Review

Herman K. van Dijk*

Challenges and Opportunities for Twenty First Century Bayesian Econometricians: A Personal View

https://doi.org/10.1515/snde-2024-0003
Received January 17, 2024; accepted January 28, 2024; published online March 12, 2024

Abstract: This essay is about *Bayesian econometrics with a purpose*. Specifically, six societal challenges and research opportunities that confront twenty first century Bayesian econometricians are discussed using an important feature of modern Bayesian econometrics: conditional probabilities of a wide range of economic events of interest can be evaluated by using simulation-based Bayesian inference. The enormous advances in hardware and software have made this Bayesian computational approach a very attractive vehicle of research in many subfields in economics where novel data patterns and substantial model complexity are predominant. In this essay the following challenges and opportunities are briefly discussed, including the scientific results obtained in the twentieth century leading up to these challenges: Posterior and predictive analysis of everything: connecting micro-economic causality with macro-economic issues; the need for speed: model complexity and the golden age of algorithms; learning about models, forecasts and policies including their uncertainty; temporal distributional change due to polarisation, imbalances and shocks; climate change and the macroeconomy; finally and most importantly, widespread, accessible, advanced high-level training.

Keywords: Bayesian econometrics; research opportunities; forecasting

JEL Classification: C11; C15; C01

Why econometrics should always and everywhere be Bayesian
–Sims (2007)

1 Introduction

As a preliminary remark to the subject of this essay, I mention how in the first quarter of the twenty first century tremendous progress has been made in many scientific fields with important practical applications. As an important illustration and major example of success of a societal challenge and research opportunity in this period, I list the discovery and production of the BioNTech-Pfizer vaccine as an effective medicine to combat the

This paper should not be reported as representing the views of Norges Bank. The views expressed are those of the author and do not necessarily reflect those of Norges Bank. A preliminary version of this paper was presented at the ESOBE 2022 Meeting. The author is indebted to Jamie Cross, Lennart Hoogerheide, Gary Koop, Dimitris Korobilis, Francesco Ravazzolo and Koene van Dijk for helpful comments and stimulating conversations about several of the topics of this paper.

***Corresponding author: Herman K. van Dijk**, Erasmus University Rotterdam & Tinbergen Institute & Norges Bank, Rotterdam, Netherlands, E-mail: hkvandijk@ese.eur.nl

pandemic caused by the Covid-19 virus. Three major features of this success are: an effective migration policy by Germany where talented persons from abroad found a living and work opportunity; the high level of the German university system that is accessible at low cost for qualified persons; and the successful international cooperation and good leadership initiative between the top management from both BioNTech and Pfizer. It is important to note that without vaccines the Delta and Omicron variants of Covid-19 virus would have been three times deadlier, see Figure 1. As a general note, I remark that the 'speed of success' in this case was substantial but it is also important to realize that viruses, some mild others very serious, will affect the health of humans and are here to stay.

In this essay I describe a similar major success for Bayesian econometric inference which has its origin in World War II. It refers to the invention of the Monte Carlo method which consists of generating pseudo-random numbers on a computer and, at the same time, to the development of computers with ever increasing computational power. This method took a longer time than the Covid vaccine to gain acceptance in several scientific fields but a fair conclusion is that the Monte Carlo method is here to stay and is widely applicable.

The Monte Carlo method was invented by Stanislaw Ulam and John von Neumann during World War II working on the Manhattan project with the purpose of improving decisions, see Metropolis (1987), and the references cited there. This approach allowed for numerical evaluation of integrals through statistical sampling methods for a wide range of model structures. Monte Carlo methods were in the early stage not used in Bayesian inference but mainly in physics and gradually in economics. The application in economics consisted of *direct simulation of artificial data* from basic probability models with *given parameter values*. Examples include the arrival of ships at ports or ambulances at hospitals. In a related development, in frequentist econometrics, simulating repeatedly artificial data on a computer from a given parametric model was used in order to explore frequentist properties of estimators. In the latter part of the twentieth century the fundamental step in Bayesian inference was one of reverse engineering. That is, *given data*, one attempts to *simulate parameter values* of posterior and predictive probability models. However, direct sampling was not feasible for a wide class of Bayesian models. One more step was to make use of *indirect* sampling by simulating from a different distribution than the distribution of interest under the condition that this different distribution should be a good approximation to the distribution of interest. This indirect sampling approach implies that a correction factor has to be applied. Two important classes of indirect simulation methods that deal with correction factors are importance

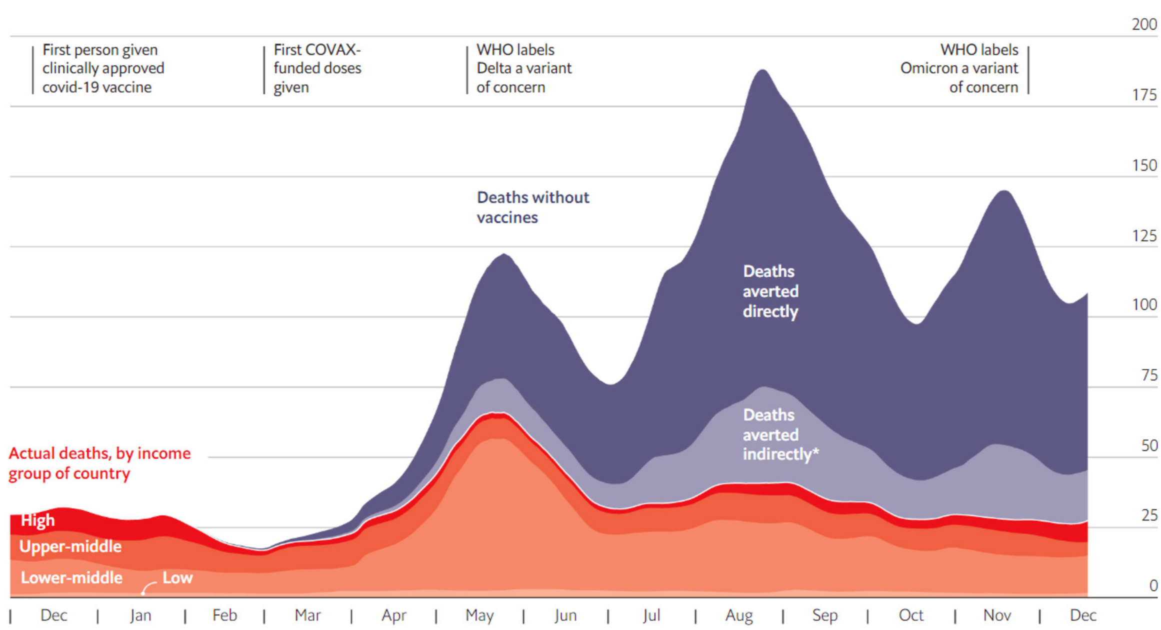

Figure 1: Estimated daily excess deaths (thousands) among individuals who avoided infection thanks to other people's vaccinations. Source: www.economist.com adapted from results in Watson et al. (2022).

sampling due to Goertzel (1949) and Kahn and Harris (1951), and introduced in econometrics and statistics by Kloek and Van Dijk (1978), and Markov Chain Monte Carlo, see Metropolis et al. (1953) and Hastings (1970), which was introduced in statistics by, amongst others, Gelfand and Smith (1990).

These two approaches enabled the performance of the integration operation in conditional probabilities in Bayesian inference more effectively as well as being operational for a wide class of complex models, say, in finance, marketing and macro-economics. This led to more accurate forecasts and a better quantification of uncertainty and risk. For a historical perspective on the rise of Bayesian econometrics, see Baştürk et al. (2014).

The world has substantially changed in the early part of the twenty first century. Communication through the internet has become so prolific that interdependence or connection between many economic fields has occurred which may give the impression of world-wide economic advancement and convergence connecting many areas of economics using data science and high tech simulation methodology. However, recently shocks, economic imbalances, polarisation, inequality and serious climate problems occurred. Both positive and negative developments imply challenges and opportunities for Bayesian econometric inference.

In this essay I discuss six challenges and opportunities for twenty first century Bayesian econometricians including the scientific advances in the twentieth century in simulation based Bayesian inference. The

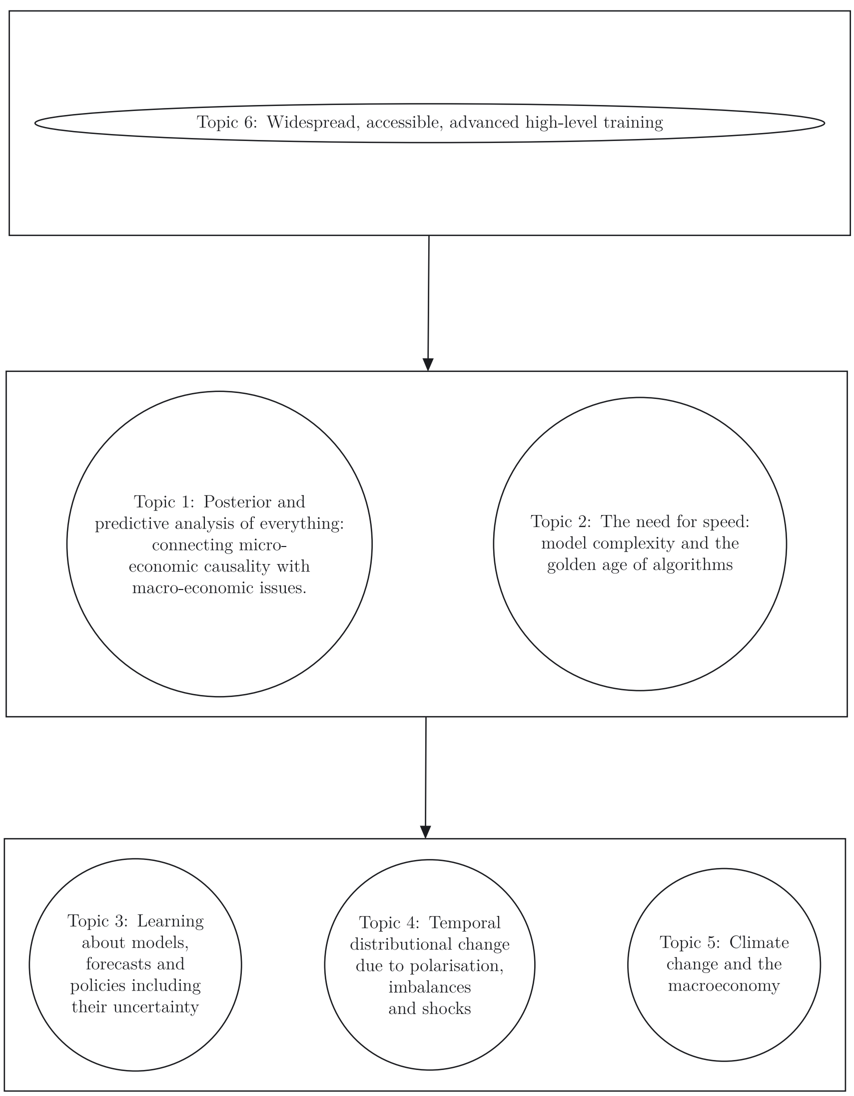

Figure 2: Connections between six research challenges and opportunities.

challenges and opportunities are listed with their interconnections in Figure 2. At the top a most important challenge and opportunity is given as: Widespread, accessible, advanced high-level training. This is crucial input to take up one of the other five challenges. In Section 2.1 the ideal challenge is presented as posterior and predictive analysis of everything:[1] connecting micro-economic causality with macro-economic issues.

This challenge needs input from and interaction with the challenge listed next to it and given as: The need for speed: model complexity and the golden age of algorithms. Here several technical topics are treated which are also needed in the three more applied challenges listed in the second block: Learning about models, forecasts and policies including their uncertainty; Temporal distributional change due to polarisation, imbalances and shocks; Climate change and the macroeconomy.

I close by noting that this essay is a personal reflection on the past, present and future of Bayesian econometrics, and is not meant to provide a survey of these topics with complete references in econometrics and statistics.

2 Six Societal Challenges and Research Opportunities for Twenty First Century Bayesian Econometricians

2.1 Posterior and Predictive Analysis of Everything: Connecting Micro-Economic Causality with Macro-Economic Issues

In the field of economics there have been, at least, two major successes in the twentieth century. In the economy at large there was a tremendous progress in the modeling and forecasting patterns in economic time series. This improved understanding of the dynamic behavior of economies at large has led first to more effective fiscal policy, e.g. New Deal Policy. Second, understanding short and long term patterns in financial time series patterns led to improved measurement of volatility and risk with implications for monetary policy. In the micro-economy the econometric analysis of causal effects using personal data led to introduction of income maintenance programs, better understanding of the search in and working of the labor market; working of education and training programs. A very interesting feature that occurred is that there exists a common structure between macro- and micro-economic models in this period. For instance, modeling the income-education effect in microeconomics; stationary combinations in macro-economic series, and information reduction in large financial models all led to the common structure of a multivariate regression model as shown in Figure 3.

$$\tilde{Y} = \tilde{X}A + E, \;\; R(A) \text{ is less than full}$$

$$\tilde{Y} = (Y \;\; X), \;\; \tilde{X} = Z$$

$$A = (0 \;\; \Pi)\begin{pmatrix} 1 & 0' \\ \beta & I_r \end{pmatrix}$$

modeling direct connectivity;
empirical analysis:
income-education effect

$$\tilde{Y} = \Delta Y, \;\; \tilde{X} = Y_{-1}$$

$$A = (I_r \;\; \beta_2')\alpha'$$

modeling stationary
random walk
combinations;
empirical analysis:
forecasting

$$\tilde{Y} = Y, \;\; \tilde{X} = I$$

$$A = F(I_r \;\; \Lambda_2)$$

$$\Sigma = D$$

modeling information
reduction;
empirical analysis: financial
momentum

Figure 3: Common structure of three classic linear econometric regression models. Source: Baştürk, Hoogerheide, and van Dijk (2017).

In macroeconomics, the integration of micro- and macro-theory developed in the mid-twentieth century and is referred to as ‘microfoundations’. This line of research uses microeconomic principles such as utility optimizing households, and profit maximizing firms, to address macroeconomic questions such as ‘what drives inflation?’ within a general equilibrium framework. This movement tended to focus on representative agents

1 Analogy with Physics’ Theory of Everything: How to combine the General Theory of Relativity (understanding the universe) with Quantum Mechanics (subatomic particles, molecules).

and low-order distributional moments of aggregate macroeconomic data. This perspective has changed in the twenty first century.

First, a major twenty first century development has been the integration of *micro data* in macro models. As discussed in the Nobel Lecture of Heckman (2001), the use of micro data in cross sections beginning in the 1960s revealed the importance of heterogeneity within individual behavior, income, consumption habits and portfolio selection. Between the 1960s and the turn of the century, the pervasiveness of this heterogeneity became more pronounced as newly available panel data revealed that these effects are persistent over time for the same persons, and have important intergenerational consequences. While cross sectional variation alone cannot be used to identify the macroeconomic effects of variables that are common across individuals, such as fiscal and monetary policies, the findings sparked an important debate about the integration of macroeconomic general equilibrium and microeconometrics for calibration and estimation purposes. Econometrically, the use of microdata has allowed us to better measure important macroeconomic variables, such as expectations and sentiment, and shed new light on frequent macroeconomic questions. Examples include the identification and quantification of fiscal policy multipliers (Ramey 2011), the causes of the Great Recession (Mian and Sufi 2011), and the ability to track the economic effects of the COVID-19 pandemic in real time (Vavra 2021).

Another important area of research has been the synthesis of *microeconometric identification methods* within macroeconomic time series models. Since the seminal work of Sims (1980), structural vector autoregressions (SVARs) have been the primary tool of empirical macroeconometricians to understand macroeconomic phenomena, however shock identification has always been the central source of debate (Cooley and LeRoy 1985). Over the past decade, there has been an increase in usage of microeconomic instrumental variables to identify macroeconomic VAR models, which have become known as 'proxy VARs'. An early example of this idea is Mertens and Ravn (2013) who use narratively identified tax changes in post-WWII data as proxies for structural tax shocks within a SVAR model. Caldara and Herbst (2019) and Arias, Rubio-Ramirez, and Waggoner (2021) were among the first to develop Bayesian frameworks for inference in proxy SVARs. More recently, Giacomini, Kitagawa, and Read (2022) propose an algorithm for robust Bayesian inference in proxy SVARs, and Mumtaz and Petrova (2023) develop Bayesian methods for the use of instruments in time-varying SVARs. This raises an interesting question: what other microeconometric identification methods can be used in macroeconometric models?

Finally, over the past few years macroeconomists have begun to recognize the importance of modeling heterogeneity not only within the behavior of individuals at a given point in time, but also within aggregate time series over a given period. Econometrically, important information about such data features can be obtained by looking at higher-order moments like skewness and kurtosis, and other important distributional features such as multimodality. For instance, changes in the quantile behavior of the conditional distribution of GDP growth have been linked to macroeconomic risk (Adrian, Boyarchenko, and Giannone 2019), and multimodality in this distribution has been linked to multiple equilibria (Adrian, Boyarchenko, and Giannone 2021). Mitchell, Poon, and Zhu (2022) show that the existence of multimodality over the business cycle is a key feature in the predictive distribution of GDP growth when conditioning on financial conditions. While tests for multimodality of distributions for continuous random variables such as real GDP have been around for some time, determining multimodality for the case of discrete variables such as survey responses to inflation expectations are rare. A first attempt to overcome this problem is provided in a recent paper by me and co-authors, see Cross et al. (2023). We propose a simple method for mode inference with discrete distributions that is illustrated in Figure 4. The top row contains observations from the University of Michigan's inflation expectations in two periods, along with a Bayesian estimated distribution. The middle and bottom rows respectively contain implied posterior probabilities of the number of modes and their locations. In both periods credible information is shown regarding the quantity and location of modes. There is also strong evidence of increased heterogeneity in survey responses by participants in 2023 relative to those in 2020. This suggests that expectations may have become unanchored in recent times, and this presents an important policy challenge for central bankers who are tasked with maintaining low and stable inflation. While one can speculate that this un-anchoring is likely due to the rising cost of living in the aftermath of the COVID-19 pandemic, identification of causal factors relating to this phenomena is an important area of future research.

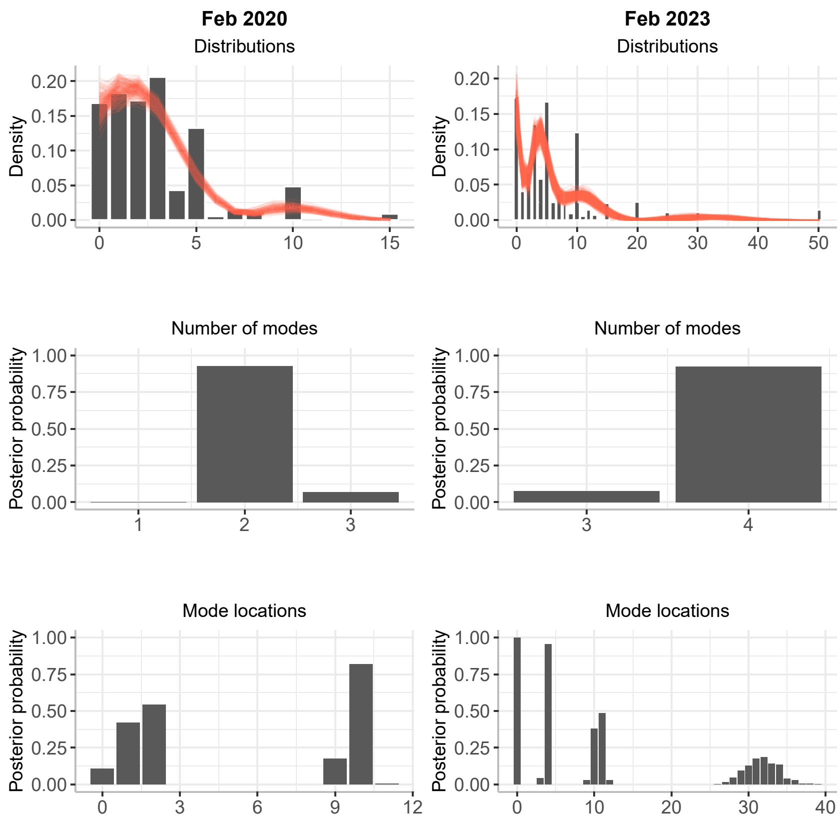

Figure 4: Empirical distribution of the University of Michigan's inflation expectations data and estimated probability mass function (top row), mode locations (center row) and number of modes (bottom row). Source: Cross et al. (2023).

I note that recently there are several other examples where a connection is found between microeconomic models, their data features and the macroeconomic implications. A detailed analysis is beyond the scope of this essay but an important example is from financial econometrics where it has been established that the financial crisis of 2007–2009 was largely due to irresponsible micro behavior of banks by lending subprime mortgages.

I conclude that the existence of multimodality in expectations, financial market conditions and over the business cycle presents an important challenge for research as well as for policy makers like central bankers tasked with maintaining low and stable inflation.

2.2 The Need for Speed: Model Complexity and the Golden Age of Algorithms

I start this discussion with a personal note, for background see Van Dijk (1999). In 1974 Teun Kloek and I started to explore the use of simulation methods in order to compute posterior moments and densities of parameters of interdependent equation systems. We realized that using the Monte Carlo simulation approach as it was known in econometrics required a step of reverse engineering: not generating data given parameters but the reverse. We worked experimentally. That is, we took a specific example of a small econometric model and considered that the posterior, being the product of likelihood and prior, is not a member of a known class of densities in our case even though an informative prior may be chosen as one. Thus, I simulated draws from this prior and evaluated likelihoods. That worked fine for the case of an informative prior that was a close approximation to the likelihood. During the Fall of 1974, I presented the method in an informal seminar at the Econometric Institute in Rotterdam. One of the participants asked whether I could handle the case of a uniform prior. When I ran our computer program for that case after the seminar I discovered that the posterior results were numerically very unstable. Only a few draws got a large weight of the likelihood function. The remaining 99 % of the draws received a negligible weight. A simple step was to generate draws from a distribution with a density that approximates the posterior/likelihood and, at the same time, to divide the posterior/likelihood by this approximate density. I discovered that evening in the library that in the book 'Monte Carlo Methods', Hammersley and Handscomb (1964), this idea was listed and framed as Importance Sampling (IS), due to Goertzel (1949) and Kahn and Harris (1951). Our results appeared in a report of the Econometric Institute (Kloek and Van Dijk 1975), presented at the World Meeting of the Econometric Society in Toronto in 1975, published in Econometrica, Kloek and Van Dijk (1978) and further developed in Van Dijk and Kloek (1980, 1983) and Geweke (1989). An advantage of importance sampling is that draws are generated independently and one can make use of standard limit theorems, like the Law of Large Numbers, to check the accuracy of the numerical results. However, finding a good approximate density in a high dimensional problem was not always easy and successful.

Another very important class of Monte Carlo methods is based on the intuition of a basic Markov property in the sense that in this approach a random draw generated from a candidate distribution depends only upon a previously generated draw but not upon older draws. Also, given the property of so-called *time reversibility* of a Markov Chain method a sample of generated draws will behave as if generated from the target/posterior distribution after a long sequence of accepted random draws. Further, it is intuitively clear that candidate draws are typically rejected in regions of the parameter space where too many candidate draws are simulated as compared with the target/posterior distribution, whereas candidate draws are typically accepted (and repeated) in regions of the parameter space where too few candidate draws are simulated as compared with the target/posterior distribution, see for an introduction Hoogerheide, Van Dijk, and Van Oest (2009). Methods based on these principles were labeled Markov Chain Monte Carlo (MCMC) methods and were developed by Metropolis et al. (1953) and Hastings (1970). The Gibbs sampling method, the most well-known MCMC method, is due to Geman and Geman (1984). These methods were introduced in statistics by Gelfand and Smith (1990). For an historical perspective on the Metropolis-Hastings method, I refer to Hitchcock (2003).

It is important to mention that parallel to the advancement of software the enormous advances in the development of hardware is crucial. For illustrative purposes I include Figure 5 showing the ENIAC, the Electronic Numerical Integrator and Computer, which was the first programmable general purpose digital computer. It is interesting to note that the first part of the name of this computer refers to numerical integration which is the most important operation in modern simulation based Bayesian inference. More important is to realize that there has been a revolutionary development in power and speed of computers since the early beginning. This continues even in recent times, see Figure 6. This development is an important cause for the substantial rise in simulation-based Bayesian econometrics since it allows the analysis of more complex economic issues than could be done with basic linear models.

I note that in both computational approaches, IS and MCMC, the choice of a good approximate distribution, named candidate/importance distribution, to the so-called target distribution, often the posterior, is crucial for more complex model structures. With a poor approximation many draws receive a negligible weight in IS or there are only very few accepted draws in MCMC. In this context it is relevant to realize that in many scientific fields the clock-shapes of the normal or student-t posterior densities are less relevant and multimodal or other non-elliptical shapes like curved ridges and asymmetric tails occur, for instance, the analysis of DNA data in bio-informatics, obtaining loans in the banking sector by heterogeneous groups and analysis of education's effect on earned income in labor economics. Here, I refer, for the sake of convenience, to a few of my own papers with different co-authors: Schaap et al. (2013), Baştürk et al. (2023), Cross et al. (2023) and Baştürk, Hoogerheide, and van Dijk (2017). The appearance of these nonnormal shapes of target densities led to the search for flexible classes of candidate densities with good approximation properties. One approach is to make use of a *mixture* of densities and I discuss briefly how this idea was implemented by several co-authors and myself. Our method, labeled Mixture of t-distributions by Importance Sampling weighted Expectation Maximization (MitISEM) (Baştürk et al. 2017; Baştürk, Hoogerheide, and van Dijk 2017; Hoogerheide, Opschoor, and

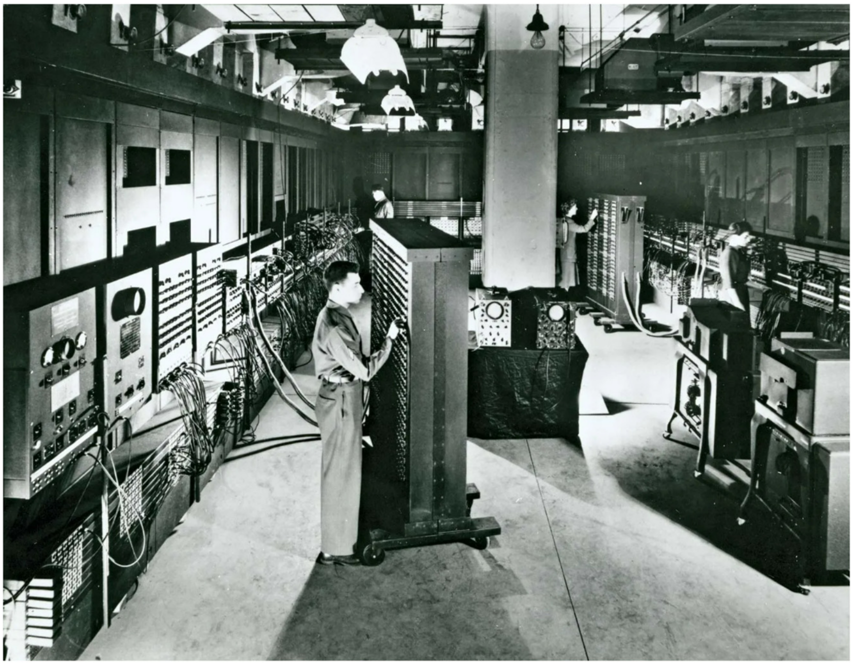

Figure 5: ENIAC (electronic numerical integrator and computer), c. 1946. Courtesy of the Moore School of Electrical Engineering, University of Pennsylvania.

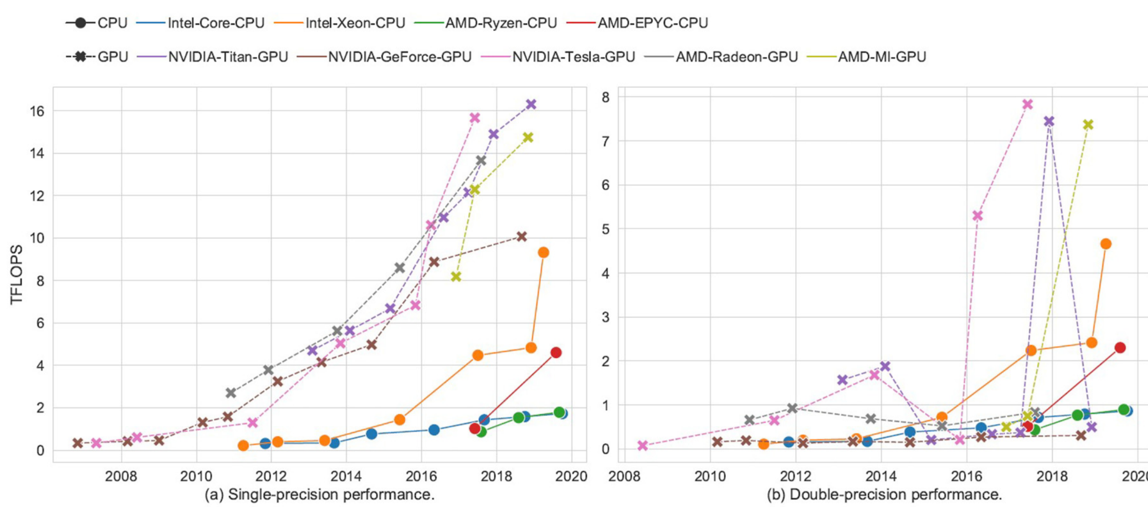

Figure 6: Comparing single-precision and double-precision performance of CPUs and GPUs. Source: Sun et al. (2019).

Van Dijk 2012) provides an automatic and flexible two-stage method to approximate a non-elliptical target density using an adaptive mixture of student-*t* densities as approximating density. In the first stage a mixture of student-*t* densities is fitted to the target using an expectation maximization algorithm where each step of the optimization procedure is weighted using importance sampling. In the second stage this mixture density is a candidate density for efficient and robust application of importance sampling or the Metropolis-Hastings (MH) method to estimate properties of the target distribution. Of course, many other researchers followed the research line of finding a good approximation to a posterior density but discussing this literature is beyond the purpose of this personal essay.

A next step was to have the candidate density to learn over time due to the changing data and model structure. A fundamental step was to introduce dynamic models with unobserved states and parameters that may change over time. For such linear dynamic systems with Gaussian noise, the optimal learning algorithm is the so-called Kalman Filter where the updating of the hidden states occurs using analytical properties of the normal distribution. I list only two papers that introduce this method in econometrics and statistics: Frühwirth-Schnatter (1994) and De Jong and Shephard (1995) but there are several more papers on this topic.

However, the temporal change in the posterior and predictive distributions of complex models with non-Gaussian noise require more sophisticated algorithms than the Normal/Kalman Filter. The nonlinear dynamic adjustments in unobserved states and parameters require simulation methods instead of analytical methods. This is a topic of much recent research, see, for instance, Herbst and Schorfheide (2016) and the references cited there, in particular the seminal paper on Sequential Monte Carlo due to Gordon, Salmond, and Smith (1993).

One example of my personal research is presented in Baştürk et al. (2019). I summarize a novel filtering approach, listed there, as follows. Filters are usually based on a recursive formula about the particles in a filter. Importantly, propagation of these particles over time leads to weight degeneracy with finally only one particle carrying all the weight. In the paper Baştürk et al. (2019) we avoid the propagation step by replacing it by an independent sampling step in each time period. Here we extend the literature about importance sampling for state space models using a very flexible approximation density based on mixtures of student-*t* densities. More research is here needed and is an opportunity for novel work.

In Table 1 a sketch is presented how a large panel of financial time series is used in a Bayesian modeling approach to yield a large set of predictive densities. As a next step a machine learning type of method is used to cluster the large set of densities into a small set of densities. The probabilistic combination weights are unobserved and have to be integrated out using dynamic filtering methods like Normal/Kalman or Particle Filters. I emphasize that the line in the middle can also be interpreted as containing a hidden layer where the elements are integrated out using, for instance, neural networks from machine learning. The figure serves, therefore, to indicate the close connection between machine learning methods with hidden layers and filtering methods in nonlinear time series models.

More generally, machine learning refers to a set of algorithms that are used to find patterns in large data sets and make predictions for future outcomes of these data. They are used, for instance, in stock trading where a trader may be informed about potential interesting future outcomes. What matters is that a similar method is used in the example of Casarin et al. (2023) but with nomenclature from nonlinear filtering. Such a close connection between the fields of machine learning and nonlinear filtering may lead to challenging research.

Table 1: Summary of predictive density combinations with time-varying weight components for large datasets and their connection with machine learning in Casarin et al. (2023).

Large set of n financial series leading to large set of n predictive densities
$\searrow$ $\downarrow$ $\downarrow$ $\swarrow$
Dynamic clustering allocates large set of n densities to small number of m groups
$\downarrow$ $\downarrow$ $\downarrow$ $\downarrow$
Cluster 1 Cluster 2 $\cdots$ Cluster m
Cluster weights are logistic transformations of unrestricted random walks weights:
$\swarrow$ $\downarrow$ $\downarrow$ $\searrow$
Mapping small set of cluster weights to large set of model weights
$\downarrow$ $\downarrow$ $\downarrow$ $\downarrow$
Construction of convex combination oflarge set of predictive densities

I close this Subsection with a remark about the issue *Can machines think?* This quote from Turing (1950) refers to the argument that technically oriented persons are of the opinion that machine learning methods automatically can perform tasks. However, it is experienced in practice and in formal inference that tuning of parameters by humans is often crucial in complex cases. A clear difference may be observed between situations where one controls the whole process and only repetitive tasks have to be performed (for example, building a series of new cars) and situations where one has to deal with humans or animals that may have unexpected characteristics and where unexpected events may happen in the task performance (for example, in healthcare or complex biological processes). Especially in the latter cases the role of human intervention may remain crucial. My personal opinion is that a combination of human and machine learning appears to be more realistic in the complex models handled in empirical econometrics.

As a final note I remark that the methods of this Subsection are needed and applicable in Sections 2.3–2.5.

2.3 Learning About Models, Forecasts and Policies including Their Uncertainty

As a preliminary remark I note that the introduction of stochastic errors in order to learn about the specification of econometric equation systems appeared less than one hundred years ago. Residual error models were added by Tinbergen to his system of equations, see Tinbergen (1939), with the aim to minimize the difference between observed values of data and forecasted values. However, Keynes was not convinced that this was the right approach, see the famous Keynes-Tinbergen debate in Keynes (1939). As a second note Trygve Haavelmo introduced the concept of probabilistic estimation of systems of equations using the likelihood approach which allowed for formal specification testing (Haavelmo 1944).

Econometric learning about model specification, forecasts and policies has come a long way since. It is nontrivial due to a large degree of ambiguity or uncertainty, making absolute conclusions difficult. I discuss three popular methods that tackle this issue of specification learning: error learning, Bayesian learning and, more recently, machine learning used on big data sets. There exists an interesting methodological and even philosophical debate about the relative merits of each approach but there does not exist a theorem about superiority of one of the three approaches in all possible situations. Learning from errors/mistakes through a process of a series of trials/tests is supposed to improve the model specification. Bayesian learning possesses the advantage of being an optimal information processing rule where probabilistic information of prior beliefs about model specification is carried through data evidence to posterior or predictive information about model features like forecasting including their uncertainty. Big data require an agnostic approach using deterministic or stochastic algorithms mostly based on machine learning that detect patterns in massive data before a modeling stage can occur. In this essay I take the point of view of a 'do-er' or instrumentalist. Handling a case of big data needs first a diagnostic step of pattern recognition; next a Bayesian learning step and finally learning from residual errors. The connections between the three procedures are sketched in Figure 7.

As mentioned learning about patterns in vast amounts of data that are too large for traditional analytical approaches requires machine learning methods, see also Section 2.2. The hope is that by exploring such datasets meaningful patterns will emerge with greater clarity, allowing researchers to identify relationships that might not be evident with smaller datasets. Many of the microdatasets used in the research outlined in Section 2.1 are examples of big data. In such cases, the immense quantity of data often means that there will be inherent noise and uncertainty. The challenge lies in distinguishing genuine patterns from random noise. To that end, big data learning has an intersection with Bayesian learning, in that prior information can be used as a regularization tool to shrink meaningless noise to zero while preserving important information (Bańbura, Giannone, and Reichlin 2010). Alternatively, advanced algorithms, often based in machine learning, have been increasingly common (Goulet Coulombe 2020; Goulet Coulombe et al. 2022). Varian (2014) provides a broad discussion of big data in econometrics while Koop (2017) provides a discussion about big data in macroeconomics. For an interesting intersection between probabilistic machine learning and bayesian nonparametric methods, see Chamberlain and Imbens (2003).

A major advantage of Bayesian learning is that it explicitly accounts for uncertainty by going beyond point estimates towards posterior distributions that depict the entire range and likelihood of possible outcomes.

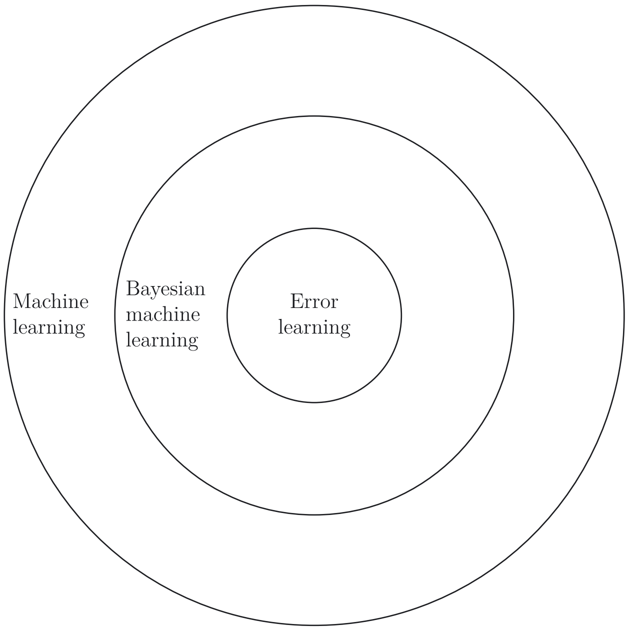

Figure 7: Connections between three learning procedures.

Another advantage is that Bayesian learning is logically coherent in that information from distinct models can be easily combined using Bayes' theorem – a method known as Bayesian Model Averaging (BMA). Set in this manner, the BMA predictive density is a weighted average of individual model densities where the combination weights are equal to posterior probabilities of the respective models. By doing this iteratively, BMA weights encompass learning in that they reflect each model's relative ability to predict the object of interest over the training period. While BMA is simple and logically coherent it has several shortcomings: (1) it assumes that the true model is included in the model set; (2) it does not account for uncertainty associated with the weights attached to each model; (3) it is extremely sensitive to prior information. Given that there does not exist a true model and that therefore any one model is always incorrect/misspecified, it is better to make use of a set of models and of probabilistic learning about features of a set of models. However, even a desired model set is likely to change over time, see Section 2.4. Therefore, a combination of Bayesian learning about model features and learning about posterior and predictive errors seems appropriate. Quantification of uncertainty and risk in this context is important, and it is precisely this rationale that makes Bayesian methods appealing. This has led to an evolution of different types of Bayesian forecast density combinations that aim to address these weaknesses (see Aastveit et al. (2019) for a review).

Error learning is predicated on the idea that prediction errors can be valuable sources of information. It involves adjusting parameter estimates, model selection, or policy related decision-making, based on the size and sign of previous and current errors. The focus here is on using historical errors to minimize discrepancies over iterations thereby achieving improvements through trial and error. The classic econometric example of this type of learning is the error correction model (ECM) of Engle and Granger (1987). Other examples which have become popular since the onset of deep learning are neural networks, especially during the backpropagation phase, and reinforcement learning where agents improve strategies based on rewards and penalties. On another personal note, my first research into neural networks for economic data was in the 1990s (Draisma, Kaashoek, and Van Dijk 1995; Kaashoek and Van Dijk 1994), however it was not until later that I realized that these methods were also useful for microeconometric identification (Hoogerheide, Kaashoek, and Van Dijk 2007). Also, while Bayesian neural networks have been around since the 1980s (Kononenko 1989), they have only recently emerged as an important tool in econometrics (Klein, Smith, and Nott 2023; Tsionas, Parmeter, and Zelenyuk 2023).

In Figure 7 it is indicated that machine learning, be it stochastic or deterministic, encompasses the other two learning procedures. Clearly, Bayesian learning is a very important class of stochastic machine learning in the context of improving model specification and forecasting in the face of uncertainty and risk. Error learning is situated in the inner circle. A quote from Nobel Laureate Clive Granger illustrates the connection between the latter two learning methods as follows: “Bayesian forecasts are better than non-Bayesian Forecasts (here error based forecasts) and better than bad Bayesian Forecasts.” This to indicate that Bayesian learning is wider than trial and error learning but that learning from errors is also an essential step in model forecasting and policy analysis.

For illustrative purposes I summarize features of three recent papers of mine: (1) seeks to address the issue of forecast uncertainty and risk with a combination of a set of models when the data exhibit several shocks, see for details Aastveit, Cross, and van Dijk (2023); (2) learns about model set composition given a large financial data panel, see Casarin et al. (2023); and (3) learns about policy combinations, their uncertainty and risk, in a portfolio strategy, see Baştürk et al. (2019).

For time series which show shock behavior like the real price of oil, see Figure 8, Aastveit, Cross, and van Dijk (2023) make use of a combination of a set of five models and an algorithm which is an econometric interpretation of the so-called Bayesian Predictive Synthesis method, see McAlinn and West (2019) and McAlinn et al. (2020), in order to provide accurate forecasts and risk measures for this oil price. Underestimation of risk could obviously cause immense problems for banks and other participants in financial markets (e.g. bankruptcy). Overestimation of risk may cause one to allocate too much capital as a cushion for risk exposures, possibly having a negative effect on productivity. Therefore, precise estimates of risk measures are obviously desirable.

Since the focus of this Subsection is on learning, Figure 9 shows results on learning about the time pattern of five individual model weights. It is seen that individual model forecasts lead to placing more weight on two models that perform well in forecasting previous periods and vice versa. So, the learning about time variation in these weights is an important source of information. The attentive reader will note that the model combination weights are not restricted to be a convex combination in the unit interval but can take on values over the real line. While this approach sacrifices the intuitive probabilistic understanding of the weights, accommodating both positive and negative weights can provide a safeguard against possible forecast uncertainties by indicating hedging opportunities, an attribute highly valued by many finance professionals.

The second illustration deals with a panel of 496 daily individual stock prices, components of the S&P500, over the sample January 2, 2014 to June 30, 2021. A first step is to make use of a stochastic machine learning

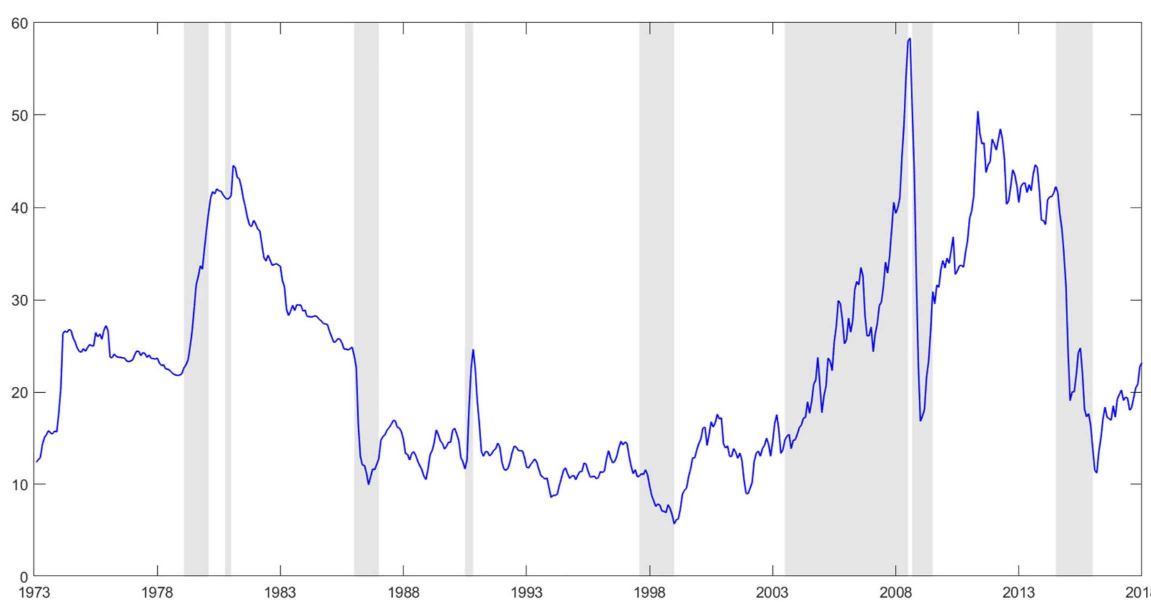

Figure 8: Real price of crude oil. Source: Aastveit, Cross, and van Dijk (2023).

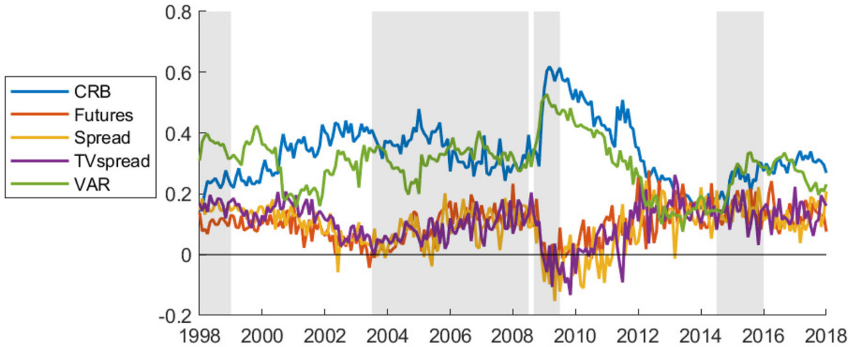

Figure 9: Model weights sequentially computed at each point in time over the forecast evaluation period 1998:03 – 2017:12. Source: Aastveit, Cross, and van Dijk (2023).

type algorithm which clusters the predictions of all stock prices in four groups. In Figure 10 the four clusters of predictive models are denoted by $n1$ and $n2$, which refer to two normalmodels (with high and low volatility, respectively), and $t1$ and $t2$, which refer to two student-t models (with low and high degrees of freedom parameters, respectively). Details are given in Casarin et al. (2023). It is seen in the top panel of Figure 10 that the percentages of stocks within models $n1$ and $t1$ are dominant. COVID-19 creates some instability in the stock allocation. In this case of financial econometrics an important motivation to obtain complete forecast distributions over outcomes is that such results provide information helpful for making economic decisions under uncertainty. I note that asset allocation decisions usually involve higher moments than just first moments.

Next, in the bottom panel it is seen that measures of model set incompleteness indicate that all models fail in the beginning of the pandemic period but the incompleteness lowers again substantially in the later period. Incompleteness is measured as the average value of the squared posterior residuals. It is seen that $n2$ has high average incompleteness after the pandemic. This diagnostic information indicates that cluster $n2$ gives low predictive accuracy in that period.

Results on learning about policy combinations in a portfolio analysis are taken from Baştürk et al. (2019). In this paper a dynamic asset-allocation model is specified in probabilistic terms with portfolio strategies based on momentum patterns in US industry returns. Figure 11 shows the posterior means of the probabilistic weights from a combination of 2 investment strategies (model momentum in blue, residual momentum in yellow). The time variation in these weights – where residual momentum appears to get larger weights in the second half of the data series – provides useful signals for improved modeling and policy, in particular, from a risk-management perspective.

Again as a final note I remark that the topic of this Subsection is closely connected to the material and topic of Section 2.1.

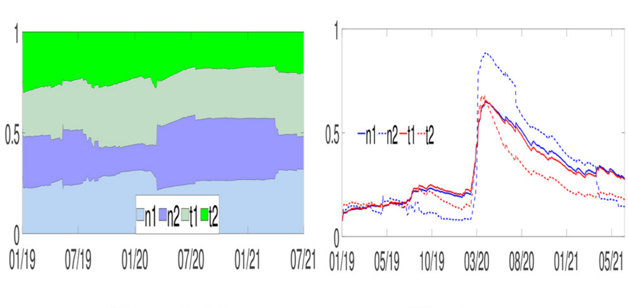

Figure 10: Dynamic weight learning (top) and error learning (bottom). Source: Casarin et al. (2023).

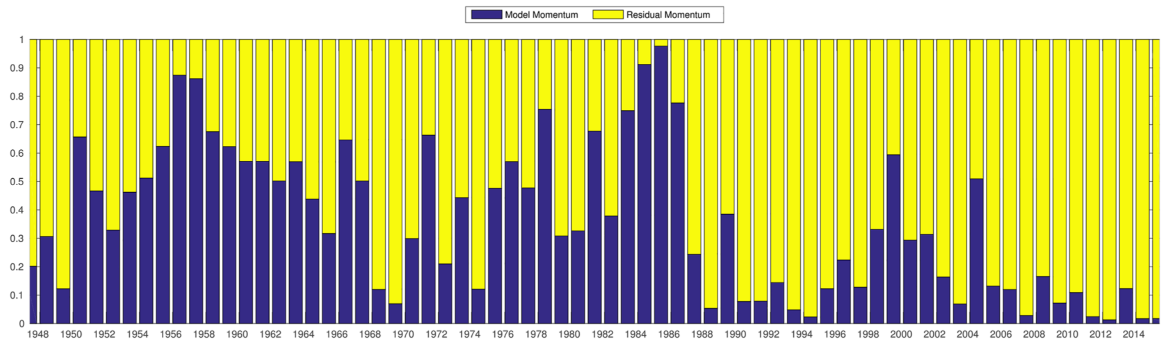

Figure 11: Posterior means of the combination weights for two investment strategies.

2.4 Temporal Distributional Change Due to Polarisation, Imbalances and Shocks

Why is it that, in capitalistic economies, aggregate variables undergo repeated fluctuations about trend, all of essentially the same character? (Lucas 1977)

This topic has fascinated econometric researchers from the beginning of business cycle analysis. A brief selection of the approaches is the following: Juglar (1862) ascribed the recurrent business crises in Europe and North America to credit crises; Yule (1971) and Slutzky (1937) suggested that the cumulative effect of *random shocks* could be the cause of cyclical patterns in economic variables; and Tinbergen explored the possible *economic causes* of the periodic upswings and downswings in economic activity, see Tinbergen (1939).

Next to this topic of exploring patterns in economic time series, several variants of regression analysis were used to estimate stable ratios like the consumption-income ratio. An example using US data is provided in

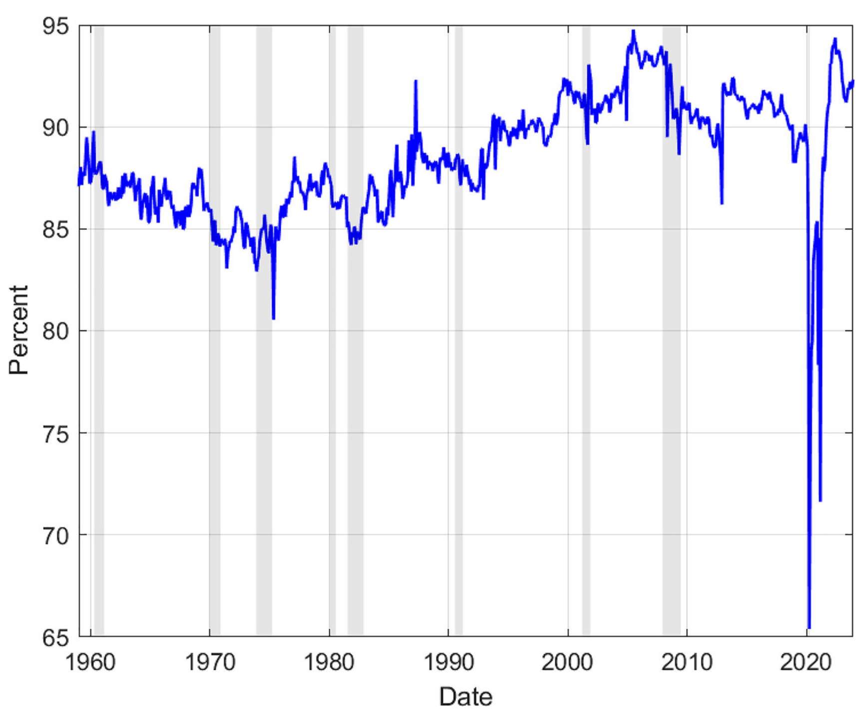

Figure 12: Consumption to income ratio in the US. Source: fred.stlouisfed.org.

Figure 12. This ratio has long been studied in econometrics (e.g. Haavelmo 1947), and has recently been used as an important statistic when modeling income distribution dynamics (Carroll et al. 2017).

Recognizing issues of endogeneity in traditional econometric regressions, Sims confronted the modeling of structural analysis in the latter part of the twentieth century by arguing that data patterns depend to a very large extent only on time series behavior of stochastic economic variables (e.g. Sims 1980; Sims et al. 1986). Lagged dependent variables and shocks were crucial for the dynamic behavior. Again a major line of research to explore stable relations in this context became known as the field of cointegration (Engle and Granger 1987; Granger 1981; Granger 2010).

In the twenty first century new phenomena occurred: A switch from stable to unstable ratios due to polarisation and economic inequality. This is illustrated in Figure 13. The figure shows clear evidence of diverging wealth between different groups since the 1980s. This creates a serious econometric challenge of how to model the connection between temporal variation in cross-sectional distributions and macroeconomic phenomena and it provides also a clear motivation for modeling with quantiles.

To further illustrate how recent imbalances lead to inequality and polarisation, Figure 14 shows the change in property price to income ratios in Western Europe. This leads to major reductions in affordable housing around the region. In 2013, only two cities (Paris and Nice) had ratios of more than 10. Ten years later, 16 cities now have ratios of more than 10, with Paris now having a ratio of more than 20, and cities like Innsbruck and Munich having ratios of more than 15. This phenomenon is not local to Western Europe. The data also show stark increases in such ratios in major cities of Asia, America, and the Oceanic regions, making it a truly global phenomenon.

Another novel feature is that the time series are not always periodical and shocks to series occur. This was discussed in Section 2.3 with reference to (Aastveit, Cross, and van Dijk 2023).

Related to much of the preceding discussion is the general question of how to model temporal distributional changes. To illustrate one such example, I include Figure 15 which shows the distributional changes in GDP per capita in an unbalanced panel of countries from 1950 to 2004, see Khaled and van Dijk (2008).

In a recent piece of work, Chang, Chen, and Schorfheide (2021) propose to accomplish this task using a state-space model in which the state-transition equation is specified as a functional vector autoregression that models

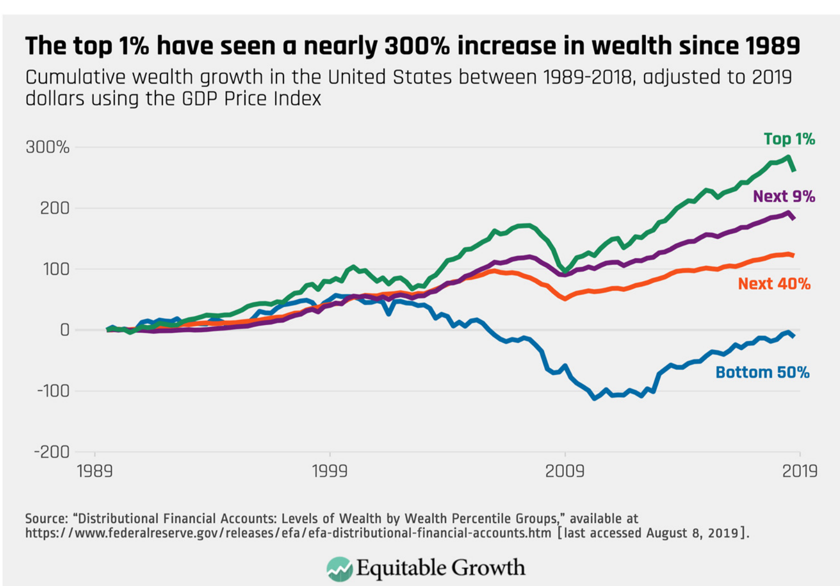

Figure 13: Diverging ratios in economic variables may lead to regime changes.

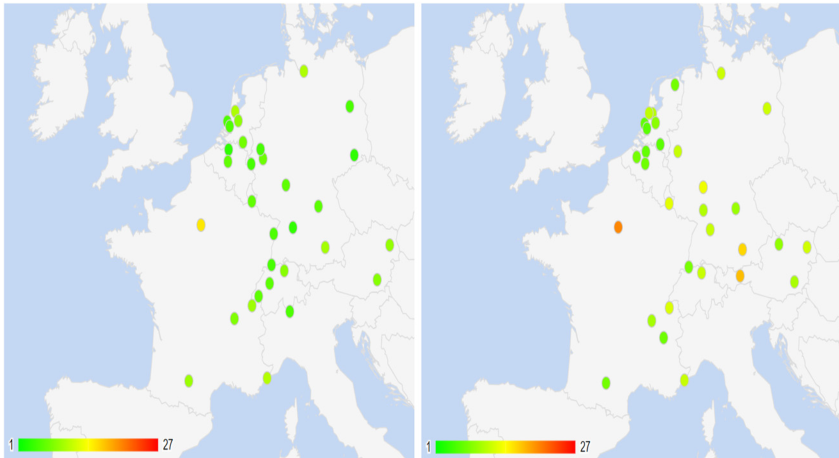

Figure 14: Property price to income ratios by city in Western Europe. Left shows data in 2013 and right shows data in 2023. Source: www.numbeo.com.

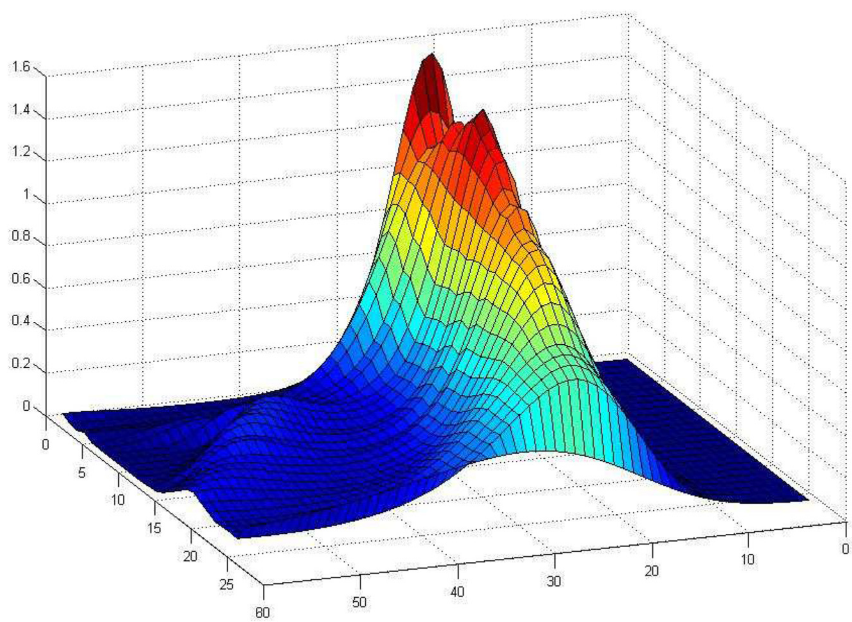

Figure 15: Distributional changes in GDP per capita in an unbalanced panel of countries from 1950 to 2004.

dynamics between macroeconomic aggregates and a cross-sectional density. In this model, the log densities and state transition kernels are approximated by spline basis functions (AKA sieves).

An alternative approach is recently provided by Bjørnland, Chang, and Cross (2023). Instead of using spline basis functions, they propose the use of functional principle component analysis (FPCA) to project the infinite dimensional density of interest to a finite set of functional principle components. These components are then used within a SVAR model that is estimated using frequentist estimation.

From a practical perspective, both the frameworks provided by Chang, Chen, and Schorfheide (2021) and Bjørnland, Chang, and Cross (2023) allow econometricians to jointly examine the distributional effects of functional cross-sections and aggregate macroeconomic time series for the first time. While this research is still in its infancy, given the desire for such methods in policy areas across multiple domains, I expect this to be a major area of theoretical and empirical research within the next two decades. Note the connection between these topics and the algorithms listed in Section 2.2.

2.5 Climate Change and the Macroeconomy

Figure 16 shows the annual quantity of global carbon dioxide (CO_2) emissions from fossil fuels and industry since the eighteenth century. Emissions prior to the Industrial Revolution were very low, and have since grown exponentially. While the growth in emissions took a dip in the COVID-19 pandemic, they have since increased to a new peak of over 35 billion tonnes each year.

While Figure 16 shows a clear link between climate change and the macroeconomy, this nexus is extremely complex, with numerous interrelated aspects. From an economic perspective, greenhouse gas (GHG) emissions are negative externalities that result from a market failure in the overuse of GHG-emitting technologies. It is well known that such externalities can be corrected through market interventions on the price, via taxation, or the quantity, via quotas and carbon trading systems (see, e.g. Stern (2008) for a literature review of the economics of climate change). Climate change therefore has both a direct effect on the economy through negative externality and indirect effects through policies.

The early work on this nexus focused on the direct long-run relationship between climate change and the economy, and it can be traced back to the twentieth century (see, e.g. Tol (2009) for a literature review). Much of this work was due to the pioneering research by Nordhaus, who was awarded the Nobel Prize for it: "integrating climate change into long-run macroeconomic analysis". His methodological contributions were in the popular classes of linear programming (LP), or computable general equilibrium (CGE) models, with his most well known model being the Dynamic Integrated model of Climate and the Economy (DICE) model (see, Nordhaus (2018) for a review). Another example of the effect of climate change on the long term distribution of international economic poverty is the expected evolution of drought by regions in Africa, see Figure 17. While most work on the link between weather conditions and economic conditions has been done using frequentist methodologies

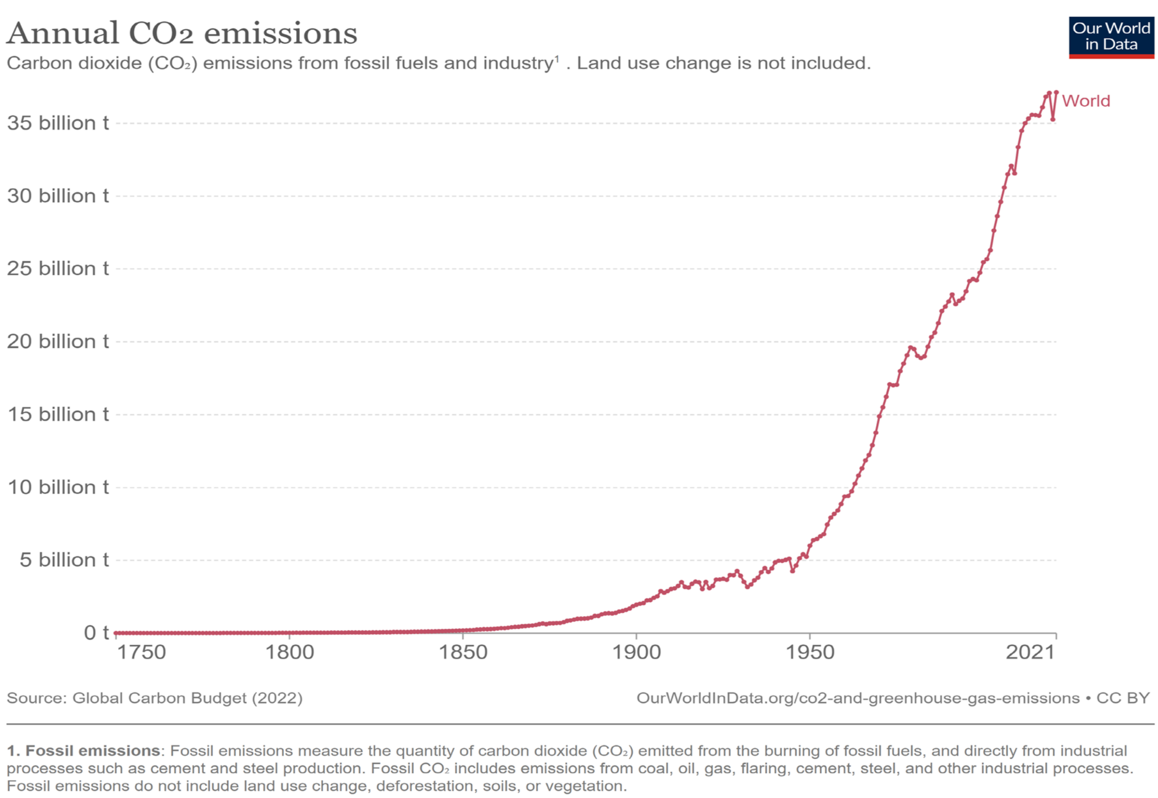

Figure 16: Global CO2 emissions from fossil fuels. Source: ourworldindata.org.

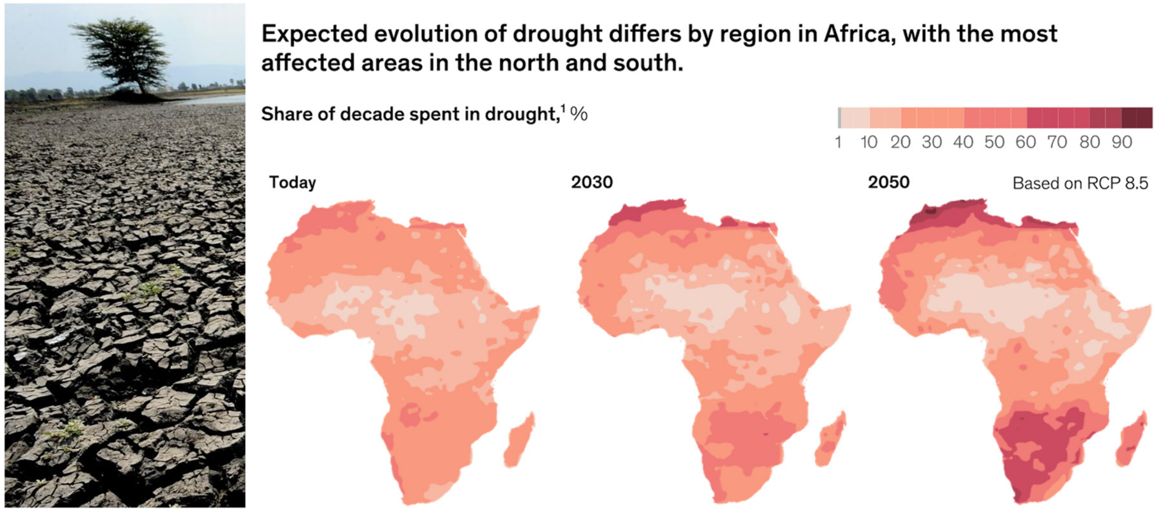

Figure 17: Example: the effect of climate change on long term changes in distribution of international economic poverty.

(e.g. Arezki and Brückner 2012; Brückner and Ciccone 2011), there is still much scope to tackle such problems using the conditional probabilistic lens that Bayesian econometrics has to offer.

In recent years, empirical macroeconomists have shifted their focus towards the short-run nexus between the economy and climate. Much of it has been with different variants of the (Bayesian) structural vector autoregression (SVAR) model. For instance, Alessandri and Mumtaz (2021) use a panel VAR model with stochastic volatility in mean (SVM) to study the impact of climate volatility on economic growth on 133 countries between 1960 and 2019. They find that increases in temperature volatility have a negative impact on GDP growth, in both rich and poor countries. Känzig (2023) estimates a proxy VAR with an external instrument to show that carbon pricing mechanisms in the European Union (EU) decrease emissions, but at the expense of lower real economic activity and greater inequality; as poorer households lower their consumption significantly more than richer households. More recently, Bjørnland, Chang, and Cross (2023) propose an SVAR model with sign restrictions to quantify the effects of demand and supply shocks underlying the EU carbon market covered by the Emission Trading System (ETS). They find that while emission supply restrictions of the EU ETS were the dominant driver of emissions reductions since its inception in 2005, two opposing emission demand factors that reflect industrial production and the transition towards a low-carbon economy have also played an important role. Given the recency of these studies, I expect to see much more work done on the climate-macroeconomy nexus using Bayesian SVARs over the next decade.

2.6 Widespread, Accessible, Advanced High-Level Training

The discussion thus far demonstrates that proficiency with Bayesian methods requires a high degree of education. While most degrees in economics or finance offer an array of frequentist econometric methods, Bayesian methods are unfortunately still not a core component of most curricula around the world. This should change. While some universities have begun to offer a single elective course in Bayesian methods at either an undergraduate or postgraduate level, the sophisticated methods discussed here demonstrate the need for increased offerings in order to move forward. Here, I list some thoughts on a possible curriculum starting at a basic level.

In the first instance, students should be taught the foundations of probability and statistics. Most universities provide this in the first year of a bachelor degree. Given this foundation, students could then engage in an introduction to a Bayesian econometrics course. This course should be accessible to anyone with a good foundation in practical probability. A first lecture could discuss philosophical differences between Bayesian and frequentist methods, and demonstrate how Bayesian estimation yields posterior probabilities that are intuitively interpreted with the notion of conditional probability. Next, analytical solutions to problems such as linear regression can be taught, and a parallel can be drawn between the OLS estimator and the Bayesian posterior mean estimate. Important here is the fact that the prior can be seen as a regularization tool within the data-driven posterior mean. Next, regression with non-normal error distributions, such as the student-t-distribution, provide a natural shift away from analytical solutions of obtaining a posterior distribution towards using numerical Monte Carlo methods for this purpose. Basic algorithms such as Gibbs Sampling, see Geman and Geman (1984) and the Metropolis-Hastings method, see Metropolis et al. (1953) can be readily introduced. Mastering such methods will provide a solid foundation to go into advanced time series methods, such as vector autoregression, panel methods, and both linear and non-linear state-space models. At the undergraduate level, the precision sampling method of Chan and Jeliazkov (2009) will be especially convenient, see Chan and Strachan (2023) for a recent survey of these methods. At postgraduate levels, drawing parallels between precision sampling and linear and non-linear filters would prove valuable. Also at the postgraduate level, alternative posterior approximation methods such as Sequential Importance Sampling, Variational Bayes, and Approximate Bayesian Computation (ABC) would round out a comprehensive treatise of simulation based Bayesian methods. Challenges when using such methods should be discussed. For instance, convergence problems when implementing variational Bayes and ABC methods, or the path degeneracy problem when using particle filters. This will spark interest in curious students and act as an open challenge to provide new breakthroughs in their PhD dissertations and beyond.

3 Final Remark

The aim of this essay was to discuss six specific topics that one may classify under the general topic *Bayesian econometrics with a purpose*. It is the author's wish and hope that twenty first century Bayesian econometricians take up (some of) the challenges discussed.

References

Aastveit, K. A., J. L. Cross, and H. K. van Dijk. 2023. "Quantifying Time-Varying Forecast Uncertainty and Risk for the Real Price of Oil." *Journal of Business & Economic Statistics* 41 (2): 523–37.

Aastveit, K. A., J. Mitchell, F. Ravazzolo, and H. K. van Dijk. 2019. "The Evolution of Forecast Density Combinations in Economics." In *Oxford Research Encyclopedia of Economics and Finance*. Oxford: Oxford University Press.

Adrian, T., N. Boyarchenko, and D. Giannone. 2019. "Vulnerable Growth." *The American Economic Review* 109 (4): 1263–89.

Adrian, T., N. Boyarchenko, and D. Giannone. 2021. "Multimodality in Macrofinancial Dynamics." *International Economic Review* 62 (2): 861–86.

Alessandri, P., and H. Mumtaz. 2021. "The Macroeconomic Cost of Climate Volatility." *arXiv preprint arXiv:2108.01617*.

Arezki, R., and M. Brückner. 2012. "Rainfall, Financial Development, and Remittances: Evidence from Sub-Saharan Africa." *Journal of International Economics* 87 (2): 377–85.

Arias, J. E., J. F. Rubio-Ramirez, and D. F. Waggoner. 2021. "Inference in Bayesian Proxy-SVARs." *Journal of Econometrics* 225 (1): 88–106.

Bańbura, M., D. Giannone, and L. Reichlin. 2010. "Large Bayesian Vector Auto Regressions." *Journal of Applied Econometrics* 25 (1): 71–92.

Baştürk, N., A. Borowska, S. Grassi, L. Hoogerheide, and H. K. van Dijk. 2019. "Forecast Density Combinations of Dynamic Models and Data Driven Portfolio Strategies." *Journal of Econometrics* 210 (1): 170–86.

Baştürk, N., C. Cakmakli, P. Ceyhan, and H. K. van Dijk. 2014. "On the Rise of Bayesian Econometrics After Cowles Foundation Monographs 10, 14." *Oeconomia* 4 (3): 381–447.

Baştürk, N., J. Cross, P. de Knijff, L. Hoogerheide, P. Labonne, and H. K. van Dijk. 2023. "Bayesmultimode: Bayesian Mode Inference in R." Technical report, Tinbergen Institute Discussion Paper.

Baştürk, N., S. Grassi, L. Hoogerheide, A. Opschoor, and H. K. van Dijk. 2017. "The R Package MitISEM: Efficient and Robust Simulation Procedures for Bayesian Inference." *Journal of Statistical Software* 79: 1–40.

Baştürk, N., L. Hoogerheide, and H. K. van Dijk. 2017. "Bayesian Analysis of Boundary and Near-Boundary Evidence in Econometric Models with Reduced Rank." *Bayesian Analysis* 12 (3): 879–917.

Bjørnland, H. C., Y. Chang, and J. Cross. 2023. "Oil and the Stock Market Revisited: A Mixed Functional VAR Approach." Technical report. Indiana University.

Brückner, M., and A. Ciccone. 2011. "Rain and the Democratic Window of Opportunity." *Econometrica* 79 (3): 923–47.

Caldara, D., and E. Herbst. 2019. "Monetary Policy, Real Activity, and Credit Spreads: Evidence from Bayesian Proxy SVARs." *American Economic Journal: Macroeconomics* 11 (1): 157–92.

Carroll, C., J. Slacalek, K. Tokuoka, and M. N. White. 2017. "The Distribution of Wealth and the Marginal Propensity to Consume." *Quantitative Economics* 8 (3): 977–1020.

Casarin, R., S. Grassi, F. Ravazzolo, and H. K. van Dijk. 2023. "A Flexible Predictive Density Combination for Large Financial Data Sets in Regular and Crisis Periods." *Journal of Econometrics* 237 (2): 105370.

Chamberlain, G., and G. W. Imbens. 2003. "Nonparametric Applications of Bayesian Inference." *Journal of Business & Economic Statistics* 21 (1): 12–8.

Chan, J. C., and I. Jeliazkov. 2009. "Efficient Simulation and Integrated Likelihood Estimation in State Space Models." *International Journal of Mathematical Modelling and Numerical Optimisation* 1 (1–2): 101–20.

Chan, J. C., and R. W. Strachan. 2023. "Bayesian State Space Models in Macroeconometrics." *Journal of Economic Surveys* 37 (1): 58–75.

Chang, M., X. Chen, and F. Schorfheide. 2021. "Heterogeneity and Aggregate Fluctuations." Technical report. National Bureau of Economic Research.

Cooley, T. F., and S. F. LeRoy. 1985. "Atheoretical Macroeconometrics: A Critique." *Journal of Monetary Economics* 16 (3): 283–308.

Cross, J., L. Hoogerheide, P. Labonne, and H. K. van Dijk. 2023. "Bayesian Mode Inference for Discrete Distributions in Economics and Finance." Technical report. Tinbergen Institute Discussion Paper.

De Jong, P., and N. Shephard. 1995. "The Simulation Smoother for Time Series Models." *Biometrika* 82 (2): 339–50.

Draisma, G., J. F. Kaashoek, and H. K. Van Dijk. 1995. "A Neural Network Applied to Embedded Economic Data." Discussion Paper, 95–20. Tinbergen Institute.

Engle, R. F., and C. W. Granger. 1987. "Co-Integration and Error Correction: Representation, Estimation, and Testing." *Econometrica: Journal of the Econometric Society* 55 (2): 251–76.

Frühwirth-Schnatter, S. 1994. "Data Augmentation and Dynamic Linear Models." *Journal of Time Series Analysis* 15 (2): 183–202.

Gelfand, A. E., and A. F. M. Smith. 1990. "The Beginning of the Monte Carlo Method." *Journal of the American Statistical Association* 85: 398–409.

Geman, S., and D. Geman. 1984. "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images." *IEEE Transactions on Pattern Analysis and Machine Intelligence* PAMI-6 (6): 721–41.

Geweke, J. 1989. "Bayesian Inference in Econometric Models Using Monte Carlo Integration." *Econometrica* 57 (6): 1317–39.

Giacomini, R., T. Kitagawa, and M. Read. 2022. "Robust Bayesian Inference in Proxy SVARs." *Journal of Econometrics* 228 (1): 107–26.

Goertzel, G. 1949. "Quota Sampling and Importance Functions in Stochastic Solution of Particle Problems." *Technical Report*.

Gordon, N. J., D. J. Salmond, and A. F. Smith. 1993. "Novel Approach to Nonlinear/Non-Gaussian Bayesian State Estimation." 140 (2): 107–13.

Goulet Coulombe, P. 2020. "The Macroeconomy as a Random Forest." *SSRN 3633110*.

Goulet Coulombe, P., M. Leroux, D. Stevanovic, and S. Surprenant. 2022. "How Is Machine Learning Useful for Macroeconomic Forecasting?" *Journal of Applied Econometrics* 37 (5): 920–64.

Granger, C. W. 1981. "Some Properties of Time Series Data and Their Use in Econometric Model Specification." *Journal of Econometrics* 16 (1): 121–30.

Granger, C. W. 2010. "Some Thoughts on the Development of Cointegration." *Journal of Econometrics* 158 (1): 3–6.

Haavelmo, T. 1944. "The Probability Approach in Econometrics." *Econometrica* 12: iii-vi+1-115.

Haavelmo, T. 1947. "Methods of Measuring the Marginal Propensity to Consume." *Journal of the American Statistical Association* 42 (237): 105–22.

Hammersley, J., and D. Handscomb. 1964. *Monte Carlo Methods*. London: Methuen & Co.

Hastings, W. K. 1970. "Monte Carlo Sampling Methods Using Markov Chains and Their Applications." *Biometrika* 57 (1): 97–109.

Heckman, J. J. 2001. "Micro Data, Heterogeneity, and the Evaluation of Public Policy: Nobel Lecture." *Journal of Political Economy* 109 (4): 673–748.

Herbst, E. P., and F. Schorfheide. 2016. *Bayesian Estimation of DSGE Models*. Princeton: Princeton University Press.

Hitchcock, D. B. 2003. "A History of the Metropolis-Hastings Algorithm." *The American Statistician* 57 (4): 254–7.

Hoogerheide, L. F., J. F. Kaashoek, and H. K. Van Dijk. 2007. "On the Shape of Posterior Densities and Credible Sets in Instrumental Variable Regression Models with Reduced Rank: An Application of Flexible Sampling Methods Using Neural Networks." *Journal of Econometrics* 139 (1): 154–80.

Hoogerheide, L., A. Opschoor, and H. K. Van Dijk. 2012. "A Class of Adaptive Importance Sampling Weighted EM Algorithms for Efficient and Robust Posterior and Predictive Simulation." *Journal of Econometrics* 171 (2): 101–20.

Hoogerheide, L. F., H. K. Van Dijk, and R. D. Van Oest. 2009. “Handbook of Computational Econometrics.” In *Simulation Based Bayesian Econometric Inference: Principles and Some Recent Computational Advances*, 215—80. New Jersey, USA: Wiley & Sons.
Juglar, C. 1862. *Des Crises Commerciales et de Leur Retour Périodique en France, en Angleterre et aux Etats-Unies*. Paris: Guillaumin.
Kaashoek, J. F., and H. K. Van Dijk. 1994. “A Neural’ Network Applied to Tlie Calculation of Lyapunov Exponents.” *Econometric Reviews* 13 (1): 123—37.
Kahn, H., and T. E. Harris. 1951. “Estimation of Particle Transmission by Random Sampling.” In *National Bureau of Standards Applied Mathematics Series*, Vol. 12, 27—30.
Känzig, D. R. 2023. “The Unequal Economic Consequences of Carbon Pricing.” Technical report. National Bureau of Economic Research.
Keynes, J. M. 1939. “Professor Tinbergen’s Method.” *The Economic Journal* 49 (195): 558—77.
Khaled, M., and H. K. van Dijk. 2008. “Distributional Dynamics Using Quartic-Based State-Space Models.” Technical report.
Klein, N., M. S. Smith, and D. J. Nott. 2023. “Deep Distributional Time Series Models and the Probabilistic Forecasting of Intraday Electricity Prices.” *Journal of Applied Econometrics* 38 (4): 493—511.
Kloek, T., and H. K. Van Dijk. 1975. “Bayesian Estimates of Equation System Parameters: An Unorthodox Application of Monte Carlo.” Technical Report 7511/E.
Kloek, T., and H. K. Van Dijk. 1978. “Bayesian Estimates of Equation System Parameters: An Application of Integration by Monte Carlo.” *Econometrica* 46 (1): 1—19.
Kononenko, I. 1989. “Bayesian Neural Networks.” *Biological Cybernetics* 61 (5): 361—70.
Koop, G. 2017. “Bayesian Methods for Empirical Macroeconomics with Big Data.” *Review of Economic Analysis* 9 (1): 33—56.
Lucas, R. 1977. “Understanding Business Cycles.” *Understanding Business Cycles* 5 (1): 7—29.
McAlinn, K., and M. West. 2019. “Dynamic Bayesian Predictive Synthesis in Time Series Forecasting.” *Journal of Econometrics* 210 (1): 155—69.
McAlinn, K., K. A. Aastveit, J. Nakajima, and M. West. 2020. “Multivariate Bayesian Predictive Synthesis in Macroeconomic Forecasting.” *Journal of the American Statistical Association* 115 (531): 1092—110.
Mertens, K., and M. O. Ravn. 2013. “The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States.” *The American Economic Review* 103 (4): 1212—47.
Metropolis, N. 1987. “The Beginning of the Monte Carlo Method.” *Los Alamos Science* 125—30.
Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller. 1953. “Equation of State Calculations by Fast Computing Machines.” *The Journal of Chemical Physics* 21 (6): 1087—92.
Mian, A., and A. Sufi. 2011. “House Prices, Home Equity-Based Borrowing, and the US Household Leverage Crisis.” *The American Economic Review* 101 (5): 2132—56.
Mitchell, J., A. Poon, and D. Zhu. 2022. “Constructing Density Forecasts from Quantile Regressions: Multimodality in Macro-Financial Dynamics.” FRB of Cleveland Working Paper No. 22-12, 2022R.
Mumtaz, H., and K. Petrova. 2023. “Changing Impact of Shocks: A Time-Varying Proxy SVAR Approach.” *Journal of Money, Credit, and Banking* 55 (2—3): 635—54.
Nordhaus, W. 2018. “Evolution of Modeling of the Economics of Global Warming: Changes in the DICE Model, 1992—2017.” *Climatic Change* 148 (4): 623—40.
Ramey, V. A. 2011. “Can Government Purchases Stimulate the Economy?” *Journal of Economic Literature* 49 (3): 673—85.
Schaap, M., R. J. L. F. Lemmers, R. Maassen, P. J. Van der Vliet, L. F. Hoogerheide, H. K. Van Dijk, N. Baştürk, P. De Knijff, and S. M. Van der Maarel. 2013. “Genome-Wide Analysis of Macrosatellite Repeat Copy Number Variation in Worldwide Populations: Evidence for Differences and Commonalities in Size Distributions and Size Restrictions.” *BMC Genomics* 14: 1—12.
Sims, C. A. 1980. “Macroeconomics and Reality.” *Econometrica: Journal of the Econometric Society* 48 (1): 1—48.
Sims, C. A. 2007. “Bayesian Methods in Applied Econometrics, or, Why Econometrics Should Always and Everywhere be Bayesian.” *Hotelling lecture, presented* June 29: 2007.
Sims, C. A. 1986. “Are Forecasting Models Usable for Policy Analysis?” *Quarterly Review* 10: 2—16.
Slutzky, E. 1937. “The Summation of Random Causes as the Source of Cyclic Processes.” *Econometrica: Journal of the Econometric Society* 5 (2): 105—46.
Stern, N. 2008. “The Economics of Climate Change.” *The American Economic Review* 98 (2): 1—37.
Sun, Y., N. B. Agostini, S. Dong, and D. Kaeli. 2019. “Summarizing CPU and GPU Design Trends with Product Data.” *arXiv preprint arXiv:1911.11313*.
Tinbergen, J. 1939. *Statistical Testing of Business Cycle Theories: A Method and its Application to Investment Activity*, 27—33. Geneva: League of Nations.
Tol, R. S. J. 2009. “The Economic Effects of Climate Change.” *The Journal of Economic Perspectives* 23 (2): 29—51.
Tsionas, M., C. F. Parmeter, and V. Zelenyuk. 2023. “Bayesian Artificial Neural Networks for Frontier Efficiency Analysis.” *Journal of Econometrics* 236 (2): 105491.
Turing, A. 1950. “Computing Machinery and Intelligence.” *Mind* LIX (236): 433—60.
Van Dijk, H. K. 1999. “Some Remarks on the Simulation Revolution in Bayesian Econometric Inference.” *Econometric Reviews* 18 (1): 105—12.
Van Dijk, H. K., and T. Kloek. 1980. “Further Experience in Bayesian Analysis Using Monte Carlo Integration.” *Journal of Econometrics* 14 (3): 307—28.

Van Dijk, H. K., and T. Kloek. 1983. "Experiments with Some Alternatives for Simple Importance Sampling in Monte Carlo Integration." Technical report.

Varian, H. R. 2014. "Big Data: New Tricks for Econometrics." *The Journal of Economic Perspectives* 28 (2): 3—28.

Vavra, J. 2021. "Tracking the Pandemic in Real Time: Administrative Micro Data in Business Cycles Enters the Spotlight." *The Journal of Economic Perspectives* 35 (3): 47—66.

Watson, O. J., G. Barnsley, J. Toor, A. B. Hogan, P. Winskill, and A. C. Ghani. 2022. "Global Impact of the First Year of COVID-19 Vaccination: A Mathematical Modelling Study." *The Lancet Infectious Diseases* 22 (9): 1293—302.

Yule, G. U. 1971. "On a Method of Investigating Periodicities in Disturbed Series with Special Reference to Wolfer's Sunspot Numbers." *Statistical Papers of George Udny Yule* 389—420, London: Griffin.