Economics Observatory.

Economic and Social Research Council

Macroeconomic modelling: A review.

Andrea Correa-Jimenez, Richard Davies, Rahat Siddique

Economics Observatory and LSE Growth Co-Lab

June 2024

www.economicsobservatory.com

Executive Summary

1. In December 2023 the ESRC commissioned the Economics Observatory to provide a background review of UK macroeconomic modelling in order to gather ideas and feedback that could inform future investments in this area. This paper presents the results of that review. We examine the current models used by the principal UK forecasting institutions, the latest academic and non-academic approaches to modelling, and the most recent developments in data, coding, and computing infrastructure. We explore opportunities to improve macroeconomic modelling and suggest some areas where the UK can make strategic investments to push the frontier of macroeconomic research.

2. The first section focuses on the modelling approaches adopted by six institutions that regularly provide policy-relevant forecasts of the UK economy: the Bank of England, the Office for Budget Responsibility (OBR), HM Treasury, the National Institute of Economic and Social Research (NIESR), the Organisation for Economic Co-operation and Development (OECD) and the International Monetary Fund (IMF). We set out the nature of the models, what they include (and what they do not) and review work that has examined the accuracy of their forecasts.

3. There is a wide range of approaches to modelling—the second section provides a taxonomy of these approaches. We review the type of models typically used, such as VARs, DSGEs and large-scale models, as well as new academic developments, such as HANK models and ABMs. We also explore examples of cutting-edge model developments outside economics, such as AI and machine learning models, and weather forecasting.

4. Models evolve, with adaptation and improvement coming in reaction to economic changes, shifts in academic thinking, and as new forms of data (and computer chip) become available. In the third section we look to the future of macroeconomic forecasting, highlighting three areas for future development: centralised databases; a repository of code and a library of models; and computer power to solve the models.

5. The review was based on a round of structured interviews[1] with modellers at the institutions named above (and more widely), desk-based research, and personal experience in building and using forecasts in a policy setting. An important caveat applies: this is a vast area and one in which model choice often rests on deep, almost philosophical, views of how the underlying economic structure operates. This means that views often differ between modellers. That said, a number of findings, common across the people we spoke to and the literature reviewed, are apparent.

[1] We held conversations with economist from the Bank of England, HM Treasury, NIESR, the IMF and the OECD.

Key findings

6. **Comparability.** While many institutions forecast the UK economy, their forecasts are not comparable. The data used, the modelling framework adopted, the frequency and forecast horizon of the modelling teams all vary. Some of the differences between forecasts are driven by deep assumptions, as forecasts are for different purposes, like fiscal sustainability and price stability, which can drive differences in approach and lead to different results. However, forecasting can be a 'black-box' to those not directly operating the toolkits. This makes it hard even for experts to tease out the differences.

7. **Infrastructure.** Over time, the recognition that models and 'infrastructure' go hand in hand has grown. Model infrastructure includes (at least): the data inputs, a code base that defines the model, computing power that solves the model, and storage. Some institutions also have a user interface to make using the model easier.

8. **Common gaps, common improvements.** The families of models that have developed over the past 60 years are diverse including the role of 'structure' (that is, economic theory) the size of the model and solution techniques. Despite this diversity, common failings have become clear in the past 15 years. Two stand out: the failure to incorporate the financial system adequately, and, in more general terms the fact that consumer- or firm-specific traits are assumed away. Newer models fill these gaps, taking finance and heterogeneity more seriously.

9. **Lessons from outside economics.** There appear to be some simple lessons that could be learned from outside economics. For example, in the field of weather prediction, the idea of an 'ensemble' is central. This refers to running many competing models (but importantly basing them all on the same data) and comparing the results: the variance of the model predictions is a way to capture uncertainty. While some macroeconomic approaches use this idea, many do not.

10. **Computing power.** Outside economics, the power of computers used has increased substantially. In weather forecasting, for example, the UK's supercomputer contains 460,000 cores. Computer power can be used to enhance the use of new economic modelling techniques and take advantage of technological advances that can be incorporated into macroeconomic modelling and forecasting.

11. **The modelling community.** The specialists that build models are embedded within policy institutions that have idiosyncratic remits. Furthermore, the forecasting process involves judgement, often from a panel or committee of senior economists. These judgments reflect individual focus, remit, interests and understanding of the economy, and can influence the final outcome of the model. There is widespread interest in building a stronger community of modellers, to exchange experience—ultimately to improve human capital—among those building the models.

Recommendations

12. **Maintain many models.** Models old and new have strengths and drawbacks. Policy contacts were clear that a range of different model are needed, so that teams can adapt to answer changing policy questions. This suggests that a diverse set of policy-focused models should be maintained, for the UK. The resource costs of this can be reduced by maintenance of open-source codebases.

13. **Data: APIs for macro (and micro).** Models are only as good as the data that is fed into them. The UK lacks a common macro API that can ensure that models are running on comparable data. APIs save time and add transparency – a central macro API could be built cheaply and quickly. In addition, some vital recent improvements—the role of finance and heterogeneity (both firm and household)—rely on statistics computed from microdata. This is often held in secure-access databases. Over time a microdata API (that fed out statistics, *without* the user seeing the raw data) would be hugely beneficial. More detail on this point is included in our report, and also in Davies (2024).

14. **Require comparability, not just open-sourcing**. While there are laudable norms—making model equations and some data open-source, for example—approaches remain finely tailored. At the same time as valuing this diversity there could be a parallel norm or requirement to aid comparability. One idea would be to run a "UK standard" forecast, perhaps annually, on the same data (so that models are based on the same vintage of data), with the same horizons (say, 2- and 5-years). This would be akin to lining all models on the same starting line, and running a race. It would aid comparison and evaluation.

15. **Enhance access—user interfaces**. Inside policy institutions the number of people that actively build and maintain models is small. It is generally not possible for policymakers themselves, academics, or specialist journalists to engage with the models in a meaningful way. This collapses the number of people that truly understand them. Simpler user interfaces would help here.

16. **Shared resources: code and compute**. There is replication of effort among UK macro modellers. While this can be beneficial (fresh eyes on a problem) in other cases it will be inefficient. A centralised codebase (the collation and presentation of macro models, on a singular website) would help. While some modellers we spoke to thought computing resources were adequate, others felt they were lacking. Shared access to cutting-edge machines and cloud-based resources would make sense.

17. **Community support**. Finally, investing in the macroeconomic modelling and research *community*—that is, in modellers' interactions and inter-relationships, rather than simply their number—would help. The people we spoke to often felt isolated and not fully abreast of developments in other institutions. A tighter-knit community would share lessons, code, resources. They could also agree on setting new common challenges to reignite interest in the field. Expanding links with international practitioners can also help to enrich research on policy-driven modelling.

Contents

Section 1: The state of play

18. This section reviews the approach taken at six institutions that produce a regular UK macroeconomic forecast: the Bank of England, the OBR, HM Treasury, NIESR, the OECD and the IMF. These institutions differ in the type of model that is used, as well as the frequency of forecasts produced. The purposes of the forecasts also vary: for example, the OBR is required by law to review and scrutinise the UK's public finances, while the IMF models the UK economy as part of their wider global forecast. Box 1 reviews the ESRC Warwick Macroeconomic Modelling Bureau, an initiative that was established to improve the accessibility of UK macroeconomic models and which was supported by several UK institutions in the 1980s.

Modelling the UK economy

The Bank of England

19. The Bank of England's current forecasting process has been in place since the establishment of the Monetary Policy Committee (MPC)[2] in 1997. The Inflation Report, which contains the Bank's forecast, was published four times a year between 1993 and 2019 following formal approval and sign off by the MPC, for the probability distributions and fan charts (BoE, 2015). Since November 2019, the Bank has produced the Monetary Policy Report, which presents the economic analysis and inflation projections used by the MPC to inform its decisions on interest rates.

20. The Bank's Monetary Analysis Directorate (MA) produces the forecast and analysis for the MPC. The MA consist of three teams: i) modelling, which oversees the development and maintenance of the models; ii) forecast, which uses the results of the models to produce the forecast; and iii) strategy, which develops optimal policy strategies. The Bank's forecast is a communication and policy device, and it is owned by the MPC.

21. The modelling approach at the Bank has evolved over time. Between 2003 and 2011, the Bank used BEQM (pronounced 'Beckham'), which was the main tool for the construction of the projections contained in the Inflation Report. Since November 2011, a system of models known as COMPASS (the Central Organising Model for Projection Analysis and Scenario Simulation) has been the primary forecast platform.

22. The main model is a New Keynesian, open-economy dynamic stochastic general equilibrium (DSGE) model, which assumes that prices and wages are sticky (Burgess et al, 2013). This is a smaller model than those previously used at the Bank, providing forecasts for 15 macroeconomic variables including GDP, inflation, interest rates and trade, over short to medium horizons. The model is estimated on UK data using Bayesian

[2] The MPC is comprised of nine individual members, who are a mixture of internal and external Bank representatives.

methods. COMPASS comprises five economic agents: firms, households, the government, the external sector (the rest of the world) and the monetary policymaker. The model comprises a system of behavioural equations derived from the optimisation problems of the agents, which means that expectations play a key role in agents' decision-making.

23. The platform also contains a suite of around 50 separate models, covering a huge range of different frameworks and ways of thinking about the economy. This supplements the projections provided by COMPASS and the models are organised into three categories: models that articulate omitted shocks and channels from COMPASS; models that produce forecasts for additional variables; and models that generate alternative forecasts to the variables included in COMPASS.

24. Alongside COMPASS and the suite of models, the Bank also has an IT infrastructure to support the different models and optimising the amount of time taken to produce the forecast (Burgess et al, 2013). The IT infrastructure has two components:

- The Economic Analysis and Simulation Environment (EASE): the user interface that Bank staff use to build and analyse the forecast.
- The Model Analysis and Projection System (MAPS): the MATLAB modelling toolbox, built and maintained by Bank economists. It is similar to other MATLAB toolboxes[3], and it is designed to support the forecasting process.

25. The Bank also publishes an open research agenda to facilitate collaboration with external researchers. In this agenda, the Bank features research topics related to the monetary toolkit, unconventional instruments and new strategy questions facing monetary policymaking. These include thinking about heterogeneity and new modelling approaches to understand how heterogeneity among banks, firms and households can affect the transmission mechanism of monetary policy. Additionally, research topics around the prudential framework highlight the importance of including macroprudential models, improving the modelling framework more broadly to estimate costs and benefits of policy and to conduct optimal policy analysis, as well as developing stress testing models to understand linkages between the financial system and the real economy, as well as situations where the financial system is under stress (Bank of England, 2024).

26. Furthermore, in May 2023 the Bank commissioned a review into the Bank's forecasting and related processes during times of significant uncertainty, with the purpose of develop and strengthen the processes in support of the MPC forward-looking approach to the formulation of monetary policy (Bank of England, 2024). This review was commissioned to Dr Bernanke and provided an assessment of current forecasting methods and the relationship between forecasts, monetary policy decisions and their

[3] Like DYNARE, IRIS, and YADA

communication, setting out 12 recommendations that provided a clear direction for the Bank's forecasting methods and the role of forecasts and broader analysis in monetary policy discussion, formulation, and presentation. The recommendations of the review are organised in three major themes (Bank of England, 2024):

- **Building and maintaining a high-quality infrastructure for forecasting and analysis:** improve and maintain the forecasting infrastructure, data management, software, and economic models.

- **Provide a forecast process that better supports the MPC decision making**: this includes equipping the MPC and Bank's staff to learn from previous forecast errors, identify and quantify risks to the outlook, and deal with uncertainty and structural change in the economy.

- **Using the forecast to communicate the MPC's outlook and policy rationale to the public:** help the MPC communicate to the public its view of the economy, the risks and uncertainties around its outlook, and its policy rationale.

The Office for Budget Responsibility

27. Set up in 2010, the OBR is the UK's official independent fiscal watchdog. Its forecasts are used for input to government decisions, as the Treasury does not produce its own forecast (Institute for Government, 2022). The model used to produce these forecasts is developed and maintained by the OBR and the Treasury jointly, although the OBR have complete freedom over the version of the model that is used. However, the OBR can only publish the forecast once asked to do so by the Chancellor.

28. The large-scale macroeconometric model is based on the original model developed by the Treasury in 1970, and it is used to provide a five-year forecast for the key components of the public finances twice a year. The forecast and model are managed by the full-time staff of the OBR, comprising 45 civil servants (OBR, 2023). The model is a simplified representation of the UK economy as described in the National Accounts published by the ONS (OBR, 2013). There are 627 variables in the model, organised by three broad groups to describe the economic relationships in the model: 213 accounting identities; 34 behavioural equations; and 378 technical relationships and exogenous variables (OBR, 2021).

29. The structure of the model is divided into seven groups, which comprise the different sectors of the model:

- Expenditure components of GDP: consumption, inventories, investment, exports, imports, public sector expenditure, gross domestic product.

- Labour market: employment, unemployment, participation rate.

- Prices, costs, and earnings: earnings, GDP(E) deflators, inflation indices.
- Balance sheets and income accounts: household balance sheet, external balance sheet, corporate balance sheet, the income account, balance of payments.
- Public sector: receipts, public sector totals.
- Domestic financial sector: interest rates, monetary aggregates.
- North Sea: relative price of oil, US dollar exchange rate.

30. As the objective of the model is to forecast the outlook for the public finances, the OBR requires a more detailed forecast of the expenditure and income measures of nominal GDP as well as detailed treatment of public sector variables. For the same reason, the use of a large-scale macroeconometric model is ideal, as it allows the use of large forecast variables and flexibility of judgement. On the latter, it is important to note that as a simplification of reality, some equations will not explain the recent past well, particularly considering recent structural changes in the economy. For this reason, the forecasters need to interpret how the unexplained behaviour will evolve during the forecast period, by adding some judgement (such as different approaches to informing the forecast, historical evidence, comparable episodes, technical relationships, use of auxiliary models or external conditioning variables).

31. One gap in this model is that it fails, as do many macroeconomic models, to include a well-articulated financial sector. However, the banking sector in this model is implicitly reflected through the forecast of potential output and cyclical movements around it, as well as in aggregate demand, through the assessment of conventional monetary policy and the wedge between policy rates and lending rates.

<u>HM Treasury</u>

32. Publications of short-term forecast (two years ahead) from the Treasury started on a regular basis from 1968. From 1975, they were required to be published twice a year and from 1980, they also included medium-term forecasts, although less detailed than the short-term forecasts (Pike & Savage, 1998). The forecast was carried out by two main teams of around ten people each[4]: the Treasury Public Sector Finances Team and the Economics Prospect Team, responsible for maintaining the macroeconomic model.

33. The model used by the Treasury had 1175 variables, of which 880 were endogenous, 500 explained by behavioural equations and the rest were identities (Mellis, 1988). A description of the structure of the model from 1986 is as follows:

- Domestic non-financial economy (285 variables).

[4] As of 1998.

- Balance of payments current account (190 variables).

- Financial model and balance of payments capital account (310).

- Public sector (390).

34. The Treasury stopped its modelling and forecasting on June 2010, when the model was passed on to be maintained and developed by the OBR. The OBR was set up to be a central part of the fiscal framework, responsible for examining and reporting on the sustainability of the public finances and providing credible fiscal and economic forecasts (OBR, 2011). However, the HMT currently plays a supporting role to the OBR analysis of the macro impacts of policy via the indirect effects process. The HMT has a set of models around different economic topics and also use the OBR's forecast model and NIGEM to undertake their own analysis outside of fiscal events.

NIESR

35. The National Institute Global Econometric Model (NiGEM) is the macroeconomic model of the National Institute of Economic and Social Research (NIESR). This is a large-scale macroeconometric model of the world economy that has been used since 1987 by policymakers and the private sector to produce economic forecasts, build scenarios and conduct stress testing. The model comprises more than 7,500 variables with over 10,000 equations, using historical data from 1997 and producing economic forecasts up until 2060 (NIESR, 2023). NIESR produces a quarterly forecast with a central forecast baseline to 2039.

36. The model is New Keynesian in its approach, with forward-looking agents in certain markets and nominal rigidities. It is structured around the national income identity and has many of the characteristics of a DSGE model (see Section 2). However, it differs in that NiGEM is based on estimation using historical data (OECD, 2011), being estimated equation by equation. There are two types of countries in the model: full countries and reduced countries, with the latter having a smaller demand side due to fewer data availability. The model has complete supply and demand sides as well as monetary and financial sectors. The countries are linked by trade, financial markets and stock of assets, and aggregation at the world level is internally consistent. Most countries in the OECD are modelled separately as well as another few countries[5], while the rest of the world is modelled in six regions: Sub-Saharan Africa, Europe, the Commonwealth of independent States, the Middle East and North Africa, Asia and the Americas (Hurst et al, 2014). The model also includes monetary and fiscal policy rules as well as debt stocks and deficits, with the possibility of including a target for the debt-to-GDP ratio. There is also a version

[5] China, India, Russia, Brazil, Hong Kong, Taiwan, Singapore, South Africa, Estonia, Latvia, Lithuania, Slovenia, Romania and Bulgaria.

of NiGEM that includes climate change and several equations to analyse policy scenarios and shocks.

<u>The OECD</u>

37. The OECD forecasts are the result of forecasting rounds that last seven to eight weeks and are published twice a year in the Economic Outlook. The forecasts rely heavily on the judgement of country experts and further reviews and discussion tables, which include not only country experts but also representatives of the countries. For the model simulations, the OECD uses the NiGEM global model and short-term indicator models, which are mostly used to 'set the scene' at the beginning of the forecast round and to evaluate scenarios around the baseline forecast (Turner, 2016). In the case of the UK, as well as the euro area and the rest of the G7 economies, the assessment includes these short-term indicator models, which comprise a set of high-frequency indicators to provide estimates of near-term quarterly GDP growth (OECD, 2011). These models try to make use of all available monthly and quarterly data, combining information from soft indicators (such as business and consumer surveys) and hard indicators (such as industrial production, retail sales and house prices).

38. The estimated indicator models are more accurate and reduce the size of errors more than autoregressive time series models, but they are still limited in their ability to forecast quarterly GDP. Nevertheless, the estimates from these models are used to guide the Economic Outlook assessment exercise and other public analyses released by the OECD. The forecast process is assisted by a Forecast Entry system that centralises the forecast data, allows country experts to view the recent outcomes and maintains the consistency and accuracy by automatically re-evaluating the National Accounts and other identities every time a forecast component is updated. Typically, a country desk will be responsible for quarterly forecasts of about 60 macroeconomic variables, another 150 variables built-up by identities and around 60 exogenous variables set according to centrally agreed assumptions (Turner, 2016).

39. In the forecast exercise for individual countries, the key variables and relationships can be grouped in:

- Domestic expenditures: private consumption, saving rates, non-financial assets, financial variables, business sentiment, residential construction.
- Employment, wages and prices: employment and labour market trends, productivity trends, capacity constraints, costs, current wages settlement data, real wages, inflation, domestic prices, output gaps and foreign prices.
- Output gaps: the difference between actual and estimated potential GDP, considering the capital stock, changes in labour supply, factor productivities and NAIRU (the non-inflation accelerating rate of unemployment).

- Foreign trade and balance of payments: aggregated import volumes of goods and services, aggregate export volumes of goods and services, export prices, import prices.

The International Monetary Fund

40. Similar to the OECD, the IMF uses a DSGE model to build scenarios around the forecast, called the IMF's Global Economy Model (GEM). The first version of GEM, which consisted of a two-country version was developed in 2001. After that, the use of the model has grown, currently being used by the central banks of Canada, Finland, Japan, Italy, Spain and the UK. The model comprises three sectors: firms, which produce goods; households, which consume goods and provide labour and capital; and the government, which taxes and spends (Laxton, 2008). Consumption and production are characterised by standard constant elasticity of substitution (CES) utility and production functions. The final goods are split into traded and non-traded goods, and the model incorporates exchange rates pass-through.

41. Some advantages of GEM include the fact that it can provide evaluations of policies in a general equilibrium setting, as the model is built from microeconomic foundations. The model also facilitates evaluation of the costs and benefits of a policy, as it is derived from the maximisation of profit and utility. GEM also has a large size, including many economic distortions, which strengthens the economic policy analysis derived from the model. The simulations of GEM have been incorporated in the IMF's analysis in the World Economic Outlook (WEO), as well as other IMF exercises, although they are not used to provide the forecasts shown in WEO. The forecast exercise is not centralised, and it is generated using expertise available on countries by aggregating projections from individual countries (Bayoumi, 2004).

42. Table 1 summarises the main characteristics of the latest models used by all the official forecaster institutions mentioned in this section.

Table 1: Summary of UK official forecast characteristics						
	Bank of England	OBR	HM Treasury	NIESR	OECD	IMF
Year in which forecasting started with the latest model	2011	2010	1968 (until 2010)	1987	2010	2001
Length of forecast	Four years	Five years	Two years	16 years (up to 39 years)	Two years	Five years
Number of variables and equations	15 variables and equations (in the COMPASS model)	627 variables and equations	1,175 variables	7,500 + variables 10,000+ equations	Around 270 variables	
Sectors included (Or agents)	Five	Seven	Four	Five	Four	Three
Trade sector	Open economy	Open economy	Open economy	Open economy – Global model	Open economy – Global model	Open economy
Model type	Structural – DSGE	Large-scale model	Large-scale model	Large-scale model	NiGEM model + short-term indicator models	Structural – DSGE
Size of modelling team	Eight in the forecast team and seven in the modelling team (with three people seconded)	3% (45)	2% (20)	14 (six in modelling and another eight at the time of forecasting)	Four on the country desk, with further teams on trade, fiscal and monetary policy	
Frequency of publication	Four times a year	Twice a year	Twice a year	Four times a year	Annual (EO)	Twice a year (WEO)

Notes: The OECD and IMF forecasts include part of the global outlook; it is assumed the UK-specific forecasts are constructed using the same modelling approach. In these institutions, the forecast is made by the judgements of the country experts. The model is used to guide the forecast and construct alternative scenarios.

Evaluation of the current models

43. Model evaluation is an important element of the forecasting process. Model evaluation enables forecasters to assess the efficacy of their models and to understand the relationship between the model design and the data that are used. Although it would be potentially beneficial for users, very few institutions publish a document evaluating the performance of their forecast. In the UK, the Bank of England and the OBR evaluate their forecasts, at irregular frequencies. While the OBR is required by Parliament to publish an annual forecast review, the Bank is not. The remaining four institutions do not have a practice of publishing a forecast evaluation, demonstrating the lack of standardisation and learnings from different models' performance.

44. Some Bank of England studies have tried to assess the forecast performance of the models used by the Bank (Independent Evaluation Office, November 2015). An exercise conducted in 2015 evaluated the judgement-free performance of the COMPASS model, during and after the global financial crisis of 2007-09, against the MPC judgemental forecast and the forecast of the suite of statistical models. The authors found that at shorter horizons, the MPC's projections were more accurate for both GDP and inflation, compared with both the forecast of COMPASS and the suite of models. However, over longer horizons, COMPASS had the most accurate forecast for inflation and the suite of models had the most accurate forecast for GDP (Fawcett et al, July 2015).

45. In terms of evaluating forecast bias, a 2013 study presented an evaluation made for projections of GDP and inflation during 1997-2013, for one quarter and one year ahead projections (Hackworth et al, 2013). For GDP, the authors found that projections were unbiased for both time horizons at 5% significance level, although there was evidence of bias for one year ahead projections at 10% significance level (projections were too high). In the case of inflation, the results were the same, finding unbiased projections for both horizons at 5% significance level and bias projections for one year ahead projection at 10% significance level (projections were too low).

46. Similar results were found in an external 2012 study analysing the MPC forecast performance for GDP and inflation, comparing average absolute errors in the MPC estimates with average absolute errors by external forecasters (Stockton, 2012). The author found that the MPC forecast has been persistently over-predicting GDP and under-predicting inflation since the global financial crisis. The narrative of the Bank to explain the persistent errors in its forecasting is mostly a consequence of bad luck in economic conditions, such as tighter than expected credit conditions, weak productivity and downward pressure of real incomes. However, these consistent errors can also be related to some inertia in the forecasting process and a lack of systematic evaluation of the forecast to identify, with more precision, the source of the forecast errors (Stockton, 2012).

47. On the other hand, the OBR regularly evaluates[6] its forecast in the Forecast Evaluation Report (FER). A comprehensive look at the OBR forecast, since its establishment in 2010, found that while the OBR's forecast differences for real GDP growth and government borrowing were small relative to the outcome, the forecast tended to underestimate government borrowing, mainly due to underestimation of government spending, which also tended to overestimate real GDP growth (OBR, 2023). The forecast record was compared to external forecasters, the Bank of England, official forecasters in Europe and the forecasts made by the Treasury prior to the OBR. The results showed that the real GDP forecast was similar to external forecasters and more accurate than official forecasters in Europe. However, the borrowing forecast was less accurate than both of them beyond the first year of forecast. But forecasts of both real GDP growth and borrowing were more accurate and less biased than those previously produced by the Treasury. In fact, the outturn of the forecast can be broken down into different factors, such as changes in the recorded data at the ONS, government policy changes or unexpected economic developments.

48. There are also some studies that have tried to assess the accuracy of the IMF forecast. A 2021 study evaluated the forecast of real GDP growth during the 2004-17 period, covering the entire five-year forecast horizon. The evaluation was carried out with

Box 1: The Macroeconomic Modelling Bureau

The Macroeconomic Modelling Bureau was established by the ESRC at the University of Warwick in 1983. The idea behind the creation of the Bureau was to set up a new centre to undertake comparative research on existing models of the UK and help to achieve greater openness and understanding of the models, and their forecasts and policy analysis (Warwick, 2014). A Macroeconomic Modelling Consortium was created to coordinate support for the research programme in macroeconomic modelling and to manage this on a four-year cycle, managed by a Research Council, HM Treasury and the Bank of England,

The Bureau's primary objectives were to improve the accessibility of the UK's macroeconomic models, to promote a general understanding of these models and to undertake its own comparative and methodological research. To support these goals, versions of models and databases were regularly deposited at the Bureau. The groups supported by the programme included the LBS, and NIESR models, the Cambridge University Small UK Model, and the Liverpool research group in macroeconomics. The HM Treasury model was deposited as part of the Bureau's contract, and the Bank of England and Oxford Economic Forecasting models were voluntarily deposited. The Bureau also arranged an annual Macroeconomic Modelling Seminar to discuss its programme of work and general matters of concern to modellers, macroeconomists and model users, as well as a regular newsletter that covered developments in macroeconomic modelling and a series of annual review volumes. The Bureau closed in 1990, although the annual seminar continued for two more years with the support of the Bank of England.

[6] To measure accuracy, the OBR compares the absolute average difference between the central forecasts and outturn (a measure of how closely the forecasts predict future outcomes); to measure bias, it compares the simple average difference between the central forecasts and outturn (a measure of how systematic the differences are).

three components in mind: accuracy, bias and efficiency[7], and compared to the WEO forecast for 1990-2003, forecast outcomes from simple time series models, and forecasts published by Consensus Economics (Celasun et al, 2021). The results showed that for a forecast of up to two years, there is no visible bias, but after the two-year forecast horizon, there is an upward bias. This tendency to have an upward-biased forecast can be explained due to the relatively common growth collapses that arise from natural disasters, wars and/or financial crises. Nevertheless, this error has been declining, as the over-prediction was higher in forecasts produced between 1990 and 2003 than in those made during the 2004-17 period. Changes in forecasting performance might also be the result of shifts in underlying economic conditions. However, it is also possible that these changes are the result of new forecasting methods or new databases that allow for better monitoring and prediction of economic outcomes. Additionally, over- and under-predictions in the current year tend to persist over time, which suggests that forecast can also be improved by learning from past mistakes. It is important to note that sometimes institutions can knowingly condition their forecasts on things they know are wrong, for example the OBR and Bank of England, which might assume government policies or market expectations for interest rates.

[7] Accuracy refers to the overall magnitude of the forecast error; bias refers to the systematically over or under prediction of the outcomes; and efficiency refers to whether the forecast can be improved by a better used of the available information.

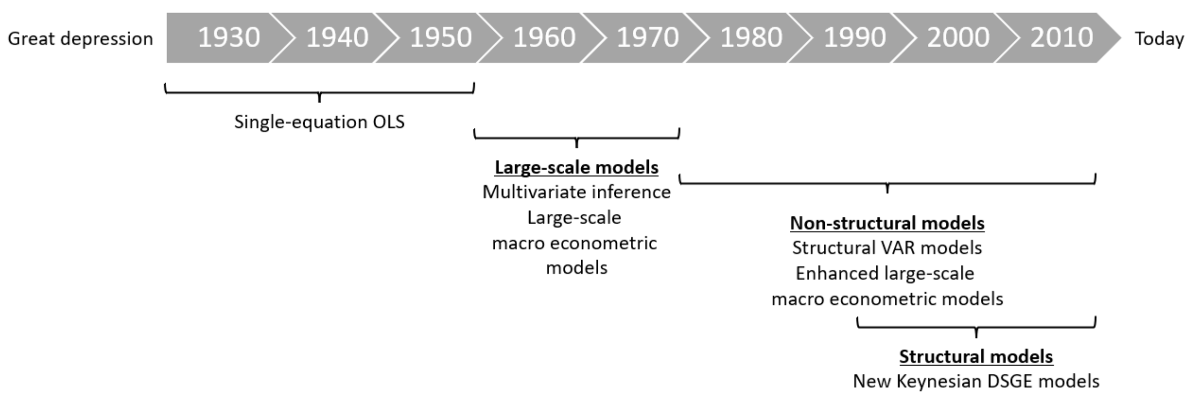
Section 2: The frontier – what do the best models do?

49. The type of models used to forecast the UK economy has varied over time. As technology and data have improved, modelling approaches have also become more complex. In academia, the availability of large datasets and greater computational power has created a new wave of models, in particular agent-based New Keynesian DSGE models. In industries such as finance and climate science, real-time data have accelerated the speed and accuracy at which forecasts can be produced. This section seeks to explore the latest developments across academia and industry, and their application to UK macroeconomic forecasting.

Background: types of models

50. Over the past 60 years, three categories of macroeconomic models have been developed: non-structural, structural and large-scale (Pescaroti & Zaman, 2011) (see Figure 1).

Figure 1: Timeline of macroeconomic models over time



Source: Authors elaboration based on (Dou et al, 2020)

Non-structural models

51. These models are data driven. They are statistical time series models that represent correlations with historical data. These models lack economic 'structure' – that is, there are no assumptions or equations specifying how a household behaves, for example. They are flexible and capture the influence of history in their forecasts, matching the economic data. This characteristic made these types of models more useful to produce unconditional forecasts, which can be accurate when the policy regime does not change frequently. These models are still widely used, the vector auto-regression models (VARs) being one of the most commonly implemented.

Structural models

52. These models differ by being based on fundamental principles of economic theory. They began in academia around the 1980s, as a way to develop an economic modelling framework that did not violate the Lucas critique[8]. This gave rise to a set of models known as dynamic stochastic general equilibrium models (DSGE), which have their origins in a combination of real business cycle (RBC) theory and New Keynesian macroeconomics (Dou et al, 2020). In these models, the variables are affected over a period of time; the models can incorporate random economic events and behave like a system (the 'general equilibrium' part of the name) in which everything in the model depends on everything else (see Figure 2).

53. These models differ from 'large-scale' models in the sense that they have microeconomic foundations, following a 'bottom-up' aggregation approach that goes from the micro to the macro. This makes these types of models better at constructing conditional forecasts and producing policy scenarios. Other advantages of these types of models are that they can incorporate the role of monetary policy and can make use of the solution methods of non-structural models, as their solutions rules are well approximated by non-linear rules. However, they are much smaller as they are very difficult to solve and analyse, leaving some variables out of the analysis.

54. These types of models are often organised featuring different economic agents, such as households, firms, the government, the monetary policy authority and the external sector, although in a representative agent setting. They can be summarised by three aggregate equations: the Phillips Curve, the Taylor Rule and the Fischer Equation (Violante, 2021). These aggregate equations specify, respectively, the link between inflation and employment/output dynamics, how the monetary authority operates the interest rate, and the link between the policy rate, the real interest rate and inflation expectations.

<u>Large-scale models</u>

55. These models are a hybrid of structural and non-structural models, often used to produce conditional forecasts. They use economic theory, but they are built from many equations that describe relationships from aggregate historical data. These models are composed of individual equations that establish connections between variables. This implies that the economic relationship can be chosen from an extensive array of data series, offering a comprehensive description of the economy, which often culminates in the creation of large models (Pescaroti & Zaman, 2011). However, their complexity can be a limitation for their understanding.

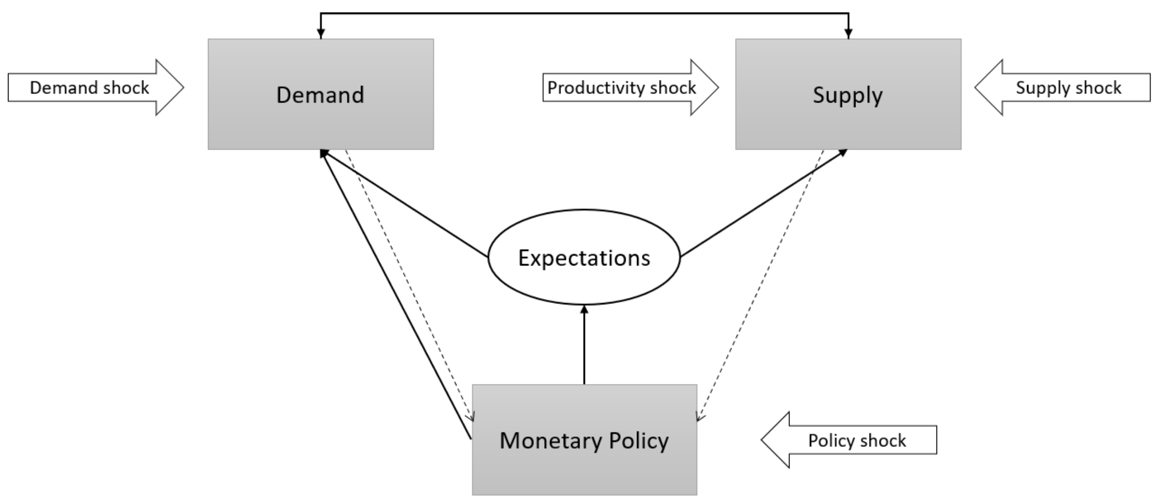
56. The development of these models began in the 1960s, with a major use for forecasting and policy analysis around the 1970s. However, the great economic changes

---

[8] The Lucas critique refers to the inability of the models to incorporate expectations and therefore their incapacity to compare alternative policy effects since their parameters are not structural and invariant.

and macroeconomic turbulence during the latter decade revealed a flaw in these models: they failed to incorporate the role of expectations, producing unreliable conditional forecasts and policy analysis. In particular, this was evident in times of economic turbulence, as the constant parameters underlying the model would change according to policymaking or expectations about policy changes. After this, further improvements of this type of models were necessary, consisting of the incorporation of forward-looking expectations and better representation of agents' decision-making.

Figure 2: Basic structure of a DSGE model



Source: FRBNY Economic Policy Review (Sborde et al, 2010)

Newer approaches: HANK models and ABM

57. Despite their differences, the older models reviewed above have some common drawbacks. These include:

- Financial system. The difficulty of accounting for important fluctuations resulting from systemic risk within the financial system, as the models lack an interbank market.

- Stressed periods. The impossibility of analysing the link between the endogenous risk premium and macroeconomic activity leads to unreliable results and an inconsistent framework when conducting experimental stress tests to understand stressed periods.

- Inequality. The difficulty of understanding inefficient allocations because of the lack of heterogeneity among agents.

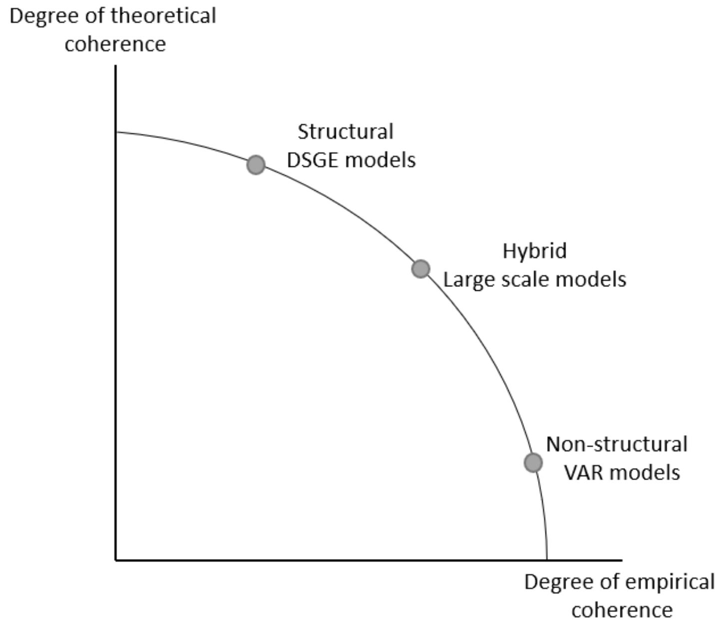
- Uncertainty. The failure to incorporate shocks, anticipate structural changes and capture uncertainty, or to address these possibilities in separate, different scenarios from the baseline (for example, using tools like fan charts).

58. Additionally, there is also a need for advances in non-linear solution methods, to address the non-linear dynamics of financial markets and the macroeconomy, as well as for better risk measurement and data collection infrastructure and availability (Dou et al, 2020).

59. Relatedly, there is a further challenge with regard to conventional and unconventional monetary policy. During recent economic crises, reaching the zero lower bound for interest rates led to central banks developing unconventional monetary policy measures such as credit easing and quantitative easing. Moreover, these measures can also have fiscal policy impacts, as they are often focused on some specific debtors and sectors, making the distributional effects of monetary policy stronger and putting at risk the government's balance sheet.

60. Regarding the economic forecast, the fact that there are no constants or invariant parameters in economics is a challenge, because the variables are always in continued interaction and their relationships do not have an easily predictable outcome (Pescaroti & Zaman, 2011). It is important to note that ultimately, the outcome can change because of the forecast itself, as forecasts can influence people's behaviour. This is why even though policymakers would like to have better forecasts, the ultimate objective should be to have better policies, for which the models can be very useful, helping policymakers in the process of decision-making. However, choosing the right economic model to guide the policymaking process can also be a challenge in itself (see Annex B). As we saw in the first part of this section, there is a trade-off between the theoretical and empirical coherence of the models, for which policymakers have to decide which type of model is more suitable for addressing their specific concerns (see Figure 3).

Figure 3: Trade-off between theoretical and empirical coherence in macroeconomic models



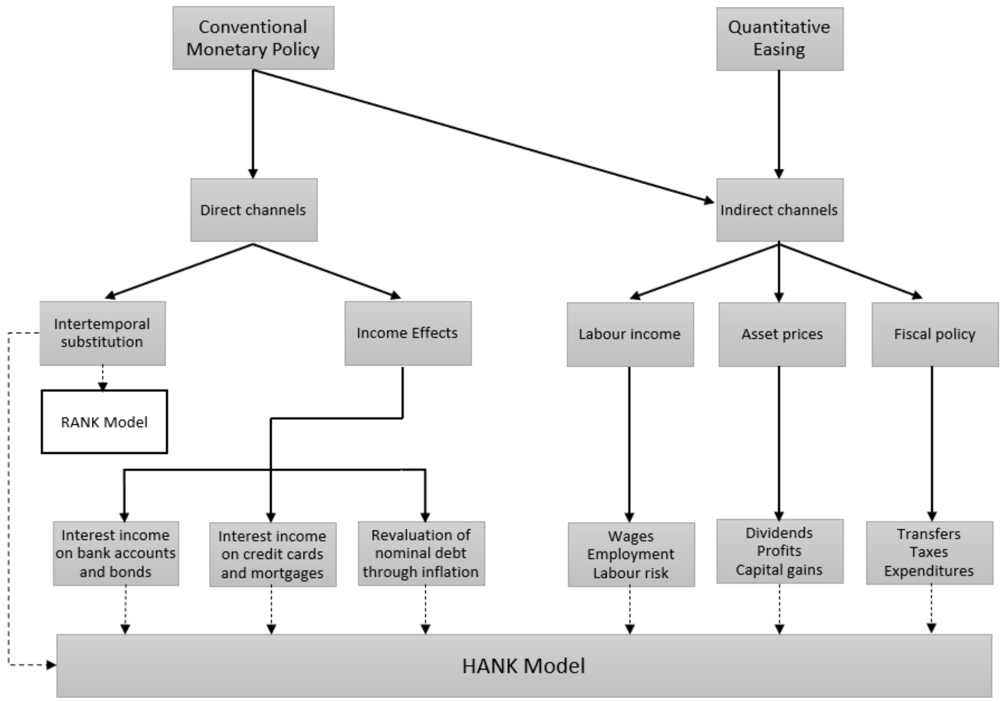
Source: Own elaboration based on the Report on modelling and forecasting at the Bank of England (Pagan, 2003)

61. To address some of these challenges, some academics have developed new economic models, such as heterogeneous agent New Keynesian (HANK) models. These models recognise the role of income and wealth inequalities, combining heterogeneous agent models with New Keynesian models. The HANK model can help with the prediction of policy impacts, as monetary policy affects households' consumption directly and indirectly. The size of these effects depends on the aggregate marginal propensity to consume, which HANK models address by being consistent with empirical evidence of consumption behaviour and including higher moments of cross-sections of wealth and income components[9] (Sargent, 2023). For this reason, they need much more information about the household side of the economy than the traditional representative agent models explored above. Additionally, the importance of the indirect effects in the HANK model means that the transmission of monetary policy comes through all the mechanisms that contribute to price formation of goods, credit, housing and financial markets (see Figure 4). This means that the central bank needs more comprehension of all those market structures and frictions (Kaplan et al, 2023) and for that, they are also required to have access to the right data, especially data on the financial system and microdata on households' balance sheets.

[9] This means that HANK models also require dynamic programming, dynamic programming squared, vector autoregressions and structural macroeconometrics to be modelled.

62. This type of model also addresses the divergent effects of monetary policy on different households, recognising that the impacts will depend on the balance sheets of the household (for example, an interest rate cut will benefit debtors but harm savers). These kinds of effects can also differ across countries. Moreover, by introducing income and wealth inequality, HANK models address the lack of linkages between fiscal and monetary policy, showing that monetary policy can have consequences for fiscal policy (for example, by raising interest rates, the government's borrowing costs will increase, which will then have to be funded by a rise in taxes or a reduction in public expenditure). The implications of monetary policy for fiscal policy, as shown in the HANK models, can have additional redistribution effects depending on whether it shifts resources from debtors or savers, and they serve as a tool to study fiscal policy effects on aggregate productivity and social insurance (Kaplan et al, 2023).

Figure 4: Monetary policy transmission mechanisms in the HANK model



Source: What have we learned from HANK models (Violante, 2021)

63. In terms of structure, HANK models operate in the same way as earlier representative agent models for the production and monetary policy blocks, and they can be summarised in the same three aggregate equations mentioned above for DSGE models. The difference lies in the household block, where the representative consumer is replaced, recognising the heterogeneity of consumers, and the aggregate IS curve[10] is

[10] Represented in most of the models by the Euler equation.

replaced by the modern theory of consumption and savings. In a simple model, there are three groups of households in the economy: households with very low liquidity; 'hand-to-mouth', households with a strong precautionary saving motive and with most of their income coming from labour; and households with high-net-worth individuals for which the precautionary savings motive is trivial and most of their income comes from capital gains. This last group has a low marginal propensity to consumer and the households are usually the consumer type in a representative agent model (Violante, 2021).

64. Another academic advance is agent-based modelling (ABM)[11], a computational technique useful for representing individual behaviour to study social phenomena. These models typically use 'agents', that are a population of heterogeneous objects in an economic or social environment (Axtell & Farmer, 2022). The individual agents are given some explicit rules of behaviour (from very general to very specific), and interact with one another through networks (social, spatial or physical, exogenously specified or endogenously generated). These models can produce an agent equilibrium, or they can perpetuate the dynamic as agents keep adjusting their behaviour. In general terms, ABM models often possess:

- One or more populations of agents, representing an individual or a group, with state information.
- Agent behavioural specifications conditional on the state of the agent.
- An external environment in which the agents are embedded.
- A scheme for agent updating.
- Data gathering and statistical facilities for assessing the state and behaviours of the agents.
- A visualisation engine or descriptive statistics to assess the activity of the agents and the performance of the model.

65. Applied to economics, this type of modelling can relax the assumptions that are used in mathematical models and provide a novel and flexible technology to approach the problems of maximising agents facing constraints. For example, instead of solving equations for equilibrium, the model is constructed specifying the agents and the restriction, and it lets the agents interact with one another. These behavioural rules produce specific agent states where the system becomes dynamic, evolving form one state to another. Each realisation of the ABM is a 'sufficient theorem', in the sense that with certain initial state S, engaging in a specific behaviour B, after some interactions N, the agents will have a new state S' (S,B,N are sufficient to produce S'). Since the early

[11] Some economists that use these methods call them Agent-based computational economics (ACE)

2000s, there has been interest in applying ABM to macroeconomic modelling (MABM), although limited by the computing power. However, there has been some progress in the past few years, where some MABM is capable of including households, firms, banks and the financial sector, with policymaking through a central bank. Most of these models can generate endogenous fluctuations/business cycles and can be used for policy analysis, although there remains much work to be done in this area (Axtell & Farmer, 2022).

Frontier ideas: AI, gaming and LLMs

66. Apart from the recent academic advances in economic modelling, there are also non-economic modelling advances that can help to shed light on modelling improvements. This is the case with Artificial Intelligence (AI) models. AI models can be described as programs that detect specific patterns using a collection of data sets, then drawing conclusions or conducting actions depending on those conclusions. They can be used for different activities, including predictive modelling and forecasting. The key to these models is to have large data sets, so the algorithm can find correlations in pattern and trends that can be used to forecast or formulate strategies (Hewlett Packard Enterprise, 2023).

67. One of the most frequently used AI models is machine learning (ML) models, which can be defined as models that give a machine the capability to imitate intelligent human behaviour and to learn without explicitly being programmed (Brown, 2021). These models can have three functions:

- Descriptive: the models used the data to explain what happened
- Predictive: the models used the data to predict what will happened
- Prescriptive: the models used the data to make suggestions about what to do

68. There are also three categories of ML models: supervised, unsupervised, and reinforcement. Supervised models are the most common category and work by taking label data while being trained to learn and grow. This type of model requires a team of experts to evaluate the results. Some examples of the algorithm used in these models are decision trees, random forest, linear regression and logistic regression. The common use of these models is to take the data to predict outcomes. Unsupervised models work by taking unlabelled data to evaluate relationships between different data points, creating clusters to give more insight to the data. These types of models can help with finding patterns and trends that people are not explicitly looking for. Finally, reinforcement models learn by taking feedback from their actions, usually being comprised of two parts: the agent that performs the action and the environment in which the action is performed. The algorithm used in these types of models works as a cycle in which the environment sends an initial signal to the agent to perform and action, and once this is performed, the environment sends another signal to the agent (positive or negative), so the agent can evaluate the next action in light to the new information (Tableau, 2023).

69. In the future, the AI reinforcement algorithm could be also used to create virtual economic environments where policymakers can understand the behaviour of complex economic and financial systems. This idea is linked to gaming environments, as there are already games with some sort of economic dynamic, such as EVE Online or World of Warcraft. There could be a multi-person dynamic game to explore the behaviour of a virtual economy, that includes the reactions of the players to policy interventions or shocks (Haldane, 2018).

70. On a different note, some studies have explored the use of AI models, more specifically Large Language Models (LLMs[12]) to produce forecast of economic variables. That is the case of an FRB of St. Louis study that produced a conditional inflation forecast during the 2019-2023 period at different horizons, using Google AI's PaLM and compared them to those made in the Survey of Professional Forecasters as well with actual inflation data (Faria e Castro & Leibovici, 2023). The results showed that LLM forecasts generate lower mean-squared errors in most years and almost all horizons, indicating that this method can be applied to other time series. The authors suggest that this approach can be less costly than a traditional DSGE model, which can be complex to run. The advantage of using these kinds of models is that they can capture complex relationships due to their scale and sophistication, having access to current data to generate real-time forecasts. However, they operate as a 'black box', which makes it difficult to understand the dynamics of the model. Additionally, other challenges related to the use of LLM models for forecasting can be related to robustness, as the results provided by the model are sensitive to the way the prompts are structured; reproducibility, as there is a degree of randomness in the results of the model; and external validity, as there is no control over the data used to train the model (Faria e Castro & Leibovici, 2023).

Outside economics: weather forecasts

71. Another type of non-economic forecast worth exploring in relation to economic modelling is weather forecasting. Weather forecasting is complex and requires constant updates. In general, it comprises three steps (Met Office, 2023):

- **Knowing what the weather is doing now:** weather variables are being recording 24 hours a day around the globe. These observations can differ in type and can be satellite, land surface, marine, upper air, radar and thunderstorm location observations. These data are passed onto weather forecast teams to be run in supercomputers several times a day.

- **Calculating how this will change in the future:** the supercomputers use models that estimate how the weather will change based on the observations as 'starting conditions', using complex equations to predict how the weather will evolve. The calculations are repeated over time, each time for a few more minutes, which

[12] Popular LLM models are Open AI's GPT-4 and Google AI's PaLM

enables the modelling team to produce forecasts from hours ahead to a hundred years ahead.

- **Using meteorological expertise to refine the details:** the meteorologist supervises the forecasts produced by the models and checks if they are running according to plan. If not, they adjust the forecasts based on scientific knowledge of how the atmosphere works and experience from previous weather forecasts and observations.

72. The Met Office has been using a unified model for weather and climate prediction since 1990. This model applies a seamless modelling approach, which means that the same dynamical core and parameterisation are used across a broad range of spatial and temporal scales. This model can be used for global and regional forecasts, and it is suitable for seasonal forecasting as well as forecasts ranging from a few days to hundreds of years (Met Office, 2023). The results of the Met Office models are compared with those from other forecasting centres in the world, which is helpful when the forecast made cannot be compared to observations as it too far away in the future. There can also be comparisons between other models that use the same starting conditions but run multiple times to give a range of possible outcomes. These models are called ensembles. If the results of all ensembles are approximately the same, confidence in the forecast would be high. Confidence in the model is also high if the models are consistent, meaning that the model run in the morning gives similar results to the one run during the night.

73. One important innovation is the supercomputing system use in the Met Office, the Cray XC40, one of the most powerful in the world dedicated to weather and climate[13]. This project was completed in 2016 and consists of three main systems: two identical machines that provide the capability to run time-critical operational weather forecasts; and a single larger system in a data centre nearby, that provides research, development and collaboration capabilities (Met Office, 2023). Additionally, regarding next generation modelling systems, the Met Office has mentioned that there is not only a need to increase the number of computer processors but also new types of processors that explore heterogeneous architectures, that is, a mix of processor types. To do that, the design of the weather and climate predictions needs to change, both the algorithms and the software for these algorithms (Met Office, 2023). As this is a challenging task, the new system needs to be flexible and portable to different architectures, and requires collaboration between computational scientists, software engineers and weather and climate scientists.

74. Another example is the European Centre for Medium-Range Weather Forecasts (ECMWF), based in Reading, which has been testing three AI models from Huawei, Nvidia

[13] The three Cray XC40 supercomputing systems are capable of over 14000 trillion arithmetic operations per second, contain 2 petabytes of memory, a total of 460000 compute cores and 24 petabytes of storage for saving data.

and DeepMind, with its own integrated forecasting system. DeepMind's GraphCast was found to be more effective than Pangu-Weather from Huawei and FourCastNet from Nvidia, as well as showing greater accuracy than some of the ECMWF's own forecasts.

75. GraphCast uses a graph neural network, trained using over 40 years of information on past weather systems from the ECMWF (Lam, et al, 2023). The model takes global atmospheric states, now and six hours prior, and produces a ten-day forecast within one minute, running on a single cloud computer. The conventional method, known as Numerical Weather Prediction (NWP), uses supercomputers to generate predictions in a process that can take several hours. This approach is not a replacement for conventional forecasting methods, but rather shows evidence that MLWP can complement and improve the current leading methods. The ECMWF maintains their own supercomputer, which is upgraded every four to five years; it was last upgraded in October 2022 at a cost of over €80m (ECMWF, 2023).

76. In terms of process and insights to data, these machine learning weather models from Google, Nvidia and Huawei maintain an open science approach: sharing methods and results with the aim of accelerating progress (UNDRR, 2023). The ECMWF maintains Github repositories with experimental code for weather and climate forecasting. These three models can be installed and run locally by any individual in a single command. The daily forecasts of each model are published on the ECMWF's public charts page.

77. Lessons learned from the weather forecast and technological advances seen at AI, gaming and LLMs are new directions towards modelling that can be incorporated into economics and particularly macroeconomics forecasts. New modelling approaches such as ABM and HANK models can also present a good addition to the set of traditional models used to evaluate policy decisions.

Section 3: The future – economic models in 2030

78. The approach to macroeconomic modelling has changed over the years. However, there are many ways in which modelling techniques could be improved and the performance of forecasts enhanced. Section two explored the novel approaches that are being used across academia and what lessons could be learned from industries such as finance, climate science and computer science. This final section identifies three areas where economic models could develop further: data, code and computing power.

Centralised databases

79. Models run on data, so ensuring that inputs are up-to-date and consistent across modellers is key. At present the inputs to competing models are opaque.

80. The growing use of Application Programming Interfaces (APIs) is one clear route for improvement here, and it could provide additional value to forecasters if it facilitated access to alternative data sources. For example, if it were provided through a new API, forecasters could trial the inclusion of novel micro-data generated measures in their models. Implementing such a system would significantly reduce the labour costs associated with manual data collection. Furthermore, it would ensure that forecasters have access to uniform, up-to-date information, enhancing the consistency and comparability of their models.

81. Big data can be characterised using the 'three V's': volume, velocity and variety. In these terms, data analytic approaches have improved some understanding of the functioning of the economy at the Bank of England. In terms of volume, some new approaches to data collection have improved the availability of data, such as the MIT's 'Billion Prices Project', which can collect around 15 million prices daily. Web scraping is also being explored by the ONS, to complement its data collection methods. The ONS has been able to collect 7,000 price quotes per day[14] just focusing on groceries and clothing (Haldane, 2018). Recently, it was also able to include information on second hand cars, while planning to add more data on private rents (ONS, 2023). These improvements can create real-time map activity flows across the economy, which can then be scoped to model.

82. Regarding velocity, higher-frequency data can give new insights into economic trends, as well as market dynamics and dislocations. After the global financial crisis, the G20 agreed to collect a greater amount of data on financial transactions in the Trade Repository, to develop a better understanding of market dynamics at stress moments. In recent years, around 11 million reports are collected each working day.

[14] This is larger than the current monthly collection for those items.

83. With reference to variety, the use of non-numerical data has been very limited in economics. Being able to capture words can give another understanding of human behaviour and decision-making. Furthermore, it can improve economic communication. The Bank of England has been able to use new techniques, like random forest, to analyse the Periodic Summary Meeting letter to financial firms and extract data on their content and tone, to see if they express a clear and consistent message (Haldane, 2018).

84. Analytic big data techniques have been the path less followed by economists, despite their rapid growth and popularity. This might be related to the tendency of economists to focus more on theoretical over empirical approaches, which has proved to be a disadvantage in some modelling approaches like DSGE models (Haldane, 2018). For this reason, there might be high returns for economics to embrace the big data analytic techniques and improve the learning process between empirics and theory.

85. An example of this is the Nomis API, provided by Durham University on behalf of the ONS. Nomis provides a comprehensive range of up-to-date and historical data from surveys and administrative sources collected by the ONS and other government departments (NOMIS, 2023). In contrast to the idiosyncratic routes that are sometimes required to access data from other public sector sources, Nomis makes batch downloading of data easy. The API provides a standardised way to access data programmatically. This standardisation is crucial for users who need to integrate these data streams into their analysis or modelling workflows. For example, a data user needing inactivity data for Darlington can always expect to find it at the same endpoint[15], which allows any pipelines or products using the data to update automatically without the need for any manual download.

86. Another example of this is Kaggle, now a subsidiary of Google, which is a data science and artificial intelligence platform, known for its data science competitions and its abundance of interesting user-provided data sets. Having a popular and central repository for novel data encourages exploratory research and innovation. For example, finding a dataset of protected designation products aided the Economics Observatory's (ECO) inflation micro-data research by helping it to segment products by country of origin. ECO has its own data API, which aggregates economic and social data from a range of different statistical agencies and other sources. For example, the endpoint https://api.economicsobservatory.com/chn/wage always returns the most recent Chinese wage data. Although it is far from the scale that would be useful for builders of large macroeconomic forecasting models, it demonstrates how different sources can be aggregated into a single API.

[15] See: https://www.nomisweb.co.uk/api/v01/dataset/NM_17_1.data.json?geography=1811939329...1811939332,&cell=404816129&measures=20100,20701

87. If a centralised solution for dissemination of the data used in forecasting was constructed, best practice would likely be an API that provides access to data from a range of different sources. This API would grant access to a diverse array of data sources, merging information from public entities like the ONS with private sector indicators, such as NatWest's house price figures. Implementing such a system would significantly reduce the labour costs associated with manual data collection. Furthermore, it would ensure that forecasters have access to uniform, up-to-date information, enhancing the consistency and comparability of their models.

88. A centralised API could provide additional value to forecasters if it facilitated access to alternative data sources. For example, if it were provided through a new API, forecasters could trial the inclusion of novel micro-data generated measures in their models. For example, with Richard Davies, the ECO is working on producing daily grocery inflation data using thousands of scraped prices each day. If data like these were available on the API, forecasters could experiment with high-frequency outputs. However, achieving stakeholder buy-in might be difficult. Migrating their data pipelines to a system using a new centralised API make them reliable on the continuous functioning of an external tool. A single point of failure for forecasting would be created – if the API were to become unavailable because of technical difficulties (or more seriously, if the project were to lose funding and the API were to disappear) modellers would be left without data.

Repository of codes and library of models

89. Modelling can be repetitious and error prone. The ability to cross-check work – both within and outside an institution – is vital. This means that having a library of models and repository of code can bring additional benefits to macroeconomic modelling and forecasting. These help forecasters to share and update their code easily, as well as to address concerns and critiques in a collaborative way. Some of the institutions mentioned in Section one already have platforms where the models and the code can be archived and used across the institution, including the COMPASS platform and the IT infrastructure of the Bank of England. However, it could be useful to have a similar platform at a larger scale.

90. A good example of this is Hugging Face, Inc., a French-American company that builds tools for machine learning development (Hugging Face, 2023). It offers a platform where users can share machine learning models and datasets and showcase their work. Additionally, it offers computing power to demo models in small web applications, termed ‘spaces' and ‘widgets,' which are designed for demonstrating small-scale machine learning applications. The open nature of the platform encourages collaboration and iterative improvement. For example, in February 2023, Meta released the weights to its Llama LLM. Other researchers have then used these weights and fine-tuned the model, making derivatives that offer better performance in different use cases (for example, CodeLlama excels at generating usable code) or operate on less powerful hardware. In

addition to hosting models, Hugging Face offers tools for building, running and operationalising ML models. Through all these functions, it acts as an accelerator for open-source AI resource.

91. A similar platform could potentially host various UK macroeconomic models, fostering a culture of collaboration where researchers could adapt existing models like those from the Bank of England or the OBR. Economists using a Hugging Face-like platforms could 'fork' existing macroeconomic models to explore various scenarios. This process involves making copies of the models and modifying them to assess the impact of different economic policies or shocks. If institution's models were more accessible to individual researchers, there could be more integration between forecasting and standard macroeconomic research.

92. Moreover, opening up models and facilitating collaboration could also improve modelling by providing scrutiny. Community engagement with models could reveal bugs, limitations or data inconsistencies, which could then be addressed. When bugs are discovered in open-source software, the party that uncovers an issue can even suggest a fix as the code is open to all. However, there would likely be frictions in getting stakeholders to bring models to a new repository. Forecasters may rely on organisational infrastructure and a proprietary codebase that they do not want to open source or transition to a new platform. Additionally, there may be less potential for iterative improvement. Given enough computational power, it is relatively easy to fine-tune AI models, but improving economic models must be motivated by theory.

Computer power to solve the models

93. Some of the modelling techniques explored here require significant computer power to be run in a timely manner. Furthermore, as technology advances, more powerful computer programs and software have been created to address previous limitations, such as working with large-scale models that required a lot of data processing. These new technological advances require more powerful computers able to support and run the models without clashing. For this reason, having an improved computer system should be one of the priorities in any macroeconomic modelling and forecasting department.

94. There are lessons from outside economics here. The weather forecasting supercomputing system CrayXC40 used by the MET Office to run their models and produce their forecast. In the world of ML, the current rate of development for LLMs and other AI models is extremely high and is dependent on the computing resources available. The companies leading in LLM development have direct access to the largest cloud

computing providers. The rapid development of compute-infrastructure, such as with AI-optimised GPUs, is facilitating the similarly rapid development of AI models[16].

95. The use of cloud computing is an important consideration. In-house compute solutions are more costly and risk going out-of-date quickly, creating a bottleneck in subsequent years. Cloud hosted computing, like that offered through AWS, Azure or Google Cloud could expand the capabilities by offering more computational power to devote to models. It is likely that some forecasts, particularly those run by large institutions, are already run through cloud providers. However, a lack of expertise among macroeconomists and smaller research groups may hold back access to cloud compute. Cloud providers are not primarily tailored towards researchers so deploying models in the cloud can require data engineering skills that some do not possess.

96. This is an issue that a centralised provider of compute for researchers could address. Platforms already exist to simplify the provision of compute in other contexts, such as Vercel wraps AWS for web apps; Vast.ai simplifies access to GPUs for AI. Researchers could be provided compute via a platform built to simplify the process of running models in the cloud. Any cloud provision of compute could, however, see issues with proprietary software. Running models written in open-source languages such as Python, R, Julia, or C++ would be easy to bring into the cloud. However, others written with proprietary solutions like STATA or MatLab may face issues. For example, STATA licences vary by core count and end-user licences would probably not cover cloud hosting. While a barrier, these are the kind of issues that a platform with institutional backing could fix.

Recommendations

97. The macroeconomic forecasting landscape in the UK is broad and evolving. Among the six institutions reviewed in this paper, there is a wide range of approaches and models. While this demonstrates independence, it also results in a lack of standardisation across several metrics such as forecast horizons and data usage. Technological developments have also expanded the tools available to forecasters, resulting in greater speed, granularity and accuracy. Models produced in the UK could improve across both operational and technical dimensions to produce better, more consistent and more precise forecasts.

98. **Standardisation of forecast and forecast evaluations.** There are inconsistencies between forecasting institutions with regard to the length of the forecast period and the frequency at which the forecasts are published. Standardising the forecast horizon would enable comparability across the six forecasts, and therefore, their expectations for the

[16] For example, according to leaks, GPT-4 was trained on 25,000 Nvidia A100 GPUs for 90-100 days. If this was trained on 25,000 of the previous generation V100 GPUs, it would have taken roughly four times longer (approx. 12 months).

UK's economic performance. Since forecasts are used to inform policy decisions, and policymakers often make policy choices now based on their expectations of the future, a longer forecast horizon would enable them to make more informed decisions. Additionally, there is no industry convention nor a requirement to produce model evaluations on a more frequent basis. Establishing such a requirement for evaluations would provide users of a forecast with an understanding of whether the forecast was upward- or downward-biased, whether due to modelling assumptions or exogenous factors. Although some of the differences between forecasts are driven by assumptions depending on the forecasts purposes, which can drive differences in approach and therefore lead to different results, producing a guidance document that acknowledge the differences but highlights the benefits of these measures can be a step forward to opening the discussion with the macroeconomic modelling community.

99. **Use of large datasets and improved data consistency.** Developments in data collection have made large and real-time data more accessible. However, the use of big data that can both demonstrate the behaviour of agents within the economy as well as improve the accuracy of models has not been fully realised. There are also inconsistencies between institutions regarding the data that are used. Developing an API that provides more frequent and detailed data, could improve the performance of economic forecasts, and facilitates the interaction between modelling approach and data required.

100. **Improved computing power.** Technological advances have provided faster and more powerful computers. Having an improved computer system is essential in any macroeconomic modelling and forecasting department, as greater computing power will be required to process larger datasets and support more complex modelling techniques. It is important to balance causality and predictive power, as in any approach there should be a way of explaining how the model got to its result to reduce fears of a black box model. Computing power can also help in the faster development of different models and facilitates the integration of a suite of models in the forecasting exercise.

101. **Macroeconomic modelling and the research community.** One of the big challenges in macroeconomic modelling is the lack of competition among institutions that develop macroeconomic models and forecasts. At the same time, the forecasting process involves a great deal of judgement. This judgement ultimately comes from the understanding of the economy and the macroeconomic relationships that can explain the outcome of the model. For that reason, is important to invest in macroeconomic research and create a community where the development of models can be shared and the lessons from other institutions and research units can work towards the improvement of macroeconomic modelling and forecasting. Initiatives such as the ESRC Warwick Macroeconomic Modelling Bureau and its annual Macroeconomic Modelling Seminar could be restored. Furthermore, this community can also help to raise awareness about the challenges of macroeconomic forecasting and facilitate funding to attract human capital to the area of macroeconomic research. Putting in place a Macroeconomic Modelling Conference can

be a reasonable first step to bring in different actors around the topic, not just government and officials, but also academics and the private sector, and could touch on development topics such as those mentioned above in the recommendations (relating to standardisation, evaluation, data and computing power). Expanding links with international practitioners can also help to enrich research on current and relevant topics around modelling forecasting such as climate change. Ultimately, building a space for the macroeconomic modelling and research community would facilitate the improvement of macroeconomic forecasts and the integration of research, theory, and practice to promote a better macroeconomic stand in the UK.

102. These steps could be implemented to improve macroeconomic models and forecasting from a technical perspective as well as their use by UK policymakers.

References

Atkins, G., & Lanskey, L. (2023). Working paper No. 19: The OBR's forecast performance. London: OBR.

Axtell, R. L., & Farmer, J. D. (2022). Agent-Based Modelling in Economics and Finance: Past, Present and Future. Oxford: INET Oxford Working Paper No. 2022-10.

Bank of England. (2024, 23 January). Bank of England Agenda for Research – Research: [Bank of England Agenda for Research | Bank of England]

Bank of England. (2024, April 12). Forecasting for monetary policy making and communication at the Bank of England: a review.

Bank of England. (2024, April 12). Response of the Bank of England to the Bernanke review of forecasting for monetary policy making and communication at the Bank of England.

Bayoumi, T. (2004). GEM A New International Macroeconomic Model. Washington DC: IMF.

Brown, S. (2021, 21 April). Machine learning, explained. Retrieved from Ideas made to matter - Artificial intelligence - MIT Management Sloan School: https://mitsloan.mit.edu/ideas-made-to-matter/machine-learning-explained

Burgess, S., Fernandez-Corugedo, E., Groth, C., Harrison, R., Monti, F., Theodoridis, K., and Waldron, M. (2013) The Bank of England's forecasting platform: COMPASS, MAPS, EASE and the suite of models. Working paper no. 471. Bank of England.

Celasun, O., Lee, J., Mrkaic, M., & Timmermann, A. (2021). An Evaluation of World Economic Outlook Growth Forecast - 2004-17. IMF Working Paper.

CFI. (2023, 20 December). Value at Risk (VaR).

Davies (2024), "Bernanke Review: Reflections", in *The Bernanke Review: Responses from Bank of England Watchers,* King's College London, April 2024.

Dou, W. W., Lo, A. W., Muley, A., & Uhlig, H. (2020). Macroeconomic Models for Monetary Policy: A Critical Review from a Finance Perspective. Annual Review of Financial Economics, 12, 95 - 140.

ECB. (2021). Review of macroeconomic modelling in the Eurosystem: current practices and scope for improvement . Occasional Paper Series No 267, European Central Bank .

ECMWF. (2023, 5 January). Supercomputer facility. Retrieved from Our facilities: https://www.ecmwf.int/en/computing/our-facilities/supercomputer-facility

Faria e Castro, M., & Leibovici, F. (2023). Artificial Intelligence and Inflation Forecasts. St. Louis: Federal Reserve Bank of St. Louis Working Paper 2023-015.

Fawcett, N., Körber, L., Masolo, R. M., & Waldron, M. (July 2015). Evaluation UK point and density forecasts from an estimated DSGE model: the role of off-model information over the financial crisis. London: Staff Working Paper No. 538, Bank of England.

Hackworth, C., Radia, A., & Roberts, N. (2013). Understanding the MPC's forecast performance since mid-2010. London: Quarterly Bulletin 2013 Q4, Bank of England.

Haldane, A. G. (2018). Will Big Data Keep Its Promise? Speech Bank of England: Data Analytics for Finance and Macro Research Centre, King's Business School.

Haldane, A. G., & Turrell, A. E. (2017). An interdisciplinary model for macroeconomics. London: Staff Working Paper No. 696, Bank of England.

Hewlett Packard Enterprise. (2023, December 15). What are AI models? Retrieved from Enterprise glossary - Key definitions - AI models: https://www.hpe.com/us/en/what-is/ai-models.html#:~:text=How%20do%20AI%20models%20work,set%20provided%20to%20the%20program.

Hugging Face. (2023, 5 January). Hugging Face. Retrieved from https://huggingface.co/

Hurst, I., Liadze, I., Lisenkova, K., & NIESR. (2014). Overview of the NiGEM-S Model: Scottish version of the National Institute Global Econometric Model. London: ESR Discussion Paper No. 422.

Independent Evaluation Office. (November 2015). Evaluating forecast performance. London: Bank of England.

Institute for Government. (2022, 28 October). Office for Budget Responsibility (OBR). Retrieved from https://www.instituteforgovernment.org.uk/explainer/office-budget-responsibility-obr

Kaplan, G., Moll, B., & Violante, G. L. (2023, March). The very model of modern monetary policy. Retrieved from IMF Finance and Development:

Lam, R., Sanchez-Gonzales, A., Willson, M., Wirnsberger, P., Fortunato, M., Alet, F., . . . Battaglia, P. (2023). Learning skillful medium-range global weather forecasting. Science, 382, 1416-1421.

Laxton, D. (2008). Getting to Know the Global Economy Model and its Philosophy. IMF Staff Papers Vol 55 No. 2.

Mellis, C. (1988). HM Treasury Macroeconomic Model 1986. Economic Modelling, 237-260.

Met Office. (2023, January). How weather forecasts are created. Retrieved from Weather and Climate: https://www.metoffice.gov.uk/weather/learn-about/how-forecasts-are-made

NIESR. (2023, 4 December). The leading global macroeconomic model. Retrieved from NiGEM Macroeconomic Model: https://www.niesr.ac.uk/nigem-macroeconomic-model

NOMIS. (2023, 5 January). Welcome to Nomis. Retrieved from nomis official census and labour market statistics: https://www.nomisweb.co.uk/

OBR. (2011). Memorandum of Understanding between the Office for Budget Responsibility and HM Treasury – the macroeconomic model. London

OBR. (2013). Briefing paper No. 5: The macroeconomic model. London.

OBR. (2021, 12 July). OBR macroeconomic model. Retrieved from OBR macroeconomic model variables: https://obr.uk/forecasts-in-depth/obr-macroeconomic-model/

OBR. (2023). Forecast evaluation report. London: OBR

OBR. (2023). Who we are. Retrieved from https://obr.uk/about-the-obr/who-we-are/

OECD. (2011, December). Forecasting methods and analytical tools. Retrieved from Economic outlook, analysis and forecast: https://www.oecd.org/economy/outlook/forecastingmethodsandanalyticaltools.htm

ONS. (2023, 6 July). Transformation of consumer price statistics: July 2023.

Pagan, A. (2003). Report on modelling and forecasting at the Bank of England. London: Bank of England Quarterly Bulletin.

Pescaroti, A., & Zaman, S. (2011, 5 October). Macroeconomic Models, Forecasting, and Policymaking. Economic Commentary, p. Number 19.

Pike, T., & Savage, D. (1998). Forecasting the Public Finances in the Treasury. Fiscal Studies, 49-62.

Sargent, T. J. (2023). HAOK and HANK Models. New York.

Sborde, A. M., Tambalotti, A., Rao, K., & Walsh, K. (2010). Policy analysis using DSGE models: an introduction. FRBNY Economic Policy Review.

Singh, M. (2020). Deep Models for Empirical Asset Pricing (Risk-premia Forecast) and their Interpretability. MIT.

Stockton, D. (2012). Review of the Monetary Policy Committee's forecasting capability. London: Presented to the Court of the Bank of England.

Tableau. (2023, 15 December). Artificial intelligence (AI) algorithms: a complete overview. Retrieved from Data insights: https://www.tableau.com/data-insights/ai/algorithms

Turner, D. (2016). The use of models in producing OECD macroeconomic forecast. Paris: Economics Department Working Papers No. 1336 OECD.

UNDRR. (2023, 6 September). How AI models are transforming weather forecasting: a showcase of data-driven systems.

Violante, G. L. (2021). What have we learned from HANK models, thus far? Princeton University.

Warwick. (2014). More about the ESRC Macroeconomic Modelling Bureau.

Annex

A. Financial market models

103. Machine learning models can be used to measure risk premia and solve asset pricing problems. Financial markets are complex, and linear relationships cannot fully capture the interactions between the factors and their returns. In modelling asset pricing, typically there are two problems to solve: explaining the variance in the cross-section of returns; and studying the assets' risk premia (Singh, 2020). ML models can help to solve these problems as they use popular methods to solve non-linear problems, for example, deep networks method to model risk-premia forecasts.

104. However, the more traditional models in finance, and in particular risk management, use the value-at-risk (VaR) approach. Conventional VaR measures the amount of potential loss that could happen in an investment over a specified period of time, assuming that the returns follow a normal or conditional normal distribution. There are three methods to calculate VaR: historical method, parametric method, and Monte Carlo method (CFI, 2023). The advantages of using VaR methods are that it is easy to understand, as it is a single number (unit or percentage) that indicates the risk in a portfolio, can be applicable to all types of assets (bonds, shares, derivates, currencies), and it is widely used and accepted. The limitations can be related to the difficulty in the calculation when there are large and diverse portfolios, the different approaches to calculate the VaR that can lead to different results, and the assumption required for the VaR to be valid. For example, the normal distribution assumption, which can lead to substantial bias in the results due to an underestimation of volatility. Other quantitative analysis to measure risk can include the standard deviation, skewness, kurtosis and the Sharpe ratio (see Figure 5). The Sharpe ratio is also one of the most popular measures used by hedge funds and compares the return of an investment with its risk.

Figure 5: Value at Risk and Sharpe ratio equations

$$VaR = v_m \frac{v_i}{v_{i-1}}$$

$$Sharpe\ ratio = \frac{(Rx - Rf)}{Std\ Dev\ Rx}$$

Source: Corporate Finance Institute (CFI)
Notes: in the VaR equation v_i is is the number of variables on day i and m is the number of days from which historical data is taken. In the Sharpe ratio equation Rx is the expected portfolio return, Rf is the risk-free rate of return, Std Dev Rx is the standard deviation of portfolio return / volatility

B. Models for policymaking

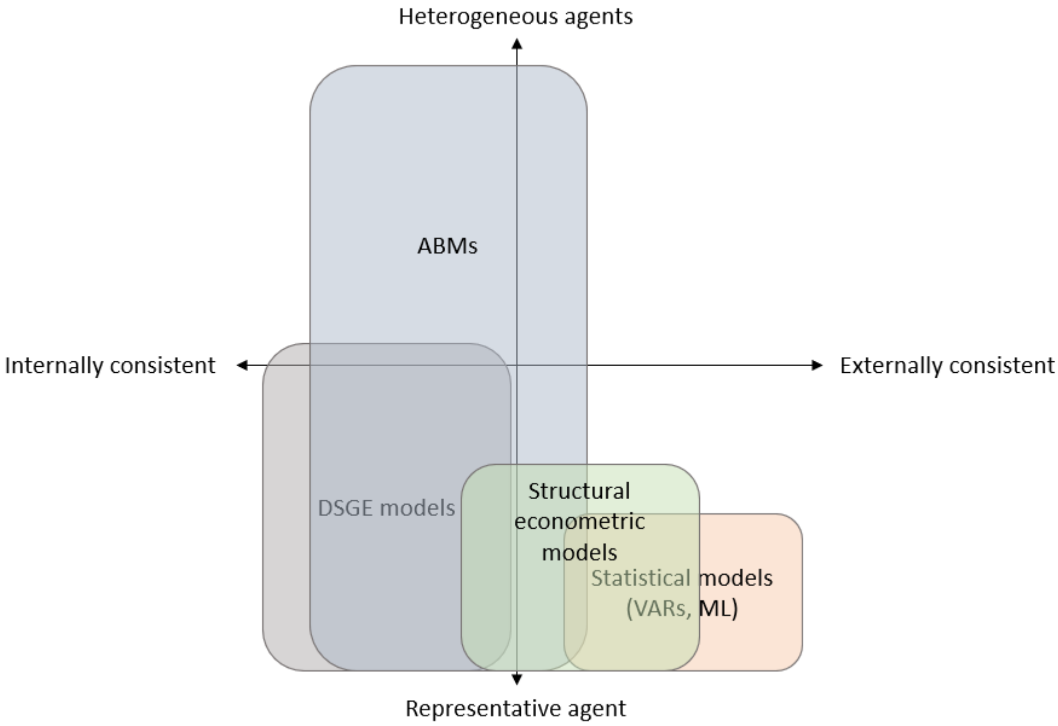
105. The history of macroeconomic modelling can be traced from the Great Depression of the 1930s and its aftermath, which inspired economists to model the new economic conditions and find policies according with the necessities of the time. Over the decades, other crises have continued to inspire economists to improve macroeconomic modelling, such as the recessions of the 1970s and the beginning of large-scale macroeconometric models and New Keynesian DSGE models, and the global financial crisis, which raised concerns about the lack of integration of the financial sector in macroeconomic models (ECB, 2021). Today, the vast majority of macroeconomic modelling is concentrated in monetary policy analysis at central banks, with most of them using large-scale semi-structural models for projections, and structural (typically DSGE) models for scenario and policy analysis.

106. The flexibility of large-scale semi-structural models has proved to be a strength over recent decades, being able to re-specify simple equations to incorporate expert views and key judgement on the projections, while the scenario and policy analysis provided by the DSGE framework complements the semi-structural models, adding an internally consistent narrative to the projections and making it possible to explore policy impacts over the aggregate economy (ECB, 2021). At the same time, a new set of improvements in macroeconomic modelling has been developed in the past decade, adjusting the existing frameworks to solve some of the main concerns, namely the expectations formation process, the transmission of non-standard monetary policy measures, the incorporation of financial sector variables, and adaptation to external shocks such as climate change and the pandemic. This has also led to the development of new models such as the HANK models and ABM mentioned in the second section of this document, which aim to solve some of these concerns by adding different agent behaviour, using more information and technology available, and drawing lessons from other sciences.

107. The trade-offs of working with these different frameworks can be shown in a modelling space between two axes (see Figure B.1):

- Consistency: internal consistency represented by a strong micro-founded behaviour and external consistency represented by agreement with the data.
- Agents: heterogeneous agents represented by the differences between the types of agents and representative agent represented by the typical rational decision-maker.

Figure B.1: Macroeconomic modelling space



Source: An interdisciplinary model for macroeconomics (Haldane & Turrell, 2017)

108. The use of these later models can help to relax the micro-founded behaviour of the DSGE models as well providing another platform to analyse policy measures and support the policymaking process. Nevertheless, there are some challenges that should be addressed first, such as the difficulty of communicating the results of these complex systems (Haldane & Turrell, 2017). However, HANK models and ABM should be considered as a complement to the already existing macroeconomic modelling frameworks, especially given the macroeconomic challenges of the last two decades.

C. Interview findings

109. As part of our research, we conducted 11 interviews with practitioners and academics, to develop a better understanding of the nature of macroeconomic forecasting in the UK. The interviews were conducted between 18 January and 6 February, and the interviewees included a mixture of academics, with several years of experience in macroeconomic modelling and forecasting research, and practitioners from four of the six institutions on which we focused our research for the first section of this document: HM Treasury, NIESR, the IMF, the OECD, and the Bank of England.

110. One important highlight of this exercise is the difference in the answers between academics and practitioners regarding the use of macroeconomic models and the main challenges for forecasting. While academics tend to be more critical about the challenges and disadvantages of the models and their use for forecasting and policymaking, practitioners emphasise the importance of the models as tools to guide both the forecast and the policy discussion and tend to focus more on the technical aspects of the forecast exercise as the main challenges to address.

111. But there is one common answer between academics and practitioners regarding the future of forecasting and macroeconomic modelling in general: the need to invest in broad macroeconomic research, human capital, and the macroeconomic community to understand the underlying mechanisms of the models and impose reliable judgments on the forecasts.

112. The interview exercise led to some common findings:

113. **The current state of macroeconomic models in the UK:** the structure and the models should depend on the questions you want to answer, since different questions require different models. Large econometric models are usually better for forecasting because they fit the data better, but they are not so good for getting insights into transmission channels. DSGE models are better at the latter, but they have proved to be too rigid, and they can fail to anticipate future risks. In terms of forecasting, forecasts are made by people, for example, international institutions use their country experts to produce their forecasts. Furthermore, it is useful to remember that forecasting and policymaking are two different things. A good institution should be using different types of models to cross-check and guide their forecasts and produce different scenarios. Lastly, the current models are big and often very difficult to understand. Even if they were open source, it would be difficult for outside economists to understand them well.

114. **What is lacking in macroeconomic forecasting:** competition and capacity in broad macroeconomics. Human capital in economics is required to understand the models and develop better solutions. Additionally, there is a lack of a macroeconomic modelling community, and initiatives comparable to the Warwick Macroeconomic Modelling Bureau of the 1980s are missing in the macroeconomic sphere. The adoption of new technology

and the diversity of new approaches are also missing in macroeconomic modelling. This includes the exploration of other types of models and techniques (HANK models, ABM, econometrics, and machine learning). Moreover, there is a lack of realism and approach to uncertainty. There is a common misunderstanding that the economy is not static and there should be better ways to introduce risks (such as financial sector risks) in the forecasting exercise.

115. **Data and computer power role in the forecast exercise:** computer power does not seem to be a constraint given the technological advances. Computer power should be used to advance in other types of models rather than focusing on forecasting. On the other hand, data collection can be a huge advantage, especially regarding models that need parameter estimation. Granularity in data is also an advantage for better understanding of the behaviour of agents in the economy. An API that can be plugged directly into the models can be very useful and facilitate the process of collecting data, which is much more a limitation than the computer power. Data collection takes time and resources, if organisations like the ONS could get some funding to improve the data availability that is useful to macroeconomics it could be a huge improvement.

116. **What can be learned from other sciences:** communication. Weather forecasts have managed to interact well with the public and communicate their forecasts in a much more effective way than economists have ever done. Data sciences and other sciences like physics and engineering are also helpful, in the sense that they are more proficient mathematically and are better at replicating crises, something that economists have not learned how to do yet. Economists have the model and the mechanism very clear, but that is not enough, the models need to be built, maintained, and improved. Modelling departments can be improved by having a diversity of backgrounds in their teams, which would help in the development and maintenance of the models.

117. **Technical aspects of the modelling and forecasting exercise:** the models are large and often hard to understand, which makes communication difficult for policymaking. Models need to be simplified and address specific questions. Theoretical rigidities in the models and their limitations need to be acknowledged, and modelling teams must be able to address them and bend them when necessary. Assumptions such as mean regression can hinder the incorporation of uncertainty in the models, as well as the difficulty of accounting for non-linearities, which limits the analysis of certain shocks. Accounting for global economic changes and the use of tools like fan charts and scenario analysis are needed to acknowledge future risks and to keep a broad perspective about economic conditions.

Economics Observatory

UKRI Economic and Social Research Council

www.economicsobservatory.com