNAR forecasts are computed using the average of the five best-performing networks at each time step. The RMSE is calculated as the average over 100 runs of the algorithm, with ± 1 standard deviation shown in smaller font. All RMSEs displayed are absolute and not relative to AR(2).

significantly more in recent years.

6. Important CPI Components for Forecasting

This final section analyses the structure of the best-performing undirected graphs to identify the most influential CPI components (nodes) and compares their time series to the inflation rate. To achieve this, we track how often each component/node is included in the stage-1 neighbour set of the CPI node for the networks ranked as the best each month.

Figure 4 displays the percentage of times each component appeared in the CPI component’s neighbour set within the top-performing networks. Throughout most periods, a single component tends to dominate: oils and fats from 2010 to 2012 and 2015 to 2016, fuels and lubricants from 2016 to 2020, and liquid fuels from 2020 onwards. Notably, as one component phases out, another takes its place. Other disaggregated series appear less frequently in the CPI node’s neighbour set, as shown in Figure 4, with the percentages for these components remaining relatively low. The exception is the period from 2012 to 2014, during which components cycled in and out of the neighbour set more frequently, when the

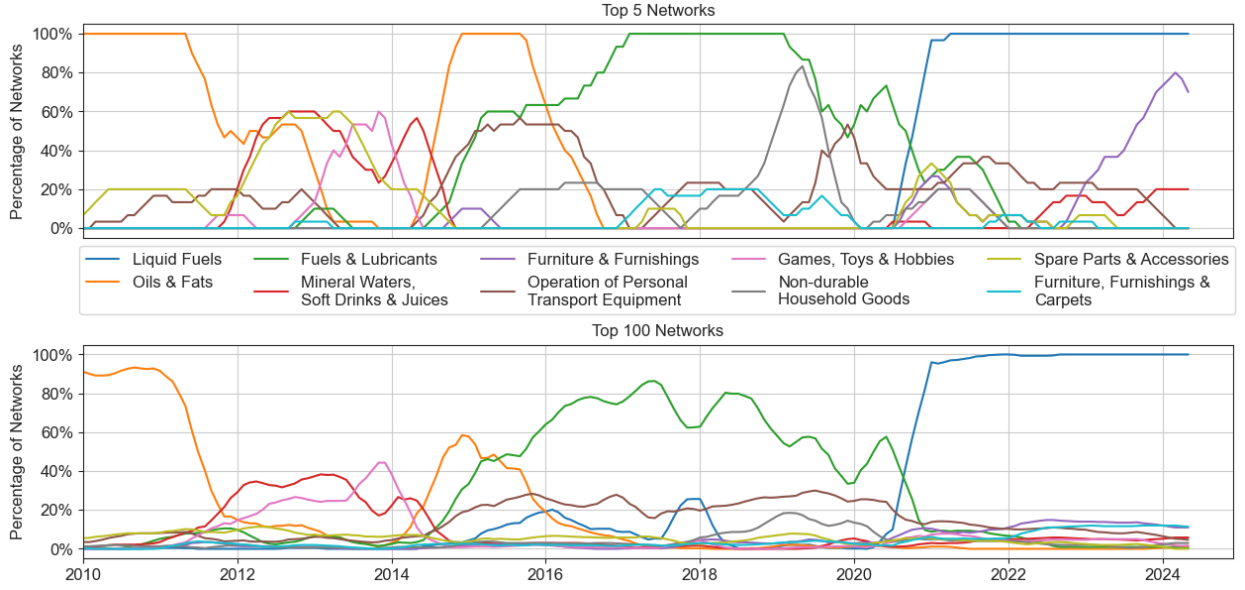


Figure 4: Percentage of times each component appears in the CPI node’s neighbour set for the best-performing networks, smoothed using a 6-month moving average. Only the components which appear most consistently are displayed.

best network selected by the algorithm was also changing more often. A similar pattern emerges between 2019 and 2020. These periods generally coincide with relatively stable inflation rates, meaning that there is less new information to be gained from the disaggregated item series.

Figure 5 displays the time series for the most common disaggregated item series, according to Figure 4. These series tend to anticipate movements in the inflation rate. For example, the year-on-year percentage change in oils & fats precedes the inflation rise in 2011. Similarly, both liquid fuels and fuels & lubricants predict the sharp increase in inflation to over 10% in 2022, followed by a return to more stable levels at the start of 2024. Although these two series display similar behaviour in the figure, this is mainly due to the rescaling and, in reality, liquid fuels exhibit much larger percentage changes than fuels & lubricants. In fact, during recent periods of significant inflationary change, networks that included liquid fuels in the stage-1 neighbour set tended to perform better. Both series also predicted the inflation rise in 2016, but since this increase was more moderate, graphs with an edge between CPI and fuels & lubricants performed better.

Figure 6 displays the neighbour set series for the corresponding best-performing networks, with colours matching those in Figure 5. For instance, the blue neighbour series corresponds to networks where liquid fuels are in the CPI’s neighbour set. These plots offer a clearer view of the dynamics described above. As shown in equation (10), the neighbour sums correspond to averages of multiple disaggregated series. This has the advantage of not only incorporating information from other CPI components, but also of stabilising the magnitudes of year-on-year percentage changes for certain series. Overall, GNAR processes tend to outperform traditional benchmark models, particularly during periods of inflation instability, as they

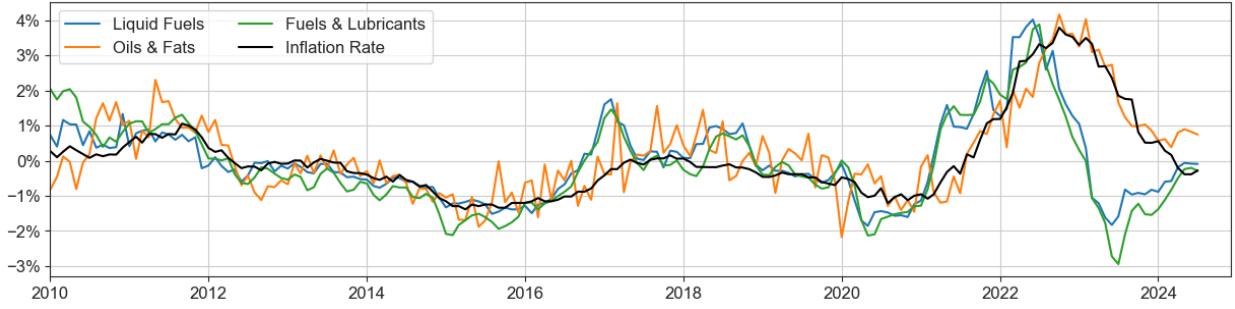


Figure 5: Time series of the components that most frequently appear in the CPI node’s neighbour set. For visualisation purposes, these series have been standardised by subtracting their means and dividing by their standard deviations.

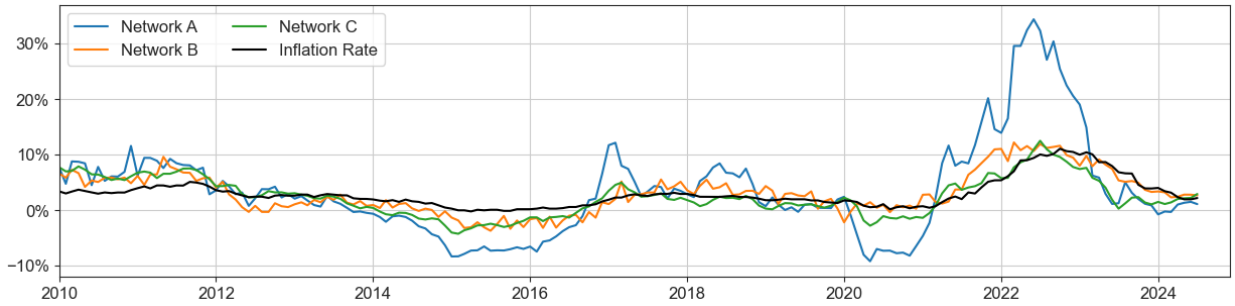


Figure 6: Time series of the stage-1 neighbour set contributions for some of the best-performing networks. The colours correspond to those used in Figure 5, where, for instance, Network A includes liquid fuels in the neighbour set of the CPI node, and similarly for the other networks.

are able to leverage information in a few key disaggregate series in order to anticipate such movements.

7. Conclusion and Future Work

This article has applied the GNAR processes from [Leeming \(2019\)](#); [Knight et al. \(2020\)](#) to networks defined on the CPI component time series to predict inflation rates in the United Kingdom up to 12 months ahead. The networks were constructed according to the Erdős–Rényi–Gilbert model ([Gilbert, 1959](#)), and then ranked based on the 1-step-ahead RMSE at the CPI node. The best graphs were then used to compute our out-of-sample forecasts. Our forecasts consistently outperformed traditional benchmark models, as well as those of the Bank of England at the four, five and six month horizons. This is especially impressive given the simplicity of the GNAR models, which rely on only a (relatively) small number of parameters, particularly when compared to the numerous attempts to outperform the classic benchmarks using more complex machine learning techniques. The strength of our approach lies in its ability to identify a few key CPI components each month while leveraging the GNAR processes’ capacity to estimate parameters from the broader network of time series.

There are several avenues for future work. One potential direction is to explore alternative methods for constructing graphs on the CPI component time series, such as using genetic algorithms—an optimisation technique inspired by natural selection in biological evolution. Starting with a set of random networks, a fitness measure such as the RMSE could be used to iteratively combine the most promising graphs. Another approach could involve applying boosting to an ensemble of graphs, where GNAR models are iteratively fitted to the residuals of previous ones. Additionally, node pruning could be investigated, which entails removing nodes that frequently appear in the worst-performing networks and generating new random graphs on the reduced set. As demonstrated in our work, the disaggregated CPI components contain valuable information that can improve inflation rate forecasts. Accurate inflation forecasting is crucial for effective economic policy-making, investment planning, and financial stability, making further research in this area highly beneficial to the broader economy.

8. Declaration of competing interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

9. Acknowledgments

We would like to thank Fergal Shortall for directing us to relevant literature on the Bank of England forecasting system and explaining some aspects in more detail. Nason gratefully acknowledges support from EPSRC NeST Programme Grant EP/X002195/1. Palasciano gratefully acknowledges support of the EPSRC studentship.

Appendix A. Proofs

Proof of Proposition 1

Let

$$\Lambda_{r-1}(N) = \left\{ \mathbf{n} = (n_1, \dots, n_{r-1}) \in \mathbb{N}^{r-1} : \sum_{s=0}^{r-1} n_s \leq N \right\} \quad (\text{A.1})$$

be the set of all possible combinations of neighbour set sizes up to stage $r - 1$. Applying the law of total probability and Bayes rule, we obtain

$$\mathbb{P}\left(|\mathcal{N}_i^{(r)}| = n_r\right) = \sum_{\mathbf{n} \in \Lambda_{r-1}(N)} \mathbb{P}\left(|\mathcal{N}_i^{(r)}| = n_r \mid |\mathcal{N}_i^{(r-1)}| = n_{r-1}, \dots, |\mathcal{N}_i^{(1)}| = n_1\right) \quad (\text{A.2})$$

$$= \sum_{\mathbf{n} \in \Lambda_{r-1}(N)} \mathbb{P}\left(|\mathcal{N}_i^{(r)}| = n_r \mid |\mathcal{N}_i^{(r-1)}| = n_{r-1}, \dots, |\mathcal{N}_i^{(1)}| = n_1\right) \mathbb{P}\left(|\mathcal{N}_i^{(r-1)}| = n_{r-1} \mid |\mathcal{N}_i^{(r-2)}| = n_{r-2}, \dots, |\mathcal{N}_i^{(1)}| = n_1\right) \dots \mathbb{P}\left(|\mathcal{N}_i^{(2)}| = n_2 \mid |\mathcal{N}_i^{(1)}| = n_1\right) \mathbb{P}\left(|\mathcal{N}_i^{(1)}| = n_1\right) \quad (\text{A.3})$$

$$= \sum_{\mathbf{n} \in \Lambda_{r-1}(N)} \prod_{s=1}^r \mathbb{P}\left(|\mathcal{N}_i^{(s)}| = n_s \mid |\mathcal{N}_i^{(s-1)}| = n_{s-1}, \dots, |\mathcal{N}_i^{(1)}| = n_1\right) \quad (\text{A.4})$$

$$= \sum_{\mathbf{n} \in \Lambda_{r-1}(N)} \prod_{s=1}^r \mathbb{P}\left(Z_{n_1:n_{s-1}}(N) = n_s\right), \quad (\text{A.5})$$

where

$$Z_{n_1:n_{s-1}}(N) = |\mathcal{N}_i^{(s)}| \mid |\mathcal{N}_i^{(s-1)}| = n_{s-1}, \dots, |\mathcal{N}_i^{(1)}| = n_1 \quad (\text{A.6})$$

is a random variable that takes values in the set of all possible sizes of the neighbour set at stage s , conditioned on the sizes of the neighbour sets from stages 1 through $s-1$. Now note that $|\mathcal{N}_i^{(1)}| \sim \text{Bin}(N-1, \pi)$, since there are a total of $n-1$ nodes to choose from, and each of these is contained in $\mathcal{N}_i^{(1)}$ with probability π . More generally, given that $|\mathcal{N}_i^{(1)}| = n_1, \dots, |\mathcal{N}_i^{(s-1)}| = n_{s-1}$, the remaining number of nodes to choose from is $N - (1 + n_1 + \dots + n_{s-1})$. Letting j denote any of the remaining nodes, the probability that this is included in $\mathcal{N}_i^{(s)}$ is given by

$$\mathbb{P}\left(j \in \mathcal{N}_i^{(s)} \mid |\mathcal{N}_i^{(s-1)}| = n_{s-1}\right) = 1 - \mathbb{P}\left(j \notin \mathcal{N}_i^{(s)} \mid |\mathcal{N}_i^{(s-1)}| = n_{s-1}\right) = 1 - (1 - \pi)^{n_{s-1}}, \quad (\text{A.7})$$

since the probability that no edge exists between node j and node $i \in \mathcal{N}_i^{(s-1)}$ is $1 - \pi$, and $j \notin \mathcal{N}_i^{(s)}$ if and only if $j \not\rightsquigarrow i$ for all $i \in \mathcal{N}_i^{(s-1)}$. Hence,

$$Z_{n_1:n_{s-1}}(N) \sim \text{Bin}\left(N - \sum_{k=0}^{s-1} n_k, 1 - (1 - \pi)^{n_{s-1}}\right), \quad (\text{A.8})$$

where $n_0 = 1$, completing the proof. $\square$

Proof of Proposition 2

To begin with, by the law of total probability,

$$\pi_r = \mathbb{P}\left(j \in \mathcal{N}_i^{(r)}\right) = \mathbb{P}\left(j \in \mathcal{N}_i^{(r)} \mid j \notin \bigcup_{s=1}^{r-1} \mathcal{N}_i^{(s)}\right) \mathbb{P}\left(j \notin \bigcup_{s=1}^{r-1} \mathcal{N}_i^{(s)}\right), \quad (\text{A.9})$$

since the probability of $j \in \mathcal{N}_i^{(r)}$ conditioned on j being a member of any $\mathcal{N}_i^{(1)}, \dots, \mathcal{N}_i^{(r-1)}$ is zero. Now

$$\mathbb{P} \left(j \notin \bigcup_{s=1}^{r-1} \mathcal{N}_i^{(s)} \right) = 1 - \mathbb{P} \left(j \in \bigcup_{s=1}^{r-1} \mathcal{N}_i^{(s)} \right) = 1 - \mathbb{P} \left(\bigcup_{s=1}^{r-1} \{j \in \mathcal{N}_i^{(s)}\} \right) \quad (\text{A.10})$$

$$= 1 - \sum_{s=1}^{r-1} \mathbb{P} \left(j \in \mathcal{N}_i^{(s)} \right) = 1 - \sum_{s=1}^{r-1} \pi_s, \quad (\text{A.11})$$

which follows from the fact that these events are disjoint, since j can only be in a single neighbour set at a time. Conditioning on the size of the previous stage neighbour set, we have

$$\mathbb{P} \left(j \in \mathcal{N}_i^{(r)} \middle| j \notin \bigcup_{s=1}^{r-1} \mathcal{N}_i^{(s)} \right) = \sum_{n=0}^{N-r} \mathbb{P} \left(j \in \mathcal{N}_i^{(r)} \middle| |\mathcal{N}_i^{(r-1)}| = n, j \notin \bigcup_{s=1}^{r-1} \mathcal{N}_i^{(s)} \right) \mathbb{P} \left(|\mathcal{N}_i^{(r-1)}| = n \middle| j \notin \bigcup_{s=1}^{r-1} \mathcal{N}_i^{(s)} \right) \quad (\text{A.12})$$

$$= \sum_{n=0}^{N-r} \left\{ 1 - \mathbb{P} \left(j \notin \mathcal{N}_i^{(r)} \middle| |\mathcal{N}_i^{(r-1)}| = n, j \notin \bigcup_{s=1}^{r-1} \mathcal{N}_i^{(s)} \right) \right\} \mathbb{P} \left(|\mathcal{N}_i^{(r-1)}| = n \middle| j \notin \bigcup_{s=1}^{r-1} \mathcal{N}_i^{(s)} \right) \quad (\text{A.13})$$

$$= \sum_{n=0}^{N-r} \{1 - (1 - \pi)^n\} \mathbb{P} \left(|\mathcal{N}_i^{(r-1)}| = n \middle| j \notin \bigcup_{s=1}^{r-1} \mathcal{N}_i^{(s)} \right) \quad (\text{A.14})$$

$$= 1 - \sum_{n=0}^{N-r} (1 - \pi)^n \mathbb{P} \left(|\mathcal{N}_i^{(r-1)}| = n \middle| j \notin \bigcup_{s=1}^{r-1} \mathcal{N}_i^{(s)} \right), \quad (\text{A.15})$$

since

$$\sum_{n=0}^{N-r} \mathbb{P} \left(|\mathcal{N}_i^{(r-1)}| = n \middle| j \notin \bigcup_{s=1}^{r-1} \mathcal{N}_i^{(s)} \right) = 1. \quad (\text{A.16})$$

Moreover, using Proposition 1,

$$\mathbb{P} \left(|\mathcal{N}_i^{(r-1)}| = n \middle| j \notin \bigcup_{s=1}^{r-1} \mathcal{N}_i^{(s)} \right) = \sum_{\mathbf{n} \in \Lambda_{r-2}(N-1)} \prod_{s=1}^{r-1} \mathbb{P} (Z_{n_1:n_{s-1}}(N-1) = n), \quad (\text{A.17})$$

since conditioning on the event $j \notin \bigcup_{s=1}^{r-1} \mathcal{N}_i^{(s)}$ is equivalent to constructing the neighbour sets from stage 1 through stage $r-1$ with one fewer node available to choose from. Hence,

$$\pi_r = \left(1 - \sum_{s=1}^{r-1} \pi_s \right) \left\{ 1 - \sum_{n=0}^{N-r} (1 - \pi)^n \sum_{\mathbf{n} \in \Lambda_{r-2}(N-1)} \prod_{s=1}^{r-1} \mathbb{P} (Z_{n_1:n_{s-1}}(N-1) = n) \right\}, \quad (\text{A.18})$$

as required. $\square$

References

- Almosova, A., Andresen, N., 2023. Nonlinear inflation forecasting with recurrent neural networks. *Journal of Forecasting* 42, 240–259. URL: <https://doi.org/10.1002/for.2901>.
- Ang, A., Bekaert, G., Wei, M., 2007. Do macro variables, asset markets, or surveys forecast inflation better? *Journal of Monetary Economics* 54, 1163–1212. URL: <https://doi.org/10.1016/j.jmoneco.2006.04.006>.
- Aparicio, D., Bertolotto, M.I., 2020. Forecasting inflation with online prices. *International Journal of Forecasting* 36, 232–247. URL: <https://doi.org/10.1016/j.ijforecast.2019.04.018>.
- Atkeson, A., Ohanian, L.E., 2001. Are Phillips curves useful for forecasting inflation? *Federal Reserve bank of Minneapolis Quarterly Review* 25, 2–11. URL: <https://doi.org/10.21034/qv.2511>.
- Barkan, O., Benchimol, J., Caspi, I., Cohen, E., Hammer, A., Koenigstein, N., 2023. Forecasting CPI inflation components with hierarchical recurrent neural networks. *International Journal of Forecasting* 39, 1145–1162. URL: <https://doi.org/10.1016/j.ijforecast.2022.04.009>.
- Batini, N., Haldane, A., 1999. Monetary policy rules and inflation forecasts. Technical Report 1. Bank of England. URL: <https://www.bankofengland.co.uk/quarterly-bulletin/1999/q1/monetary-policy-rules-and-inflation-forecast>.
- Bernanke, B.S., 2024. Forecasting for monetary policy making and communication at the Bank of England: a review. Technical Report. Bank of England. URL: <https://www.bankofengland.co.uk/independent-evaluation-office/forecasting-for-monetary-policy-making-and-communication-at-the-bank-of-england-a-review/forecasting-for-monetary-policy-making-and-communication-at-the-bank-of-england-a-review>.
- Bernanke, B.S., Woodford, M., 1997. Inflation forecasts and monetary policy. Technical Report 6157. National Bureau of Economic Research. Cambridge, Massachusetts, USA. URL: <https://doi.org/10.2307/2953656>.
- Binder, C.C., Sekkel, R., 2024. Central bank forecasting: A survey. *Journal of Economic Surveys* 38, 342–364. URL: <https://doi.org/10.1111/joes.12554>.
- Boaretto, G., Medeiros, M.C., 2023. Forecasting inflation using disaggregates and machine learning. arXiv preprint URL: <https://doi.org/10.48550/arXiv.2308.11173>.
- Boero, G., Smith, J., Wallis, K.F., 2008. Uncertainty and disagreement in economic prediction: The Bank of England Survey of External Forecasters. *The Economic Journal* 118, 1107–1127. URL: <https://doi.org/10.1111/j.1468-0297.2008.02162.x>.
- Brockwell, P.J., Davis, R.A., 1991. *Time Series: Theory and Methods*. Springer, New York.
- Bullmore, E., Sporns, O., 2009. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nature Reviews Neuroscience* 10, 186–198. URL: <https://doi.org/10.1038/nrn2575>.
- Burgess, S., Fernandez-Corugedo, E., Groth, C., Harrison, R., Monti, F., Theodoridis, K., Waldron, M., 2013. The Bank of England’s forecasting platform: COMPASS, MAPS, EASE and the suite of models. Technical Report. Bank of England Working Paper No. 471. URL: <https://dx.doi.org/10.2139/ssrn.2266506>.
- Chu, B., Huynh, K., Jacho-Chávez, D., Kryvtsov, O., 2018. On the evolution of the United Kingdom price distributions. *The Annals of Applied Statistics* 12, 2618 – 2646. URL: <https://doi.org/10.1214/18-AOAS1172>.
- Clements, M.P., 2024. Do professional forecasters believe in the phillips curve? *International Journal of Forecasting* 40, 1238–1254. URL: <https://doi.org/10.1016/j.ijforecast.2023.11.004>.
- Clements, M.P., Rich, R.W., Tracy, J.S., 2023. Chapter 3 - surveys of professionals, in: Bachmann, R., Topa, G., van der Klaauw, W. (Eds.), *Handbook of Economic Expectations*. Academic Press, pp. 71–106. URL: <https://doi.org/10.1016/B978-0-12-822927-9.00009-4>.
- Consumer Price Inflation Team, 2024. Consumer price inflation time series. URL: <https://www.ons.gov.uk/economy/inflationandpriceindices/datasets/consumerpriceindices>.
- Consumer Prices Team, 2024. Consumer Prices Indices Technical Manual, 2019. Technical Report. Office for National Statistics (ONS). URL: <https://www.ons.gov.uk/economy/inflationandpriceindices/methodologies/consumerpricesindicatestechnicalmanual2019>.

- Domit, S., Monti, F., Sokol, A., 2019. Forecasting the UK economy with a medium-scale Bayesian VAR. *International Journal of Forecasting* 35, 1669–1678. URL: <https://doi.org/10.1016/j.ijforecast.2018.11.004>.
- Eugster, P., Uhl, M.W., 2024. Forecasting inflation using sentiment. *Economics Letters* 236, 111575. URL: <https://doi.org/10.1016/j.econlet.2024.111575>.
- Faust, J., Wright, J.H., 2013. Forecasting inflation, in: Elliott, G., Timmermann, A. (Eds.), *Handbook of Economic Forecasting*. Elsevier. volume 2, pp. 2–56. doi:<https://doi.org/10.1016/B978-0-444-53683-9.00001-3>.
- Gilbert, E.N., 1959. Random graphs. *The Annals of Mathematical Statistics* 30, 1141–1144. URL: <https://www.jstor.org/stable/2237458>.
- Heckerman, D., 2008. A tutorial on learning with Bayesian networks, in: Holmes, D.E., Jain, L.C. (Eds.), *Innovations in Bayesian Networks: Theory and Applications*. Springer Berlin Heidelberg, pp. 33–82. URL: https://doi.org/10.1007/978-94-011-5014-9_11.
- Hendry, D.F., Hubrich, K., 2011. Combining disaggregate forecasts or combining disaggregate information to forecast an aggregate. *Journal of Business & Economic Statistics* 29, 216–227. URL: <https://www.jstor.org/stable/25800794>.
- Hubrich, K., 2005. Forecasting euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy? *International Journal of Forecasting* 21, 119–136. URL: <https://doi.org/10.1016/j.ijforecast.2004.04.005>.
- Joseph, A., Potjagailo, G., Chakraborty, C., Kapetanios, G., 2024. Forecasting UK inflation bottom up. *International Journal of Forecasting* 40, 1521–1538. URL: <https://doi.org/10.1016/j.ijforecast.2024.01.001>.
- Kang, X., Ganguly, A., Kolaczyk, E.D., 2021. Dynamic networks with multi-scale temporal structure. *Sankhya A* 84, 218–260. URL: <https://doi.org/10.1007/s13171-021-00256-1>.
- Knight, M., Leeming, K., Nason, G.P., Nunes, M., 2020. Generalized Network Autoregressive Processes and the GNAR package. *Journal of Statistical Software* 96, 1–36. URL: <https://doi.org/10.18637/jss.v096.i05>.
- Knight, M.I., Nunes, M.A., Nason, G.P., 2016. Modelling, detrending and decorrelation of network time series. arXiv preprint URL: <https://doi.org/10.48550/arXiv.1603.03221>.
- Kolaczyk, E.D., 2009. *Statistical Analysis of Network Data: Methods and Models*. Springer.
- Kolaczyk, E.D., Csárdi, G., 2020. *Statistical Analysis of Network Data with R (Use R!)*. Springer.
- Koski, T., Noble, J., 2009. *Bayesian Networks: An Introduction*. Wiley.
- Leeming, K., 2019. *New Methods in Time Series Analysis: Univariate Testing and Network Autoregression Modelling*. Ph.D. thesis. University of Bristol. URL: <https://hdl.handle.net/1983/a07ed087-ec7a-4929-8423-2e1b1726e4fe>.
- Lorenzo-Valdés, A., 2024. Network analysis of the Mexican stock market. *Investigación Económica* 83, 55–78. URL: <https://www.jstor.org/stable/48766939>.
- Lucchese, L., Pakkanen, M.S., Veraart, A.E.D., 2023. Estimation and inference for multivariate continuous-time autoregressive processes. arXiv preprint URL: <https://doi.org/10.48550/arXiv.2307.13020>.
- Makridakis, S., Spiliotis, E., Assimakopoulos, V., 2018. Statistical and machine learning forecasting methods: Concerns and ways forward. *PLOS ONE* 13. URL: <https://doi.org/10.1371/journal.pone.0194889>.
- Marcellino, M., Stock, J.H., Watson, M.W., 2006. A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics* 135, 499–526. URL: <https://doi.org/10.1016/j.jeconom.2005.07.020>.
- Medeiros, M.C., Vasconcelos, G.F.R., Veiga, Á., Zilberman, E., 2021. Forecasting inflation in a data-rich environment: the benefits of machine learning methods. *Journal of Business & Economic Statistics* 39, 98–119. URL: <https://doi.org/10.1080/07350015.2019.1637745>.
- Mohammad, S., Khochiani, R., 2022. Comparison of classical and dynamic network models efficiency in the application of generalized network autoregressive models. *Journal of Economic Research* 21, 171–194. URL: <https://doi.org/10.22054/joer.2022.64422.1018>.
- Monetary Policy Committee, 2024. *Monetary Policy Report*. Technical Report. Bank of England. URL:

- <https://www.bankofengland.co.uk/monetary-policy-report/2024/august-2024>.
- Nason, G.P., Salnikov, D., Cortina-Borja, M., 2023. New tools for network time series with an application to COVID-19 hospitalisations. arXiv preprint URL: <https://doi.org/10.48550/arXiv.2312.00530>.
- Nason, G.P., Salnikov, D., Cortina-Borja, M., 2024. Modelling clusters in network time series with an application to presidential elections in the USA. arXiv preprint URL: <https://doi.org/10.48550/arXiv.2401.09381>.
- Nason, G.P., Wei, J.L., 2022. Quantifying the economic response to COVID-19 mitigations and death rates via forecasting purchasing managers' indices using generalised network autoregressive models with exogenous variables (with discussion). *Journal of the Royal Statistical Society Series A: Statistics in Society* 185, 1778–1792. URL: <https://doi.org/10.1111/rssa.12875>.
- Peach, R., Arnaudon, A., Barahona, M., 2019. Semi-supervised classification on graphs using explicit diffusion dynamics. arXiv preprint URL: <https://doi.org/10.48550/arXiv.1909.11117>.
- Peach, R.L., Arnaudon, A., Schmidt, J.A., Palasciano, H.A., Bernier, N.R., Jelfs, K.E., Yaliraki, S.N., Barahona, M., 2021. HCGA: Highly comparative graph analysis for network phenotyping. *Patterns* 2. URL: <https://doi.org/10.1016/j.patter.2021.100227>.
- Pearl, J., 2009a. Causal inference in statistics: An overview. *Statistics Surveys* 3, 96–146. URL: <https://doi.org/10.1214/09-SS057>.
- Pearl, J., 2009b. *Causality*. Cambridge University Press.
- Ralph, J., O'Neill, R., Winton, J., 2015. *A Practical Introduction to Index Numbers*. Wiley.
- Scarselli, F., Gori, M., Tsoi, A.C., Hagenbuchner, M., Monfardini, G., 2008. The graph neural network model. *IEEE Transactions on Neural Networks* 20, 61–80. URL: <https://doi.org/10.1109/TNN.2008.2005605>.
- Stock, J.H., Watson, M.W., 1999. Forecasting inflation. *Journal of Monetary Economics* 44, 293–335. URL: [https://doi.org/10.1016/S0304-3932\(99\)00027-6](https://doi.org/10.1016/S0304-3932(99)00027-6).
- Stock, J.H., Watson, M.W., 2003. Forecasting output and inflation: The role of asset prices. *Journal of Economic Literature* 41, 788–829. URL: <https://www.jstor.org/stable/3217532>.
- Stock, J.H., Watson, M.W., 2007. Why has US inflation become harder to forecast? *Journal of Money, Credit and Banking* 39, 3–33. URL: <https://doi.org/10.1111/j.1538-4616.2007.00014.x>.
- Stock, J.H., Watson, M.W., 2009. Phillips curve inflation forecasts. Technical Report 14322. National Bureau of Economic Research. Cambridge, Massachusetts, USA. URL: <https://www.nber.org/papers/w14322>.
- Stock, J.H., Watson, M.W., 2020. Slack and cyclically sensitive inflation. *Journal of Money, Credit and Banking* 52, 393–428. URL: <https://doi.org/10.1111/jmcb.12757>.
- Tucker, J., Gooding, P., 2017. Consumer price indices, a brief guide: 2017. Technical Report. Office for National Statistics (ONS). URL: <https://www.ons.gov.uk/economy/inflationandpriceindices/articles/consumerpriceindicesabriefguide/2017>.
- Wu, Z., Pan, S., Chen, F., Long, G., Zhang, C., Philip, S.Y., 2019. A comprehensive survey on graph neural networks. arXiv preprint URL: <https://doi.org/10.48550/arXiv.1901.00596>.
- Yao, L., Chu, Z., Li, S., Li, Y., Gao, J., Zhang, A., 2021. A survey on causal inference. *ACM Transactions on Knowledge Discovery from Data* 15, 1–46. URL: <https://doi.org/10.1145/3444944>.
- Zhou, J., Cui, G., Hu, S., Zhang, Z., Yang, C., Liu, Z., Wang, L., Li, C., Sun, M., 2020a. Graph neural networks: A review of methods and applications. *AI Open* 1, 57–81. URL: <https://doi.org/10.1016/j.aiopen.2021.01.001>.
- Zhou, Z., Chen, X., Zhang, Y., Hu, D., Qiao, L., Yu, R., Yap, P., Pan, G., Zhang, H., Shen, D., 2020b. A toolbox for brain network construction and classification (BrainNetClass). *Human Brain Mapping* 41, 2808–2826. URL: <https://doi.org/10.1002/hbm.24979>.
- Zhu, X., Pan, R., Li, G., Liu, Y., Wang, H., 2017. Network vector autoregression. *The Annals of Statistics* 45, 1096 – 1123. URL: <https://doi.org/10.1214/16-AOS1476>.