

TI 2014-085/III

Tinbergen Institute Discussion Paper



On the Rise of Bayesian Econometrics after Cowles Foundation Monographs 10, 14

Nalan Basturk¹

Cem Cakmakli²

S. Pinar Ceyhan¹

Herman K. van Dijk¹

¹ *Erasmus School of Economics, Erasmus University Rotterdam, and Tinbergen Institute, the Netherlands;*

² *Koc University, Turkey.*

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at <http://www.tinbergen.nl>

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: <http://www.dsf.nl/>

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

On the Rise of Bayesian Econometrics after Cowles Foundation Monographs 10, 14*

Nalan Basturk^{1,5}, Cem Cakmakli^{2,4}, S. Pinar Ceyhan^{1,5}, and
Herman K. van Dijk^{1,3,5}

¹*Econometric Institute, Erasmus University Rotterdam*

²*Department of Quantitative Economics, University of Amsterdam*

³*Department of Econometrics, VU University Amsterdam*

⁴*Department of Economics, Koc University*

⁵*Tinbergen Institute*

Abstract

This paper starts with a brief description of the introduction of the likelihood approach in econometrics as presented in Cowles Foundation Monographs 10 and 14. A sketch is given of the criticisms on this approach mainly from the first group of Bayesian econometricians. Publication and citation patterns of Bayesian econometric papers are analyzed in ten major econometric journals from the late 1970s until the first few months of 2014. Results indicate a cluster of journals with theoretical and applied papers, mainly consisting of *Journal of Econometrics*, *Journal of Business and Economic Statistics* and *Journal of Applied Econometrics* which contains the large majority of high quality Bayesian econometric papers. A second cluster of theoretical journals, mainly consisting of *Econometrica* and *Review of Economic Studies* contains few Bayesian econometric papers. The scientific impact, however, of these few papers on Bayesian econometric research is substantial. Special issues from the journals *Econometric Reviews*, *Journal of Econometrics* and *Econometric Theory* received wide attention. *Marketing Science* shows an ever increasing number of Bayesian papers since the middle nineties. The *International Economic Review* and the *Review of Economics and Statistics* show a moderate time varying increase. An upward movement in publication patterns in most journals occurs in the early 1990s due to the effect of the ‘Computational Revolution’. The paper continues using a visualization technique to connect papers and authors around important empirical subjects such as forecasting in macro models and finance, choice and equilibrium in micro models and marketing, and around more methodological subjects as model uncertainty and sampling algorithms. The information distilled from this analysis shows names of authors who contribute substantially to particular subjects. Next, subjects are discussed where Bayesian econometrics has shown substantial advances, namely, implementing stochastic simulation methods due to the computational revolution; flexible and unobserved component model structures in macroeconomic and finance; hierarchical structures and choice models in microeconomics and marketing. Three issues are summarized where Bayesian and frequentist econometricians differ: Identification, the value of prior information and model evaluation; dynamic inference and nonstationarity; vector autoregressive versus structural modeling. A topic of debate amongst Bayesian econometricians is listed as objective versus subjective econometrics. Communication problems and bridges between statistics and econometrics are summarized. A few non-Bayesian econometric papers are listed that have had substantial influence on Bayesian econometrics. Recent advances in applying simulation based Bayesian econometric methods to policy issues using models from macro- and microeconomics, finance and marketing are sketched. The paper ends with a list of subjects that are important challenges for twenty-first century Bayesian econometrics: Sampling methods suitable for use with big data and fast, parallelized and GPU, calculations, complex economic models which account for nonlinearities, analysis of implied model features such as risk and instability, incorporating model incompleteness, and a natural combination of economic modeling, forecasting and policy interventions.

*Nalan Basturk (basturk@ese.eur.nl), Cem Cakmakli (ccakmakli@ku.edu.tr), S. Pinar Ceyhan (ceyhan@ese.eur.nl), correspondence to Herman K. van Dijk (hkvandijk@ese.eur.nl). A very preliminary version of this paper, Basturk, Cakmakli, Ceyhan and Van Dijk (2013b), was presented at the NBER Workshop on Methods and Applications for DSGE Models at the Philadelphia Federal Reserve Bank, September 2013. Useful comments from the editor, Jean Sébastien Lenfant, two anonymous referees, Luc Bauwens, Gary Chamberlain, John Geweke, Ed Leamer, Harald Uhlig, Neil Shephard, Frank Schorfheide, Rui Almeida and Lennart Hoogerheide are gratefully acknowledged. Remaining errors are the authors’ own responsibility. Nalan Basturk and Herman K. van Dijk are supported by NWO Grant 400-09-340 and Cem Cakmakli by the AXA Research Fund.

Keywords: History, Bayesian Econometrics
JEL Classification: B4, C1

1 Introduction

Bayesian econometrics is now widely used for inference, forecasting and decision analysis in economics, in particular, in macroeconomics, finance and marketing. Three practical examples of this use are: In many modern macro-economies the risk of a liquidity trap, defined as low inflation, low growth and an interest rate close to the zero lower bound, is relevant information for the specification of an adequate monetary and fiscal policy; international corporations that sell their goods abroad want to know the risk of foreign exchange rate exposure that they incur in order to specify an optimal time pattern for the repatriation of their sales proceeds; evaluating the uncertainty of the effect of a new pricing policy is highly relevant in advertising strategies of supermarket chains. Particular references and more examples, for instance Bos, Mahieu and Van Dijk (2000), are given in textbooks like Lancaster (2004); Geweke (2005); Rossi et al. (2005) and Koop et al. (2007). More formally stated, there is now a widespread interest in and use of conditional probability assessments of important economic issues given available information that stems from different sources such as data information and experts knowledge. This has come a long way from the early steps of Bayesian econometrics in the 1960s following the likelihood based inference reported in the brilliant Cowles Foundation monographs 10 and 14, see Koopmans (1950) and Hood and Koopmans (1953). Papers in these monographs applied the likelihood approach introduced by R.A. Fisher, see Fisher (1912, 1922), to, predominantly, a system of simultaneous equations where immediate feedback mechanisms posed substantial methodological challenges for econometric inference. Several shocks occurred to this line of research in the early and middle part of the 1970s: Data series exhibited novel features like strong persistence over time, regime changes and time varying volatility. New modeling concepts, in particular the vector autoregressive approach, see Sims (1980), were developed to accommodate the observed economic data features without imposing very detailed restrictions on the model structure. Novel computational techniques based on stochastic simulation methods known as Importance Sampling and Markov chain Monte Carlo were introduced, see Kloek and Van Dijk (1975, 1978) and Metropolis and Hastings, see Metropolis et al. (1953) and Hastings (1970). This freed Bayesian analysis from very restrictive modeling and prior assumptions and opened a wide set of new research lines that have had substantial methodological and practical consequences for economic analysis, forecasting and decision strategies. Structural economic models based on dynamic stochastic general equilibrium concepts, flexible models based on unobserved components allowing for time varying parameters and data augmentation and stochastic simulation methods making use of filtering and smoothing techniques allowed researchers in academia and professional or-

ganizations to analyze complete forecast distributions, in particular, the tails like in Value-at-Risk and allowed an increased focus on measuring policy effects. These topics are becoming more and more prominent research fields.

In the present paper we start with a brief review of the introduction of the likelihood approach to econometrics as presented in Cowles Foundation Monographs 10 and 14. A sketch is given of the criticisms on this approach mainly from the first group of Bayesian econometricians in the 1960s and early 1970s. Next, in section 3 we describe historical developments of Bayesian econometrics from 1978 until the first few months of 2014 by collecting and analyzing the publication and citation patterns of Bayesian econometric papers in ten major econometric journals over that period. The number of pages written on Bayesian econometrics have been recorded as a percentage of the total number of pages for each year for all journals. There appear four prominent results: The existence of two clusters of journals, one with a high number and one with a low number of Bayesian econometric papers; a substantial scientific impact of a few papers in the more theoretical journals; a large number of citations from papers published in special issues and, fourthly, an increase in the number of published Bayesian econometric papers since the early nineties due to the ‘Computational Revolution’. More specifically, journals which contain both theoretical and applied papers, such as *Journal of Econometrics*, *Journal of Business and Economic Statistics* and *Journal of Applied Econometrics*, publish the large majority of high quality Bayesian econometric papers in contrast to theoretical journals like *Econometrica* and *the Review of Economic Studies*. These latter journals publish, however, some high quality papers that had a substantial impact on Bayesian research. The journals *Econometric Reviews* and *Econometric Theory* publish key invited papers and/or special issues that received wide attention, while *Marketing Science* shows an ever increasing number of papers since the mid nineties. The *International Economic Review* and the *Review of Economics and Statistics* show a moderate time varying increase. It is noteworthy that since the early nineties there exists an upward movement in publication patterns in most journals probably due to the effect of the ‘Computational Revolution’.

In section 4, we apply a visualization technique, using data from the JSTOR digital archive, data from leading journals and the references in the Handbook of Bayesian Econometrics edited by Geweke, Koop and Van Dijk (2011), which will henceforth be referred to as ‘the Handbook’. In this way, papers and authors are connected around important subjects in theoretical and empirical econometrics. The proximity of subjects that we consider is defined by the number of times that keywords or names appear together or in relation to all keywords and names cited together. The results show the interconnections of several subjects of interest. The macroeconomics and finance literature is related to simulation and filtering methods as well as to methods dealing with model uncertainty. Macro models used for policy purposes are related to fundamental identification issues. Marketing and micro economic panel

data models are linked to flexible model and prior structures such as hierarchical Bayes and Dirichlet processes with implications for treatment effects. The information distilled from this analysis shows also names of authors who contribute substantially to particular subjects.

Influential papers in Bayesian econometrics have been earlier analyzed in Poirier (1989, 1992), where quantitative evidence is provided of the impact of the Bayesian viewpoint as measured by the percentage of pages devoted to Bayesian topics in leading journals. We contribute to this literature by extending the bibliographical data with more recent papers and additional leading journals. Our contribution differs from the literature in several ways. First, regarding the influential papers in the field, we consider an alternative measure, the number of citations of each paper in addition to the percentage of pages devoted to Bayesian topics. The impact of papers are found to be different according to the criteria chosen for this purpose. Second, we define a set of influential papers in the field relying on the references in the Handbook. Third, we consider clustering of papers without defining a measure for their influence. This analysis is based on online bibliographic databases and the results are not affected by a very personal definition of influential papers in the field of Bayesian econometrics.

After the bibliographic analysis, we follow in section 5 with a discussion of those subjects that pose interesting challenges for discussion amongst Bayesian econometricians. These refer to: The effects of the computational revolution on Bayesian econometrics; advances in modeling and inference such as flexible and unobserved component models in macroeconomics and finance and hierarchical and choice models in microeconomics and marketing. In section 6 we summarize advances in the following three research areas where Bayesian and non-Bayesian econometricians differ: Endogeneity, instrumental variables, prior information and model evaluation; dynamic inference, nonstationarity and forecasting; vector autoregressive versus structural modeling. Then in section 7, we discuss the internal debate on the relative merits of objective versus subjective econometrics. There is probably not a subject area in the literature where everything goes smoothly. Bayesian econometrics is no exception. Further, a description of communication mechanisms between statistics and econometrics is given and we indicate a few non-Bayesian papers which have had a substantial influence on the development of Bayesian econometrics in the past 30 years. In section 8 we sketch recent developments in the applications of simulation based Bayesian econometrics in the field of decisions and policy analysis in macroeconomics, finance and marketing and in the field of treatment effects in microeconomics. In section 9, a list of important themes is given that we predict to be a challenge for twenty-first Bayesian econometrics. These refer to big data, model complexity, parallel computing, model incompleteness and the natural connection between Bayesian forecasting and decision analysis. This list contains the authors' personal expectations about the future of Bayesian econometrics.

We end this introduction with two remarks. First, many more interesting Bayesian economic and

econometric papers exist and are published in other major economics journals like the *American Economic Review*, *Journal of Monetary Economics*, *Journal of Political Economy* and *Quarterly Journal of Economics*. These are, however, outside the scope of the present paper. Second, detailed and excellent discussions of early historical developments of Bayesian econometrics are available. We refer to Pagan (1987, 1995), Poirier (1992, 1989, 2006), Qin (1996, 2013), Koop (1994), Gilbert and Qin (2005), Sims (2007, 2012), Zellner (2009) and the references cited in these papers for more background. Our paper sketches several historical developments but we emphasize that we do not provide a complete bibliographic analysis.

2 Cowles Commission Research and Early Bayesian Econometrics

Cowles Foundation Monographs 10 and 14, published in 1950 and 1953 respectively, contain foundations for modern inference in econometrics after the famous Haavelmo papers of 1943 and 1944, see Haavelmo (1943a,b, 1944). Haavelmo took up the Keynes-Tinbergen debate on the possible use of economic data with the purpose of estimating sets of equations that adequately describe the dynamic behaviour of an economy, in particular its feature of a periodic business cycle. Keynes characterized Tinbergen's approach, Tinbergen (1939), of estimating the parameters of an econometric model and computing quantitative policy scenarios as '...statistical alchemy ...', arguing that this approach ... is a means of giving quantitative precision to what, in qualitative terms, we know already as the result of a complete theoretical analysis ...', see Keynes (1939, p. 560) and Keynes (1940). Tinbergen (1940), on the other hand, argued that economic theories cannot present a complete description of data patterns. Error terms that bridge the gap between equation systems and observations were added and basic econometric methods, correlation and regression, were used by Tinbergen to estimate the numerical values of the coefficients in dynamic models that determine the cyclical and stability properties of a model; see for an informal summary of this debate Van Dijk (2013b) and for a detailed background analysis Sims (2012). Haavelmo extended and formalized the early econometric work by specifying a complete probability model for the set of economic variables and by proposing an accompanying inferential approach to empirically analyze economic phenomena like business cycles within a system of equations. Although Haavelmo listed two interpretations of probability: the 'frequentist' concept and the 'a priori confidence' one, see Haavelmo (1944, p. 48), the first one was mainly used in the research of the Cowles Commission. The focus in Monographs 10 and 14 was on applying the method of maximum likelihood, due to R.A. Fisher and developed in the early nineteen-twenties, see Fisher (1912, 1922), for the case of a system of simultaneous equations. Identification issues and estimation procedures such as full information maximum likelihood, limited information maximum likelihood and

the corresponding numerical optimization methods to find the maximum were the predominant topics that were analyzed and developed. Among the major contributors to this area are Cowles Commission researchers: Koopmans (1945), Anderson (1947), Anderson and Rubin (1949), Hurwicz (1950), Chernoff and Divinsky (1953) and Chernoff (1954).

As stated, much of this research followed the Fisherian likelihood approach. Fisher rejected Bayesianism and had as alternative so-called ‘fiducial inference’ (see Fisher (1973) and Aldrich (1995) for a review), which was Fisher’s attempt to use inverse probability and to analyze the shape of the likelihood without stochastic prior information. This has been characterized by Savage as: ‘A bold attempt to make the Bayesian omelette without breaking the Bayesian eggs’, Savage (1961).

The frequentist interpretation of estimators obtained by using the likelihood approach became known as the ‘classical approach’ in econometrics. This classical approach has some very restrictive assumptions. In Cowles Foundation Monograph 23, Rothenberg (1973), argued that for efficiency, accuracy, and credibility one usually makes use of the Cramér-Rao lower bound of the inverse of the Fisher Information matrix (or variance) of the estimator. This holds only for unbiased estimators and is in most cases, in particular in dynamic models, only asymptotically valid. Second, one makes use of exact restrictions as prior conditioning information, which is often unrealistic and overly restrictive. Further, Sims (2012) very recently argued that Haavelmo’s research program contained, apart from the limitations of using the frequentist approach, also an unclear treatment of policy analysis. Given that there exists uncertainty about policy effects as well as about parametric structures, it is only natural that ‘policy behaviour equations should be part of the system’, Sims (2012, p. 1189). More restrictions of the frequentist approach like the extensive use of sequential testing of a large number of hypotheses in order to ‘accept’ a particular model specification without specifying an alternative and without taking into account the so-called ‘pretest’ problems are discussed below.

Given the work on the implementation of the likelihood approach to econometrics and the early recognition of its limitations, it is natural that the Bayesian approach would follow.¹ Among major contributions to this literature are three important books. The first one is Raiffa and Schlaifer (1961) who introduced the concept of conjugate analysis as a way to construct informative prior information on model parameters. The idea is that the model is already in existence in the period before the data are observed or alternatively that the model is in existence for related data sets in other countries or for a different set of agents with similar features and it is most useful to incorporate this source of information in a sensible information processing technique. A second book was Schlaifer (1959), later summarized in Pratt et al. (1995), where practical decision problems were explained and analyzed. Here

¹Note that in post World War-II econometrics one has, apart from the likelihood approach, the (dynamic) regression methods and GMM as major schools of econometric inference. We only refer to the likelihood approach in the present paper.

a connection with the field of finance and business was made. Thirdly, there came the very influential ‘Bible’ of analytical results in Bayesian econometrics by Zellner (1971). All econometric models that were in use at that time were analyzed in this classic book from a Bayesian perspective. Analogies and differences between the classical and Bayesian approach were discussed, often using a weak or non-informative prior approach.

During and after these early Bayesian steps, there were several issues that attracted attention in the econometric literature. We document five of these in this section.

Structural modeling using conjugate priors

The focus on likelihood inference of structural equation systems led researchers in Bayesian econometrics in the early period to analyze ways of incorporating prior information that was in natural connection or ‘conjugate’ to the information of the equation system. This ‘natural conjugate approach’ specified a family of prior distributions that was analytically tractable and convenient for the case of the standard linear regression model. This density is known as the normal-inverted gamma density and it allows to update prior information using Bayes theorem in a simple way: The posterior mean of the regression parameters is a weighted average of prior mean and data mean with weights that are given as the relative accuracy of prior and data mean, respectively. It maybe of interest to note that Raiffa and Schlaiffer’s book was used at the Econometric Institute in Rotterdam with an emphasis on chapters 1, 2, and 3 that dealt with ‘Experimentation and Decision’ and Chapters 7, 8, 11, 12, and 13 that dealt with ‘Distribution Theory and Conjugate Analysis for the Normal and the Regression Process’.

However, several restrictions came forward regarding properties of the natural conjugate family of which we list three. Inequality restrictions on the range of equation system parameters, such as income elasticities in the unit interval and stability restrictions in dynamic models, are very natural restrictions in economics and these cannot easily be dealt within the conjugate approach. Secondly, for the basic linear regression model, Richard (1973) and Bauwens (1991) showed that a natural conjugate prior density that is non-informative on the variance of the disturbances but informative on part of the regression parameters gives the paradoxical result that the posterior of the latter becomes equal to the prior with probability one. Rothenberg (1963, 1973) showed that for a system of regression equations a prior density that belongs to the conjugate family and that is proportional to and of the same functional form as the likelihood implies that the variances of the parameters in the r^{th} equation have to be proportional to the variances of the corresponding parameters in the s^{th} equation. There is no a priori economic reason why this mathematical restriction should be the case. This creates a problem for inference in systems of equations, like Seemingly Unrelated Regression Equations (SURE) and Vector AutoRegressive (VAR) models, and Simultaneous Equations Models (SEM) and it is known

as ‘Rothenberg’s problem’. These results show that one has to think carefully about prior information and not mechanically apply natural conjugate priors.

In order to tackle the restrictions that conjugate priors impose on simultaneous equations systems, Drèze extended the natural conjugate family for such a system, see Drèze (1962) and Drèze and Richard (1983). Drèze followed also a different path to limit this restriction by concentrating the analysis on a single equation within a system of equations which is known as the Bayesian limited information approach, see Drèze (1976) and Bauwens and Van Dijk (1990) for details. However cavalier these approaches were, the restrictions that the natural conjugate prior family impose often limit inference and forecasting and different attempts in research were started to free the analysis from these analytical restrictions.

Fragility of structural inference and incredible structural restrictions

The possibility of incorporating parameter restrictions in structural inference through the use of prior information has been an appealing feature of Bayesian inference. It was acknowledged that the size of most data is too small to rely on asymptotic properties of frequentist analysis and that for most applied cases, economically interpretable ranges for structural coefficients can be defined. Furthermore, in the absence of such ranges for structural coefficients, there is often a gap between economic theory and econometric analysis.

The introduction of feasible ranges for structural coefficients in frequentist methods relied on exact parameter restrictions. These ad hoc restrictions, thus the implied information supplied by the modeler, to guide econometric analysis was first criticized by Drèze (1962). Specifically, Drèze advocated Bayes’ principle as a natural framework with appropriate generality to quantify the uncertainty around the prior information in cases where the model estimation required ‘substantial prior information’.

During the seventies, Leamer (1974, 1978) continued to question strongly the precise inferential conclusions that were drawn using structural models and the only asymptotically valid frequentist methods. Leamer’s criticisms on the inclusion of structural restrictions were in line with those of Drèze, in defining Bayesianism as the approach to explicitly account for uncertain prior information revealing the modeler’s information, see Leamer (1978, p. 510). Unlike Drèze, Leamer’s arguments to explicitly account for the modeler’s information was concentrated on the ‘mapping’ between the prior and posterior, where the former is mostly judgemental. For a sequence of priors, this mapping is called the ‘information contract curve’ and it should be reported explicitly. In other words, ‘the mapping is the message’, see Leamer (1978).

The fragility of structural estimation led Leamer to concentrate on two topics in econometric analysis: The relation between priors and the issue of collinearity; and specification search procedures

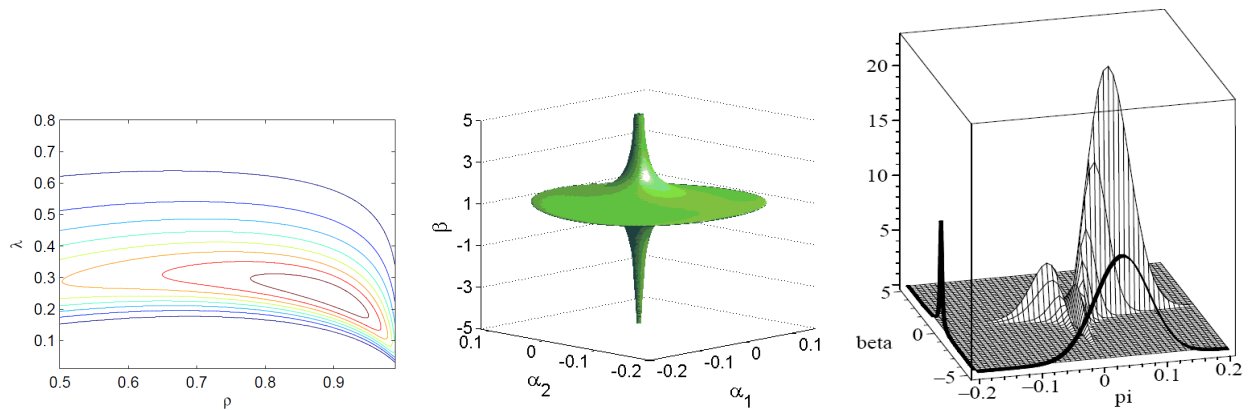
particularly with respect to modeler information. The issue of collinearity and its consequences in parameter estimation were linked to ‘uncertain prior information’ rather than weak data evidence, and an adequate definition of modeler’s information through the prior was advocated to enable the interpretation of data evidence in a parameter-by-parameter fashion, see Leamer (1973). Regarding the model choice given structural restrictions, Leamer specified six specification search procedures that indicate and accommodate the sensitivity and/or the plausibility of the obtained estimates. These specification search procedures are the ‘hypothesis testing search’ to choose a ‘true model’; the ‘interpretive search’ to interpret multidimensional evidence; the ‘simplification search’ to obtain a ‘fruitful’ and interpretable model; the ‘proxy search’ to find a quantitative facsimile; the ‘data selection search’ to analyze model implications in subsamples of data; and the ‘postdata model construction’ improve an existing model. In all procedures but ‘postdata model construction’, the final empirical statistical evidence had to be discounted as ‘part of the data evidence is spent to specify the model’, see Leamer (1978, p. 12) and Leamer (1974). In particular, Leamer’s ‘extreme bounds’ analysis, aimed to measure the fragility of parameter estimates to prior information, has been a prominent feature of his work, see Leamer (1985), Leamer (1978, ch. 10), Pagan (1987, 1995) and Qin (1996, 2013). Following Leamer’s analysis, formalization and utilization of all prior information, and to assess the degree of ‘domination’ of prior information became major concerns in Bayesian analysis.

Sims (1980) took a different line of attack on structural restrictions and criticized the ‘incredibility’ of many of the theoretical restrictions that were a priori imposed on the parametric structure of systems of equations as exactly known without checking their ‘plausibility’. Sims proposed a vector autoregressive model structure for systems of variables where the data information and dynamic properties are more adequately handled and supported by data information. This modeling approach evolved along several lines: Convenient priors were developed in Doan et al. (1984): More soft structural information was added in the so-called structural VAR approach and connections were made with economic models like the DSGE models using recursive VAR models with Kalman Filters and using formal Bayesian MCMC techniques for estimation purposes. It is noteworthy that this line of research found in recent years many applications in academia and professional organizations like Central Banks and the Federal Reserve System. More details are presented in sections 5–8.

Start of the Computational Revolution

A very effective approach to free Bayesian econometrics from the analytical restrictions that are inherent in the conjugate analysis and the tight structural equations turned out be the use of stochastic simulation methods that are known as Monte Carlo simulation. Figure 1 shows examples of typical shapes of likelihoods and posterior distributions of models for realistic problems in economics. These

Figure 1: Examples of complex (non-elliptical) posterior distributions



shapes are very different from the usual elliptical one that is a feature of conjugate analysis. These distributions occur in finance (modeling daily stock returns), macroeconomics (modeling the joint behavior of variables with a long run equilibrium relationship), and microeconomics (modeling the effect of education on income), see Hoogerheide, Opschoor and Van Dijk (2012) and Basturk, Hoogerheide and Van Dijk (2013c) for more details on the shape and properties of these densities. Given that for many realistic economic models the shape of the likelihood and posterior is such that they cannot be accurately approximated by the clockwise shape of basic Gaussian densities, Monte Carlo methods were developed where one simulates random draws from a flexible distribution that is a good approximation to the shape of the likelihood and posterior. This indirect sampling procedure needs then a correction step that takes into account the distance between the true posterior and the approximate distribution in order to yield numerically correct results. Three methods became very popular in this context that are known as Rejection Sampling (Rej), Importance Sampling (IS) and Markov Chain Monte Carlo (MCMC). Using these novel simulation methods one can obtain reliable and accurate estimates of the properties of interest of such posterior distributions. Rejection sampling is the oldest method and introduced by von Neumann (1951). IS was introduced into statistics and econometrics by Kloek and Van Dijk (1975) and later published as Kloek and Van Dijk (1978) and the independence chain Metropolis-Hastings algorithm was introduced by Metropolis et al. (1953) and Hastings (1970). All these methods have been further developed, see section 5.

Apart from the need to free the Bayesian approach from restrictive priors and models, there also existed interest in evaluating uncertainty of policy effectiveness. An important example was to obtain the posterior distribution of the multiplier in a system of simultaneous equations, see Brainard (1967). Here one faces the issue that this multiplier is usually a ratio or more general a rational function of structural parameters. Given that the multiplier is a regular function of structural parameters, Monte Carlo simulation methods give an easy operational approach to obtain its finite sample distribution, see e.g. Van Dijk and Kloek (1980). A similar result holds for dynamic properties of models regarding

stability and forecasting. There is no need for a ‘plug-in’ estimation step that is used in frequentist analysis. More details are given in sections 5–9.

Testing, signifying nothing, sequential testing and credibility of the final chosen model

Despite attempts to apply the Bayesian approach to econometrics, frequentist testing became a major line of research with many applications in the past fifty years. However, it has not yielded substantial confidence in obtained final model specifications for a finite set of data and does not give immediate reliable indications of the uncertainty of implied policies. One reason for this is that the focus on statistical testing regularly means that the researcher does not raise the issue whether the results matter from an economic point of view. Statistical significant but economically almost meaningless is something that decision makers will often not accept as a sound basis for policy analysis, see McCloskey and Ziliak (1996) and The Economist (2004). In this context, Leamer’s argument is relevant that too much confidence was given to results obtained using advanced econometric methods, that are only asymptotically valid under restrictive conditions, instead of performing serious empirical research, see his statement ‘Let’s take the con out of econometrics’ in Leamer (1983). A second fundamental statistical weakness of the ‘classical approach’ is the testing of many different hypotheses in econometric models by a sequential testing procedure. In the analysis it is usually not taken into account that the distribution of the second test depends on the outcome of the first one and so on for further tests. One more problem is that no alternative is considered in testing hypothesis. Only falsification of the null hypothesis is usually considered although in the end a model is ‘accepted’ without stating a measure of ‘credibility’ of the final result. One natural Bayesian solution is to give weights to particular model features by using posterior odds analysis and Bayesian model averaging. Then one may pursue forecasting with a weighted average of model structures. This line of research is shown in, e.g. Wright (2008) and Strachan and Van Dijk (2013). A second Bayesian approach is to study policy effects of using alternative model structures some of which may be misspecified. A substantial loss may be obtained from a decision strategy using a wrong model. This clearly gives an important warning signal to modelers, for details, see Geweke (2010) and for an example where a wrong model yielded substantial financial losses in the 2008 financial crisis, see Billio et al. (2013).

Conditional probability statements and the measurement of policy effectiveness

Another fundamental problem with the frequentist approach is the difficulty of dealing with the issue of conditional probability statements which is a concept that is widely used in practice. Given a set of data, decision makers are usually interested in the probability of an unknown feature of a problem at hand. The earlier listed examples are clear indications: What monetary policy to be used in the

face of a liquidity trap given data for countries like Japan and the European Union; how to hedge currency risk for international corporations given data on exchange rate behavior; and which advertising policy to be implemented given scanner data about customer behavior are relevant supermarkets. Sims (2012) discusses the methodological connections between inferential procedures and economic policy that already existed in the 1930s and 1940s and continued to play an important role in the rational expectations literature in the late 70s and 80s. The frequentist hypothesis testing method ... ‘inhibits combination of information from model likelihood functions with information in the beliefs of experts and policymakers themselves’ Sims (2012, p. 1188). The modeling of policy interventions should be a part of the economic modeling process. Both should make use of probability distributions and this can be naturally done in a Bayesian framework. For the operational procedures and to free from implausible restrictions, the simulation based Bayesian approach is very suitable and many stochastic simulation methods are nowadays available. The effect that model incompleteness may have on plausible policy scenarios is also relevant in this context, see Geweke (2010). This issue will also be dealt with in section 8.

3 Exploratory Data Analysis

In this section we analyze the advance of Bayesian econometrics since the late 1970s from a descriptive point of view. Specifically, we analyze how Bayesian econometrics got in the mainstream and high quality econometric journals using the publication and citation patterns of 999 papers in leading journals during the period between 1978 and 2014 (March). We select these papers on the basis of their contributions to theoretical and/or applied topics in Bayesian econometrics and denote them by ‘Bayesian papers’. The list of leading journals consists of 10 journals: *Econometrica* (Ectr), *Econometric Reviews* (ER), *Econometric Theory* (ET), *International Economic Review* (IER), *Journal of Applied Econometrics* (JAE), *Journal of Business and Economic Statistics* (JBES), *Journal of Econometrics* (JE), *Marketing Science* (MS), *Review of Economic Studies* (RES) and *Review of Economics and Statistics* (ReStat). Our analysis extends the one from Poirier (1992) by including more journals and a longer period. We also make use of citation patterns. Detailed statistics of the papers considered in this section are provided in Appendix A, Tables A.1–A.4 and Figure A.1.

3.1 Publication Patterns

The first criterion we use to analyze the advances in Bayesian econometrics is the percentage of Bayesian pages in leading econometrics and quantitative economics journals that were listed above.

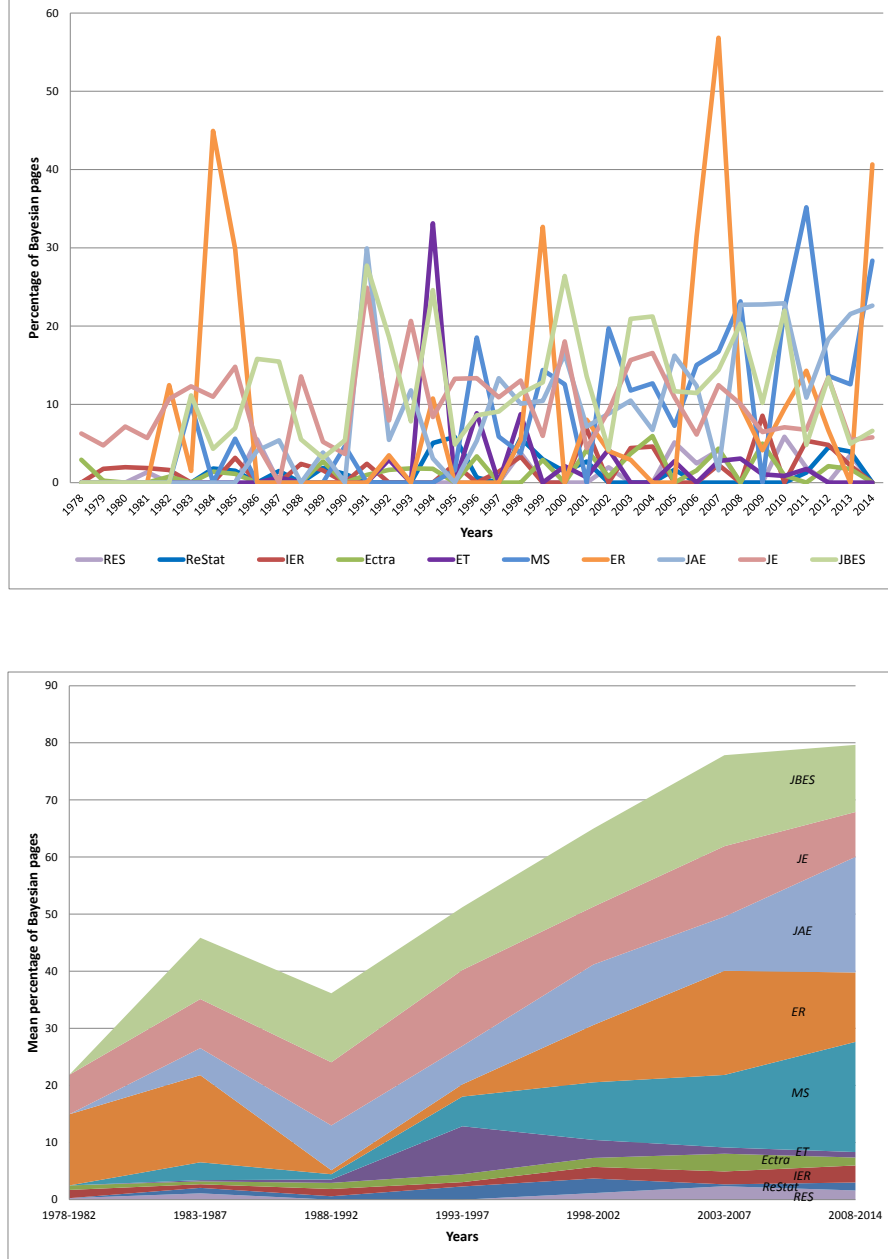
The top panel in Figure 2 presents the annual percentages of the pages allocated to Bayesian papers

for each journal. These percentages are usually below 30%, with exceptions in ER, ET and MS. There are three journal issues which have more than 40% Bayesian content. The ER issue in 1984 has four Bayesian papers constituting 44.93% of the total number of pages with as most influential paper the one by Doan et al. (1984). The 2007 issue of ER also has a high percentage of Bayesian pages, where 56.83% of the issue is devoted to 18 Bayesian papers, including An and Schorfheide (2007) as one of the largest papers. In the 2014 issue of ER, 40.65% of the total number pages is devoted to 15 Bayesian papers.

Special issues yield the highest values reported in the top panel of Figure 2. These are the ER issues in 1984, 1999, 2007, 2014; the ET issue of 1994 (on Bayes methods and unit roots); the JAE and JE issues in 1991, and to a lesser extent the JE issues in 1985, 2004, and 2012.

The bottom panel in Figure 2 presents the average percentage of pages for Bayesian papers in each journals over 5-year intervals. These average percentages provide general publication patterns compared to the top panel of Figure 2 since the influence of special journal issues related to Bayesian econometrics is now more limited due to the 5-year averaging. This panel of Figure 2 shows that the influence of Bayesian econometrics in terms of the percentage of allocated pages is time varying and journal dependent. Journals such as *Ectra*, *ET*, *RES*, *ReStat* and *IER* typically have low percentages of Bayesian pages, below 3% over the whole period. On the other hand, *JBES*, *JAE* and *MS* typically have high percentages of Bayesian pages, with a substantial increase in these percentages after 1990s. The set of journals with a large share of Bayesian papers show that Bayesian inference is mainly present in a combination of theoretical and applied papers rather than in sole theoretical papers. Figure 2 indicates two main clusters of journals in terms of their focus on Bayesian econometrics. The first cluster consists of journals with relatively low average number of percentages of Bayesian pages: *Ectra*, *ET*, *IER*, *RES* and *ReStat*. The average percentages of Bayesian pages in these journals are less than 3%. The second cluster consists of journals with relatively high average percentages of Bayesian pages: *ER*, *JAE*, *JBES*, *JE* and *MS*. The increased share of Bayesian pages is most visible for *MS*, then for *ER* and *JAE*, particularly for the period after 1992. To a lesser extent, this increasing pattern holds for *Ectra*, *IER* and *RES*. This general increasing influence of Bayesian econometrics after 1990s can be attributed to computational advances making Bayesian inference easier and the increased number of applied papers using Bayesian inference.

Figure 2: Percentages of pages allocated to Bayesian papers for all journals



Note: The figures present the annual percentage of pages (top panel) and 5-year averages of pages (bottom panel) of Bayesian papers for the period between 1978 and 2014 (March). In the bottom panel, the final period consists of 7 years. Abbreviations of journals are as follows: *Econometrica* (Ectra), *Econometric Reviews* (ER), *Econometric Theory* (ET), *International Economic Review* (IER), *Journal of Applied Econometrics* (JAE), *Journal of Business and Economic Statistics* (JBES), *Journal of Econometrics* (JE), *Marketing Science* (MS), *Review of Economic Studies* (RES) and *Review of Economics and Statistics* (ReStat).

3.2 Citation Patterns and Impact

We next focus on citation patterns of papers in the ten journals, as an additional criterion to define the advances in Bayesian econometrics after the late 1970s.²

The top panel of Figure 3 shows that the impact analysis is substantially different when we compare the citation patterns with the number of Bayesian pages. Although the journals with a high share of Bayesian pages, ER, JAE, JBES, JE and MS, also have a high share in the total number of citations, there are two high quality theoretical journals, Ectra and RES, with a high influence in the field in terms of the citation numbers despite their relatively low numbers of Bayesian pages. This high impact is more visible when we focus on papers with more than 400 citations, shown in the figures on the right panel of Figure 3. There are 31 papers satisfying this criteria. Note that ER has a large share in total citations although its share in terms of the percentage of pages is more time varying.

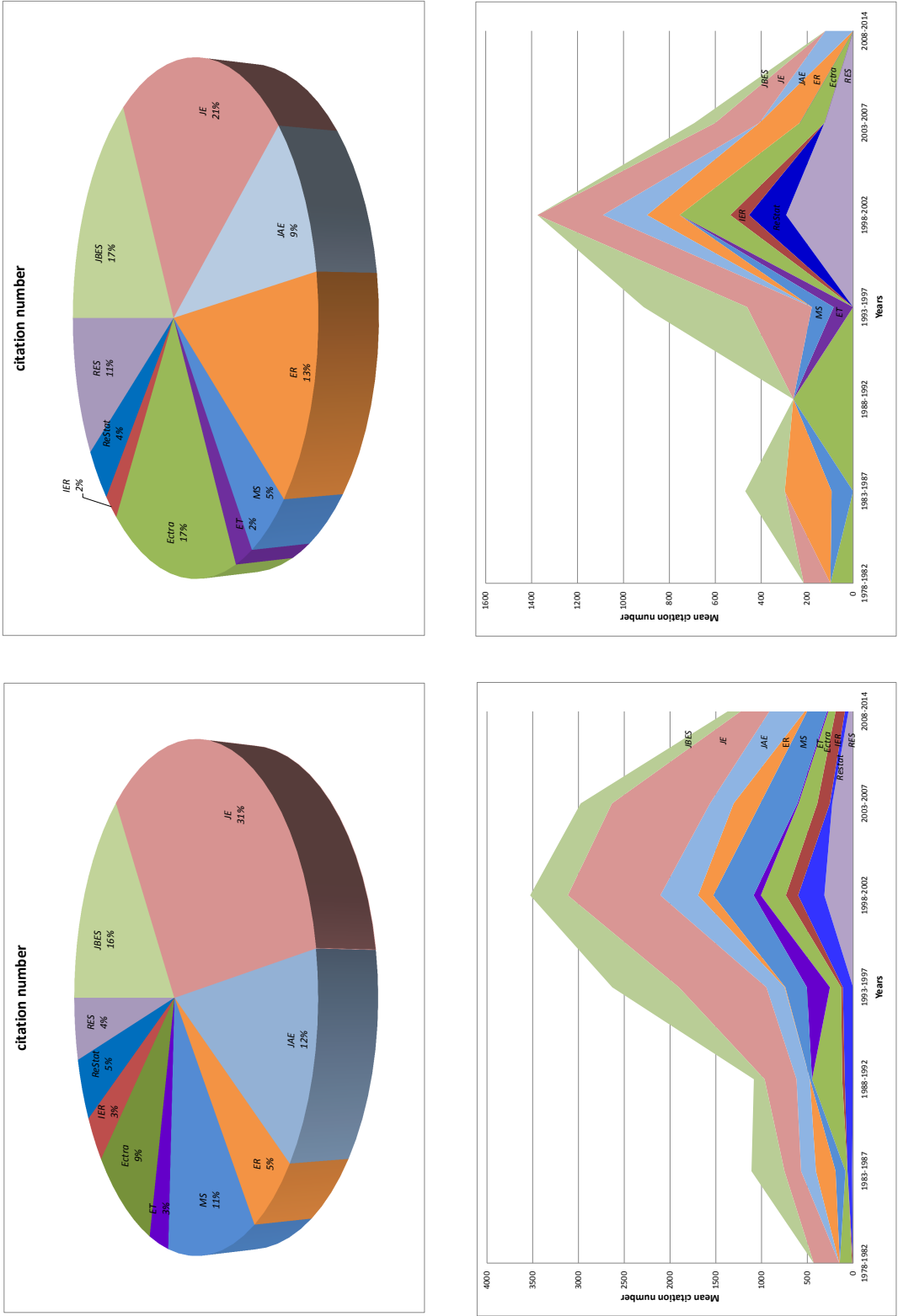
In order to compare influential papers in terms of their shares of pages and in terms of the number of citations, we next consider two clusters of journals, with high and low numbers of pages devoted to Bayesian econometrics, according to the publication patterns in section 3.1 and report the number of citations separately for the journals in these clusters. Furthermore, we report the citation patterns for highly influential papers, papers with at least 400 citations, in these clusters of journals. These citation patterns are provided in Figure 4. The top (bottom) left panels in Figure 4 present the number of citations for papers in leading journals with a low (high) number of pages devoted to Bayesian econometrics during the period between 1978 and 2014 (March). The top (bottom) right panels in Figure 4 present the number of citations for the highly influential papers with at least 400 citations in the leading journals with a low (high) number of pages devoted to Bayesian econometrics.

The left panel in Figure 4 shows that papers in cluster 2, which includes journals with a high number of pages devoted to Bayesian econometrics, are on average cited more times than those in cluster 1. Despite this similarity, the right panel in Figure 4 shows that highly influential papers with at least 400 citations are more evenly distributed across cluster 1 and cluster 2 journals. Particularly, Ectra and RES have papers that are highly cited.

We note that four papers are highly influential in the field with more than 1000 citations: Geweke (1989) (Ectra) with 1303 citations, Kim et al. (1998) (RES) with 1459 citations, Jacquier et al. (1994) (JBES) with 1347 citations and Doan et al. (1984) (ER) with 1020 citations. These papers refer to *computational advances*, *macroeconomics* and *financial econometrics* with a focus on time varying data patterns. We first note that the reported average number of citations is naturally low at the end of the sample period, especially between 2008–2014, since these papers are relatively new. When these recent

²The citation numbers are collected in the last week of April 2014. The number of citations are based on Google Scholar records available at <http://scholar.google.com/>.

Figure 3: Average citation patterns for papers in leading journals



Note: The figures show average citation numbers for the period 1978–2014 for all papers in leading journals. Figures on the left panel are based on all papers published in leading journals. Figures on the right panel are based on a subset of influential papers with at least 400 citations. Reported years correspond to the years that the cited papers are published. Journal abbreviations are as in Figure 2.

papers are not taken into account, an increasing pattern in the overall number of citations for Bayesian papers is visible: The total numbers presented in the bottom panel of Figure 3 clearly increase between 1978 and 2002. These figures also show that *Marketing Science* papers started to be cited much more heavily after 1996 and in general the 1990s bring more citations to each journal in our data set.

We present four conclusions from this exploratory data analysis:

1. *High and low cluster of journals:* A high cluster which contain both theoretical and applied papers, such as JAE, JE, JBES, ER and MS, publish the large majority of high quality Bayesian econometric papers. A low cluster of published Bayesian papers appear in theoretical journals, such as Ecetra and RES.
2. *Scientific impact through citations:* Theoretical journals like Ecetra and RES published few papers, but with substantial scientific impact through a large number of citations.
3. *Effect of special issues:* Special issues of the journals like *Econometric Reviews* and *Econometric Theory* receive more citations than usual issues since they contain influential papers.
4. *Increasing trend since 1990's:* A structural break, indicating a switch to an increasing number of Bayesian papers, occurs since the early 1990s, which is due to the use of novel computational techniques.

4 Subject and Author Connectivity

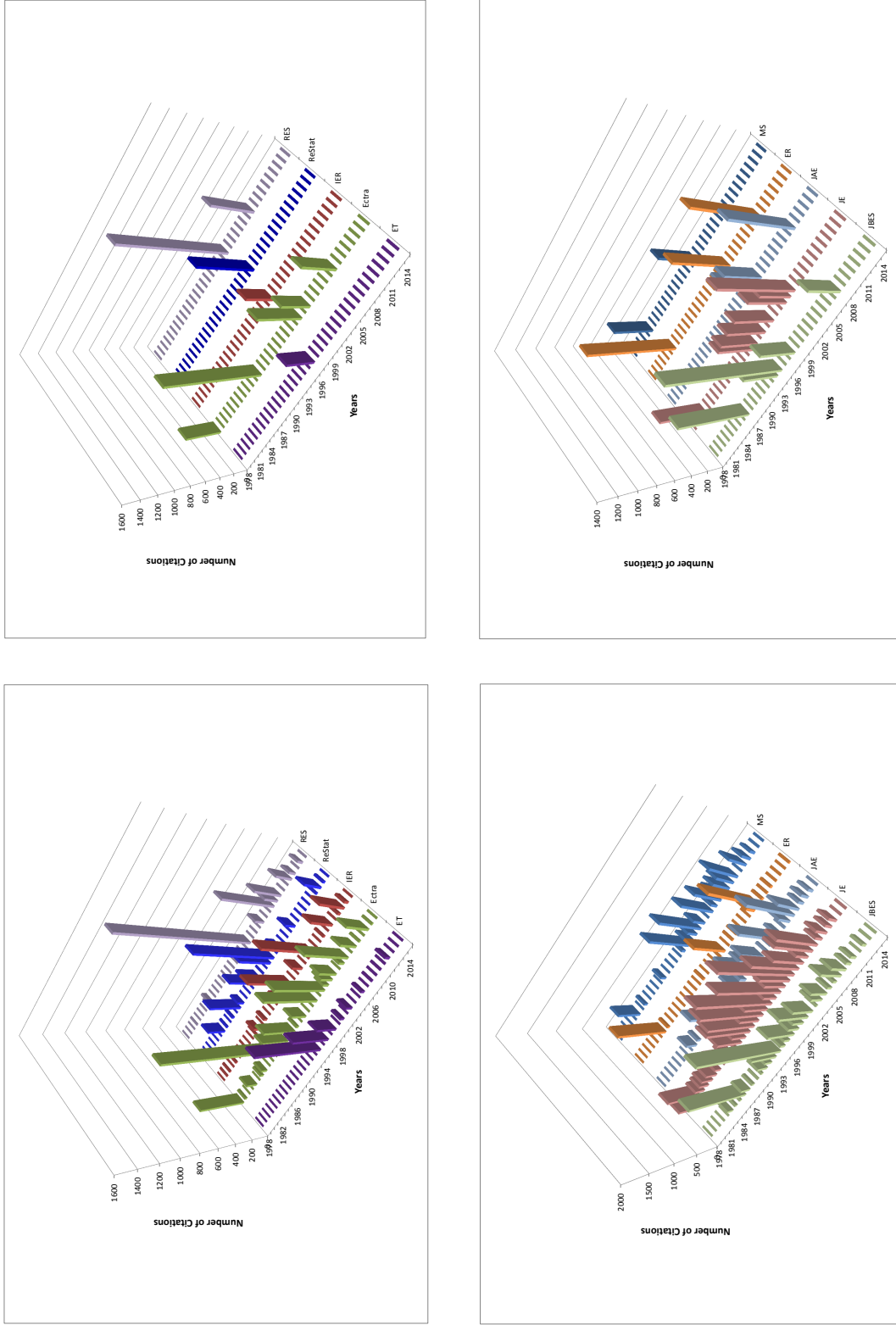
This section considers connectivity of subjects and authors in Bayesian econometric papers. The list of scientific papers in Bayesian econometrics is extensive. First, we first consider a large set of papers relying on digital archives in order to analyze subject connectivity. We use a data set of 1000 papers and *key terms* extracted from each paper provided by the JSTOR digital archive in the field of Bayesian econometrics.³ Next, we focus on influential papers taken from the Handbook of Bayesian Econometrics which have more than 100 citations according to Google Scholar or Web of Knowledge.⁴ This selection of papers uses *expert information*, since the set of papers is based on the careful selection of the authors of the Handbook. We summarize the connectivity of the keywords and the associated subject in the Handbook for each paper. Thirdly, we choose the set of influential papers presented in section 3. For each of the three data sets, the proximities are defined by the number of times that keywords or key terms appear together and in relation to their pairwise concurrence.⁵

³<http://www.jstor.org/>. JSTOR provides a randomly selected set of 1000 papers from their digital archive.

⁴<http://scholar.google.com/>, <http://thomsonreuters.com/web-of-science/>.

⁵The network maps we present are obtained from the VOSviewer program, available at <http://www.vosviewer.com> and their software to address proximity, see Waltman et al. (2010); Van Eck and Waltman (2010).

Figure 4: Citation numbers for journals with high and low numbers of pages devoted to Bayesian econometrics



Note: The figures show annual citation numbers for the period 1978–2014 for all papers in leading journals. Top (bottom) left panel shows citations of papers in journals with a low (high) percentage of Bayesian pages. Top (bottom) right panel shows the citations of papers with at least 400 citations in journals with a low (high) percentage of Bayesian pages. Reported years correspond to the years that the cited papers are published. Journal abbreviations are as in Figure 2.

Figure 5: Connectivity of subjects in JSTOR database

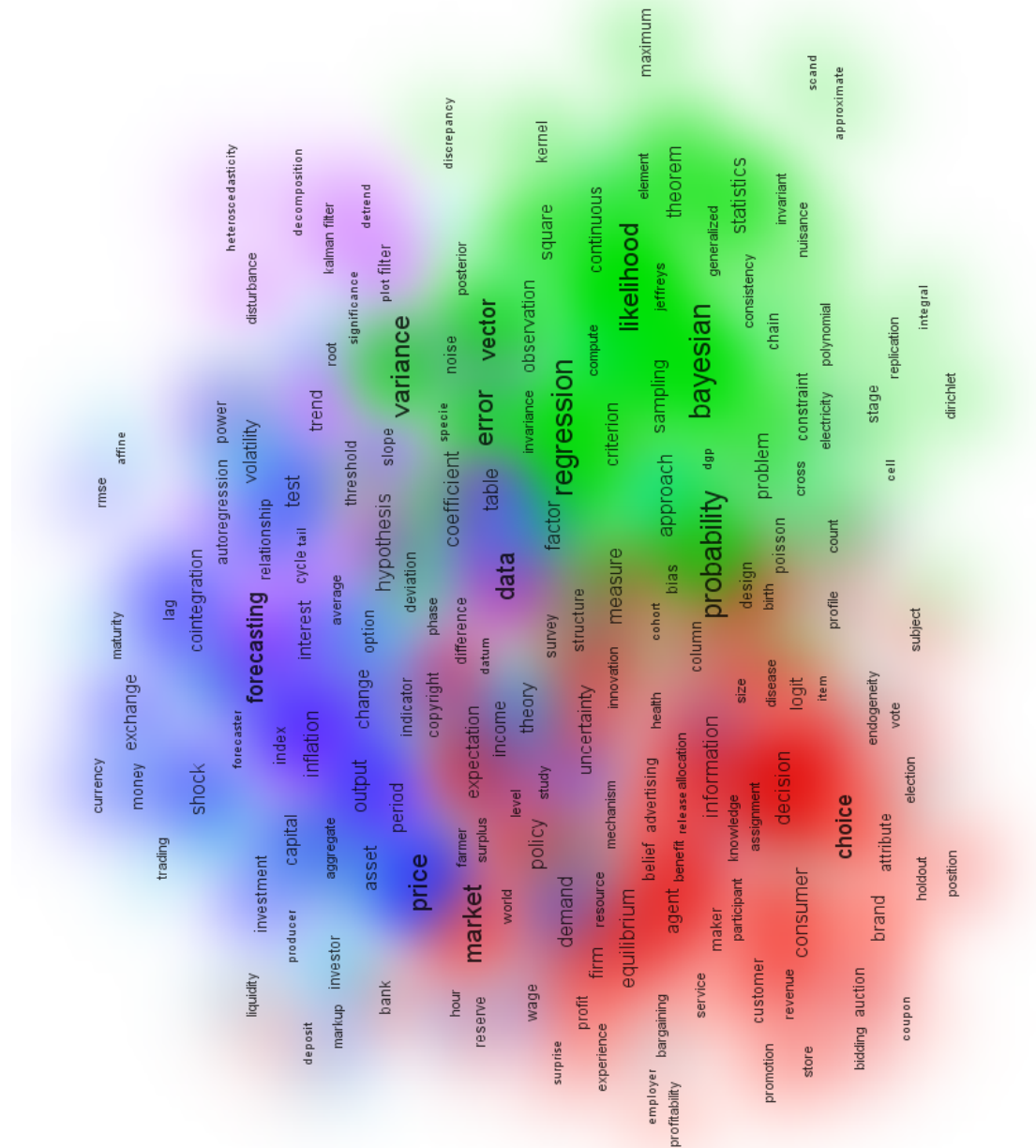


Figure 5 presents the network and connectivity map for the key terms based on 1000 selected papers in Bayesian econometrics, published since 1970. This connectivity analysis is solely based on a random sample of papers that JSTOR provides. Three major areas emerge from this connectivity analysis and these are presented in different colors in Figure 5.

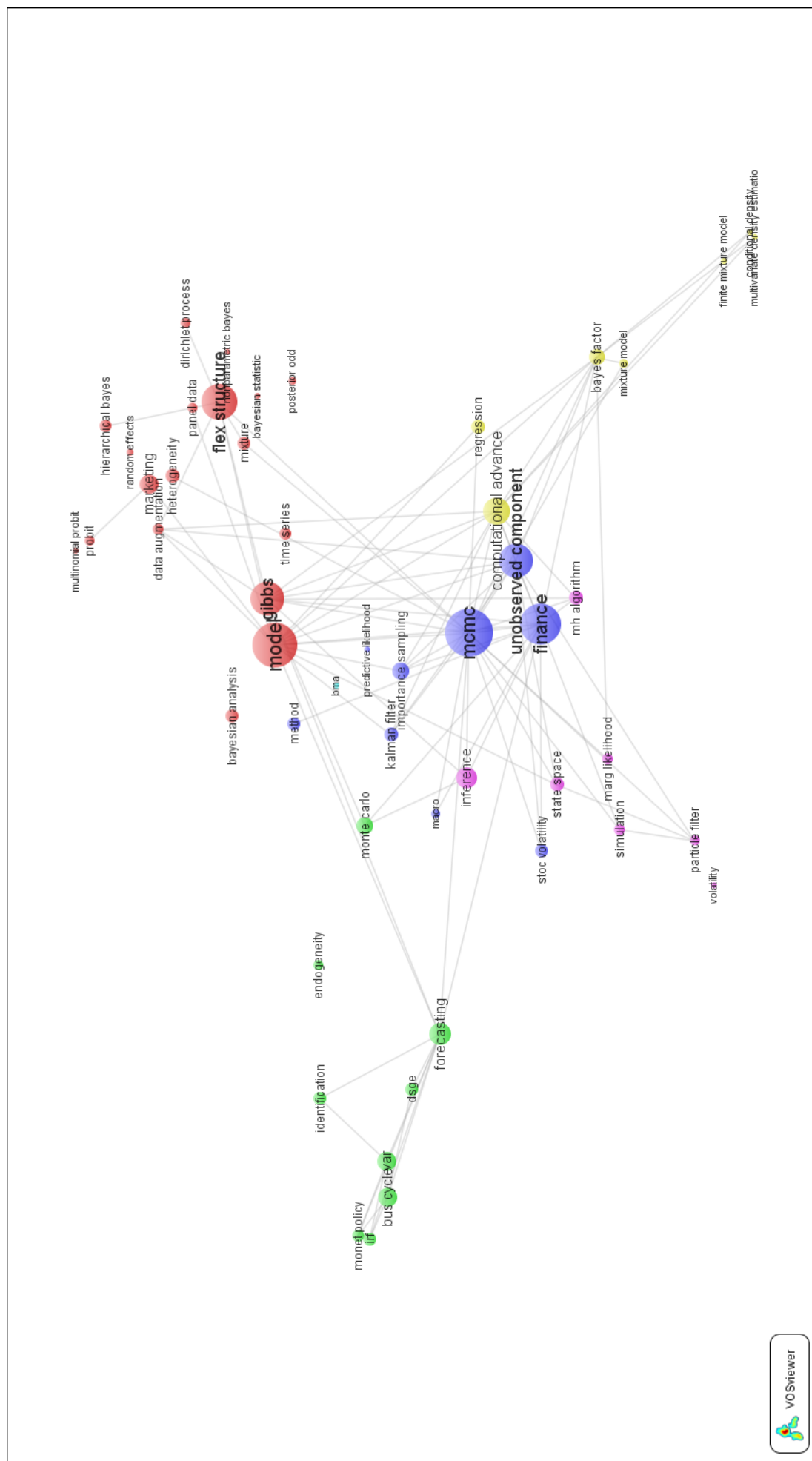
1. The first cluster of keywords, plotted in dark and light green in the figure, corresponds to theoretical subjects, with related keywords of likelihood, moment, statistics, assumption and probability. This cluster is naturally linked to all remaining clusters.
2. The second area, consisting of clusters colored in blue and purple, is centered around the key terms ‘forecasting’ and ‘price’. This cluster shows that particularly forecasting is central to the analysis of macroeconomic and financial data, such as (economic) growth, exchange (rates), (financial) returns and interest (rates). Most common models for these data include autoregressive models. (Forecast) horizon, regime (changes) and testing are related and important issues for this area.
3. The third area, shown in light and dark red in Figure 5 has as most prominent key terms ‘market’, ‘choice’, ‘information’ and ‘equilibrium’. Other keywords in this area, such as decision, brand, profit, behavior, equilibrium and utility signal market equilibrium models as well as choice models.

The connectivity analysis presented so far does not take into account the amount of influence of each paper, the papers’ original keywords or any extra information on the subject area. We next consider a refined set of influential papers in Bayesian econometrics, based on the citations in the Handbook. Note that the subjects and the references covered in the Handbook are divided to 9 chapters according to the subfields: endogeneity & treatment (Chapter 1), heterogeneity (Chapter 2), unobserved components & time series (Chapter 3), flexible structures (Chapter 4), computational advances (Chapter 5), micro & panel data (Chapter 6), macro & international economics (Chapter 7), marketing (Chapter 8) and finance (Chapter 9). We consider keywords of Bayesian papers cited in each chapter, and include the corresponding subfield as an additional keyword for each paper.

Figure 6 shows that the subfields defined in the Handbook are connected to several keywords. This is an expected outcome since we use the chapter information in the Handbook as ‘expert knowledge’ to relate each paper to a subfield. Besides these subfields, sampling techniques such as the Gibbs sampler, Markov Chain Monte Carlo (MCMC), Metropolis Hastings (MH) algorithm and importance sampling have very large weights indicated by the sizes of the circles in Figure 6 and they lie in the middle of the keyword connection map. This indicates that sampling algorithms are central to research in all subfields of Bayesian econometrics covered here.

An interesting result from Figure 6 is the connectivity of Bayesian methods and economic subfields.

Figure 6: Connectivity of subjects in papers cited in the Handbook, Koop and Van Dijk (2011)



Papers in the area of marketing are closely related to flexible model structures (flexible functional forms), and particularly hierarchical Bayes, Dirichlet processes, panel data methods and heterogeneity. Given the increased amount of consumer data in the marketing field, more complex model structures which can handle heterogeneity across consumers are becoming important for this field.

Figure 6 also indicates a strong relation between the macroeconomics and finance literature and Bayesian methods. First, the subject of forecasting is central for macroeconomics and finance as this keyword occurs very frequently and is linked to both areas. Second, state space models, particle filters, Monte Carlo methods, Kalman filter, predictive likelihood analysis and Bayesian Model Averaging (BMA) are closely related to the macroeconomics and finance literature. These close relations indicate the need for sophisticated simulation techniques, such as particle filters, for the estimation and forecasting of complex models used for financial and macroeconomic data. Furthermore, the issue of (parameter) identification is central for macro models used for policy analysis, such as the VAR, Impulse Response Functions (IRF), and business cycle models. This relation is shown in the lower right corner of Figure 6.

We finally note that computational advances have a large weight according to Figure 6. This subject is naturally linked to simulation methods, as speeding up computations is a central topic for the wide applicability of simulation methods. Computational advances are central especially for finite mixture models, and are in close relation to the areas of marketing and macro models.

We next select the influential papers in Bayesian econometrics based on the highly (more than 100 times) cited papers published in leading journals in section 3, and analyze the connectivity of the authors and the keywords of each paper. The connectivity of keywords and authors of these papers are shown in Figure 7 using a heatmap of the terms' density estimated by the concurrence of each keyword and author.⁶

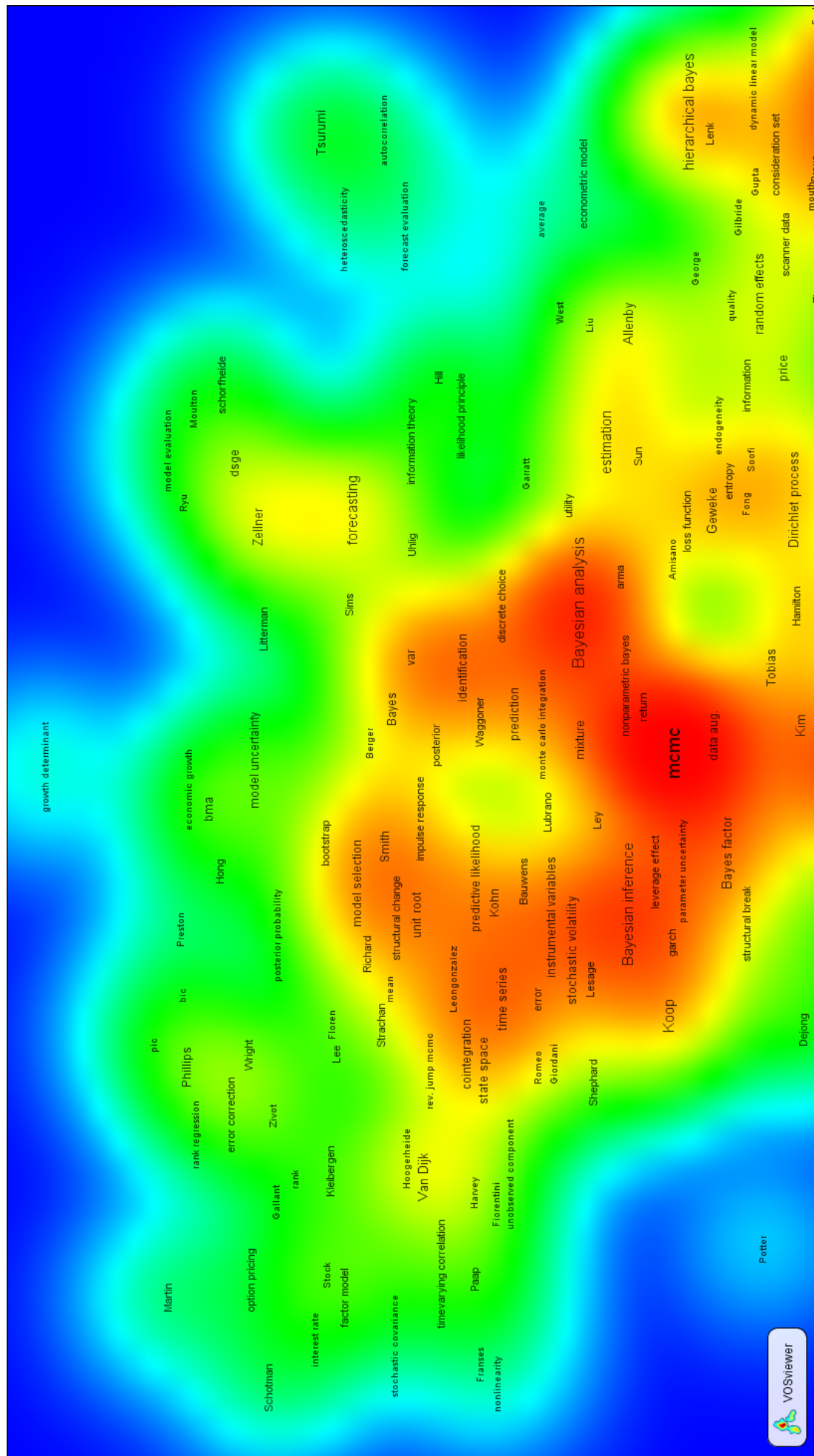
According to Figure 7, MCMC is central for Bayesian inference and Bayesian analysis. Macroeconomic and finance topics, such as stochastic volatility, time series, DSGE and option pricing, occur frequently in Bayesian econometrics. Marketing and choice models also occur frequently since the second dense area in the heatmap is centered around keywords such as pricing, choice model and advertising.

5 Subjects with substantial advances of Bayesian econometrics

In this section we distill three subjects from the connectivity analysis from section 4 where Bayesian econometrics has shown tremendous progress by itself and also compare it to the frequentist approach.

⁶We note that we leave some authors, such as Atkinson, Dorfman, Gelfand, Griffith and Trivedi, outside Figure 7 for visualization purposes. Despite a high number of papers by these authors, our clustering method separates these authors from the central part of the heatmap, most probably due to the diversity of the keywords in these authors' papers.

Figure 7: Connectivity of subjects and authors in papers in leading journals



These refer to the computational revolution, flexible structures and unobserved component models in macroeconomics and finance, and hierarchical structures and choice models in microeconomics and marketing.

The Computational Revolution

Applicability of Bayesian methods in econometric analysis relies heavily on the feasibility of the estimation of models. For most econometric models of interest the posterior distribution is not of a known form like the normal one and analyzing this distribution and its corresponding model probability using analytical methods or deterministic numerical integration methods is infeasible. Stochastic simulation methods, known as Monte Carlo (MC) methods, have been very useful for tackling these problems. One may characterize this as a ‘Computational Revolution’ for Bayesian inference leading to statements like ‘Monte Carlo saved Bayes’. The popular Markov chain Monte Carlo method, known as Gibbs sampling, contributed in particular to this. Therefore ‘Gibbs saved Bayes’ is a more appropriate statement.

There are at least three features of MC simulation techniques that make it attractive for Bayesian inference:

1. Given random drawings from the posterior of structural parameters, posterior densities and/or probabilities of regular functions of these parameters are also directly obtained;
2. There has been tremendous progress in the construction of indirect sampling methods where random draws are generated from an approximation of the posterior and a correction step is included to account for this;
3. Econometric models where the likelihood contains an integral that can be evaluated by MC methods can be naturally analyzed with simulation based Bayesian inferential methods that also make use of MC draws.

We elaborate on these features as follows. First, there exists the traditional one of being able to directly simulate a nonlinear function of parameters of a model. Obvious examples are: Given a set of generated parameter draws from a posterior of a structural model one can directly evaluate the distribution of a forecast and the distribution of a multiplier in order to study forecast properties and the uncertainty of policy effectiveness; and given a set of generated parameter draws from the posterior of a dynamic econometric model one can directly obtain the distribution of the eigenvalues to study the stability of that system and the random walk nature of the process. Early examples of implied dynamic features of trends and cyclical properties of an estimated model - using parameter draws from the posterior - are presented in Van Dijk and Kloek (1980) for US data and by Geweke (1988) for

nineteen OECD countries. Compared to frequentist methods, the Bayesian approach has the advantage that there is no need for plug-in estimates with a delta method added in order to evaluate estimation accuracy in an approximate manner.

A second, more methodological feature is how to obtain these parameter draws from models where the posterior is not of a known form and it is not known how to generate draws directly from the posterior. In fact, apart from direct simulation all sampling methods are indirect methods using approximations to the posterior density that are labeled: *importance* densities or *candidate* densities. Rejection sampling, introduced by von Neumann (1951), was the first method that was used widely, for instance for generating normal random variables, see Marsaglia and Bray (1964). Importance Sampling (IS), see Hammersley and Handscomb (1964), was introduced in Bayesian inference by Kloek and Van Dijk (1975) and published in Kloek and Van Dijk (1978), further developed by Van Dijk and Kloek (1980, 1985) and given a complete detailed treatment in Geweke (1989). Importance sampling is attractive since the generated draws are independently and identically distributed (IID) and this gives relatively easy ways to assess the numerical accuracy of the MC estimators using the Law of Large numbers and the Central Limit Theorem. However, one has to deal with generated *weighted* draws and the construction of an importance functions in high dimensions that is both appropriate (the weight function has bounded variance) and efficient (computations can be done in reasonable time) is not always trivial. The theory of Markov chain Monte Carlo (MCMC) was developed by Metropolis et al. (1953) and Hastings (1970), and extended in several influential papers such as Tierney (1994). This simulation method generated random draws (and not weighted draws). Although the evaluation of numerical accuracy is more involved than with IS, the MCMC procedures became the popular ones. A major pioneering advance in this first computational revolution is Gibbs sampling developed in Geman and Geman (1984) and extended in Tanner and Wong (1987) and Gelfand and Smith (1990). See Robert and Casella (2004) for a recent and detailed discussion on the Gibbs sampling method and its extensions. It is interesting to observe that in recent Bayesian econometric analysis, MCMC methods also struggle with finding appropriate and effective candidate densities for models where the posterior is multimodal and rather irregular with ridges in the surface due to weak identifiability.

The use of sampling methods turned out to be crucial for a third feature of Monte Carlo. Limited dependent variable models including Probit and Tobit models in panel data and unobserved component models, in particular, State Space models in macroeconomic time series became popular due to their added flexibility in describing nonlinear data patterns. However, these models have an integral in the likelihood that refers to the underlying unobserved continuous data for the limited dependent variable models and unobserved state for the state space models. The success for Bayesian simulation methods has been that the MC methods already used for integration in the parameter space can easily be

extended and are the natural technical tools to also integrate these unobserved data and states. Basic papers in these fields are Chib (1992); Albert and Chib (1993); De Jong and Shephard (1995); Chib and Greenberg (1996). Panel data models with random effects became a interesting field of research for simulation based Bayesian methods, see Chamberlain (1984); Lancaster (2000, 2002); Hirano (2002).

Importance of Hardware Developments

These three features of Monte Carlo contributed greatly to the development of Bayesian econometrics, however, Monte Carlo became operational only with the improvements in the hardware of computing power, i.e. how fast a computer can perform an operation. The issue of computing power is central in econometric analysis in general, but it is even more central to Bayesian econometrics when the MC methods are applied. The improvements in computing power since the 1970s are clearly substantial, but a recent improvement has been observed with the introduction of clusters of computers, super computers and the possibility of performing operations in Graphics Processing Units (GPUs). These computing power improvements show tremendous possibilities for handling complex problems and big data, see e.g. Aldrich et al. (2011) for a non-Bayesian approach. So far, they have been adopted in the Bayesian econometrics literature in a few cases but they show tremendous potential. Using such computing power efficiently requires usually careful engineering of and/or modifications in the posterior sampler. Certain sampling methods, such as the importance sampler, are naturally suited for efficient use of computational power, see Cappé et al. (2008) for a discussion. A recent study specifically focusing on enabling Bayesian inference using the GPU is Durham and Geweke (2013). For a case where parallel computing is used to combine information from many models for improved forecasting, we refer to Casarin et al. (2013).

Flexible structures, unobserved components models and data augmentation in macroeconomics and finance

As mentioned above, unobserved component models constitute a field in econometrics where Bayesian inference is heavily used. We focus on the state space models using macroeconomic and financial time series. The reason for the extensive use of Bayesian methods in this context is that simulation based Bayesian inference allows for much flexibility in the model structure as well as in the distributional assumptions. Flexible nonlinear structures can be modeled by introducing an extra latent space in such a way that the conditional structure of the model is linear given this unobserved state, see the local level model from Harvey (1990). Then from an estimation point of view, since the unobserved patterns underlying the behavior of observables need to be integrated out of the model, simulation based Bayesian integration methods can be used for inference and are very suitable for this class of models. That is, from the inference point of view, Bayesian inference takes the uncertainty of the unobserved

patterns into account while estimating the model parameters. This is an important issue where the frequentist approach is more restrictive since the unobserved patterns are estimated conditional on the estimates of the model parameters (one takes the mode of the distribution rather than the whole distribution). Carlin et al. (1992) provide an exposition of the estimation methodology based on simulation to estimate the unobserved components and the model parameters jointly. Shortly after, Jacquier et al. (1994) show how an exact inference, unlike the quasi maximum likelihood approximation, can be obtained for the stochastic volatility models, a popular class of models in finance for modeling time varying volatility, using a similar approach. While the basic Bayesian inference principle remains unchanged, more efficient simulation algorithms are proposed in Carter and Kohn (1994), Frühwirth-Schnatter (1994), Carter and Kohn (1996), De Jong and Shephard (1995) and Koopman and Durbin (2000).

Although standard models using unobserved components allow for only continuous changes, models with discrete changes in parameters allowing for structural changes or discrete Markov processes are also feasible using Bayesian techniques. Gerlach et al. (2000) and Giordani and Kohn (2008), among others, provide efficient algorithms for obtaining Bayesian inference in case of such discrete changes in parameters. Interesting applications on regime analysis in economics are provided by Paap and Van Dijk (2003) and Sims and Zha (2006).

When the observed variables to be modeled using unobserved components do not follow the standard normal distribution or dependence structures in the model are not linear, other estimation strategies, denoted as Particle Filter or Sequential Monte Carlo techniques that approximate the target distribution to be estimated well, can be conducted. Bayesian inference backed up with advanced simulation algorithms have proved to be very useful in these circumstances, see for example Gordon et al. (1993), Pitt and Shephard (1999), Andrieu and Doucet (2002) and Andrieu et al. (2010). This type of inference is also the key ingredient of the volatility modeling in finance and micro founded macroeconomic models, among others, if the researcher does not resort to linear approximations to estimate the model. This makes it feasible to obtain exact online inference in these settings providing more accurate outcomes. Omori et al. (2007), Fernández-Villaverde and Rubio-Ramírez (2007) and Fernández-Villaverde and Rubio-Ramírez (2008) are some examples of this approach.

Hierarchical structures and choice models in microeconomics and marketing

Hierarchical models that refer to choice processes are a prominent research topic in recent decades due to the, often, unobserved heterogeneity of individual agents in commodity and labor markets. Flexibility in structure and distribution like the Dirichlet process are important features of the modeling process. Latent processes such as Probit models are used to describe unobserved components in models. Panel

data models are more and more used with scanner data giving rise to massive computing. Basic papers that deal with these issues are McCulloch and Tsay (1994); Rossi et al. (1996) and Hansen et al. (2006). More references are given in chapter 8 of the Handbook and in Rossi et al. (2005).

Econometric issues in this area are the presence of endogenous regressors, treatment effect problems, latent variables and many parameters, see Heckman and Vytlačil (1999) for the relation between treatment effect models, latent variable models, and the issue of endogeneity. Gibbs-based MCMC sampling are standard but new simulation based Bayesian techniques are being developed, see e.g. Frühwirth-Schnatter et al. (2004); Imbens and Rubin (1997) and Li et al. (2004). Recently, these models are developed using Dirichlet process priors in order to gain robustness of results and the inference relies on semiparametric Bayesian approaches, see Hirano (2002). Furthermore, shrinkage priors are regularly used in this class of models. It is expected that parallel processing will become important in this area of Bayesian research.

6 Issues of Debate

We next summarize three issues of research where Bayesian and non-Bayesian econometricians differ, and we discuss the advances in these research areas. Note that there is some overlap with the material of the previous section.

Endogeneity, Instrumental Variables, Prior Information and Model Evaluation

As stated in Section 2, early work in Bayesian econometrics focused on the simultaneous equations model with as prominent research issue *endogeneity* of a set of economic variables in a market or basic macroeconomic model. The technical issue is how correlation between a right hand side variable in an equation and the disturbance of that equation affects inference. Apart from the likelihood approach pursued in Cowles Commission monographs 10 and 14, another approach to this problem is the use of instrumental variables (IV), originally developed in the 1920s and 1930s by Wright (1928, 1934) which was followed by the work of Goldberger (1972). For details, see Stock and Trebbi (2003). One may characterize the instrumental variable (IV) regression model as a single equation Simultaneous Equations Model. Bayesian analysis of the latter model was introduced by Drèze (1976), see also Bauwens and Van Dijk (1990). However, the issue of endogeneity does also occur in microeconomic models. A prominent empirical example is the income-education effect of number of school years. Starting with Angrist and Krueger (1991), this literature was focused on the issue of measuring a treatment effect in a model with data from a randomized experiment. A relationship between the instrumental variable approach and the (local average) treatment effect literature was established by

Heckman and Robb (1985) and a Bayesian analysis presented in Imbens and Rubin (1997). Many other empirical applications using panel data models have occurred in the microeconometrics literature. A detailed analysis of the development of the literature in this field is beyond the scope of the present paper.

Specification of prior distributions in these model structures that leave the model information dominant became an active field of research. The choice of a reasonable informative or noninformative prior distribution is a crucial part of any Bayesian analysis and often subject to criticism by frequentist econometricians. However, sensible prior distributions provide valuable improvements in inference for many of the econometric issues that are hard to tackle. Hence, it is a blessing rather than a curse.

As argued above for the presented examples of empirically relevant models, data are often only weakly informative about the appropriate parameter values and may yield similar likelihood values with different parameter combinations, which is referred to as ‘the weak identification problem’. Weak identification is the common characteristic in models with nearly reduced rank, which occurs in simultaneous equations models, instrumental variable regression models, dynamic models with cointegration and in factor models. Weak identification gives usually irregular behavior of the likelihood; see papers by Bauwens and Van Dijk (1990), Kleibergen and Van Dijk (1994), Kleibergen and Van Dijk (1998) and Hoogerheide, Kaashoek and Van Dijk (2007). In such cases, two lines of research for specifying prior information took place. One is to assign reasonable informative priors from other studies or other evidence in order to alleviate the identification problem. This is often used in micro founded macroeconomic studies where priors are constructed using economic theory or from other studies or from micro data such as households surveys, see for example Del Negro and Schorfheide (2008). Another well known example of prior information is the use of reasonable regions in the parameter space, restricted by inequality conditions. Frequentist inference is extremely difficult using such restrictions. Examples of Bayesian inference where the implied prior on the range of an economics multiplier or a prior on the length of the period of oscillation of the business cycle yield plausible restrictions are given in Van Dijk and Klok (1980) and Harvey, Trimbur and Van Dijk (2007).

Another line of research has been the specification of diffuse priors in order to let the likelihood information dominate strongly. One important issue in this context is the existence of the posterior distribution and its first and higher order moments in case of an almost flat likelihood. Early Bayesian literature suggests that the posterior densities in the class of SEM and IV models under flat priors may be improper, see e.g. Zellner, Bauwens and Van Dijk (1988) and Bauwens and Van Dijk (1990) and Kleibergen and Van Dijk (1998). It is recently shown in Zellner, Ando, Baştürk, Hoogerheide and Van Dijk (2014) that the posterior distribution and its higher order moments exist for improper flat priors depending on the number of instrumental variables present in the model. Due to these identification

issues in IV regression models, the use of alternative prior structures, such as the Jeffrey’s prior, were also proposed. Kleibergen and Van Dijk (1998), Martin and Martin (2000) and Hoogerheide, Kaashoek and Van Dijk (2007) show examples of models in macro- and microeconomics. More recent advances in Bayesian estimation of these models are the introduction of semiparametric models by Conley et al. (2008) and Florens and Simoni (2012) among others, and efficient posterior sampling algorithms as in Zellner, Ando, Baştürk, Hoogerheide and Van Dijk (2014). For a more detailed discussion of Bayesian approaches to IV models, we refer to Lancaster (2004) and Rossi et al. (2005). For a panel data study reference is given to Hirano (2002).

Prior distributions are also the key ingredients for flexible modeling strategies in Bayesian analysis. This is especially important for density estimation. Bayesian nonparametric analysis is one evolving area where such prior distributions or processes are heavily used. While some of the theoretical achievements were already accomplished during 1970s, see Ferguson (1973), Antoniak (1974) for example and Sethuraman (1994) for a more recent paper, extensive use of such priors were only possible with the advance of computing power. Faster simulation schemes, due to increased computing power, proved to be very useful for such complex analysis, see for example Escobar and West (1995), Neal (2000) and Walker (2007). Currently, many applications have emerged in different fields using flexible prior distributions, see for example Chib and Hamilton (2002); Hirano (2002); Griffin and Steel (2004); Jensen (2004); Jensen and Maheu (2010).

Model evaluation is a crucial ingredient of Bayesian inference in econometrics. Lindley’s paradox - or Bartlett’s or Jeffreys’ paradox; see Lindley (1957) and Bartlett (1957) - implies that one has to choose very carefully the amount of prior information compared to the amount of sample information when comparing alternative hypotheses on model structures with the intention to let the information from the data in the likelihood dominate that of the prior. Typically a naive or malevolent researcher could ‘force’ the posterior probability of a certain model M , the ‘restricted model’ in case of two nested models, to tend to go to unity by letting the priors in all alternative models tend to diffuse priors, thereby decreasing the marginal likelihoods of all alternative models, even if the particular model M does not make sense and poorly describes the data. In an attempt to make the posterior model probabilities ‘fair’, one could use predictive likelihoods instead of marginal likelihoods; see Gelfand and Dey (1994), O’Hagan (1995), and Berger and Pericchi (1996). However, the use of predictive likelihoods brings several questions and issues. First, one must choose the training sample and the hold-out sample. Examples of important questions are: How many observations are included in these samples? Is one training sample used or does one average over multiple (or all possible) training samples? In the latter case, what does one average – e.g., marginal likelihoods, logarithms of marginal likelihoods, Savage-Dickey Density Ratios or posterior model probabilities? Second, if one chooses to average results over

multiple (or all possible) training samples, then the computing time that is required for obtaining all Monte Carlo simulation results for all training samples may be huge. In other words, the Lindley paradox and the computation of predictive likelihoods enlarge the relevance of simulation methods that efficiently provide reliable and accurate results in case of non-elliptical credible sets. A suitable method must deal with large numbers of different non-elliptical shapes in a feasible computing time. For time series models computing the marginal likelihood for a random subsample implies that the estimation must be performed for an irregularly observed time series (with many ‘missing values’), which is typically only feasible using an appropriate formulation and estimation of a state space model. In future research computationally efficient and accurate simulation methods need to be developed here. We also refer to Sims (2005) for an approach to make use of dummy variables in this context.⁷

Dynamic inference, Nonstationarity and Forecasting

Inference in dynamic models constitute a second Bayesian issue subject to criticism by frequentist econometricians. As a start we summarize the perfect duality between Bayesian inference in the parameter space and frequentist inference in the sample space for the well-known class of the linear regression model $y = X\beta + \epsilon$. In both frequentist and Bayesian econometrics, the parameter β has a student- t density. However, the interpretations are different. A graphical illustration of the difference for this model is provided in the left panel in Figure 8. Table 1 presents a summary of the duality and differences between Bayesian and frequentist inference. In practice one finds that many empirical researchers are ‘closet’ Bayesians in interpreting the obtained value of a t-test as indicating possible positive strength of the empirical result. In the strict frequentist sense one can only reject a null hypothesis if there is sufficient evidence for it.

This equivalence breaks down for dynamic regression models. Bayesian and frequentist inference take then different routes. In stationary dynamic models, Bayesian econometrics suggests the student- t density for the parameters, but frequentist econometrics faces finite sample bias problems and finite sample frequentist densities of estimators that have usually no known properties. This divergence between the Bayesian and the frequentist inference is even more pronounced for nonstationary dynamic models, see Sims and Uhlig (1991) and for a clear survey Koop (1994). In the frequentist case, the inferential statement: ‘No falsification of the Null Hypothesis of a Unit Root leads to the acceptance of the Unit Root’ yields fragile and often incorrect conclusions, for example, if a break occurs in the series. That is, several alternatives may be more plausible.

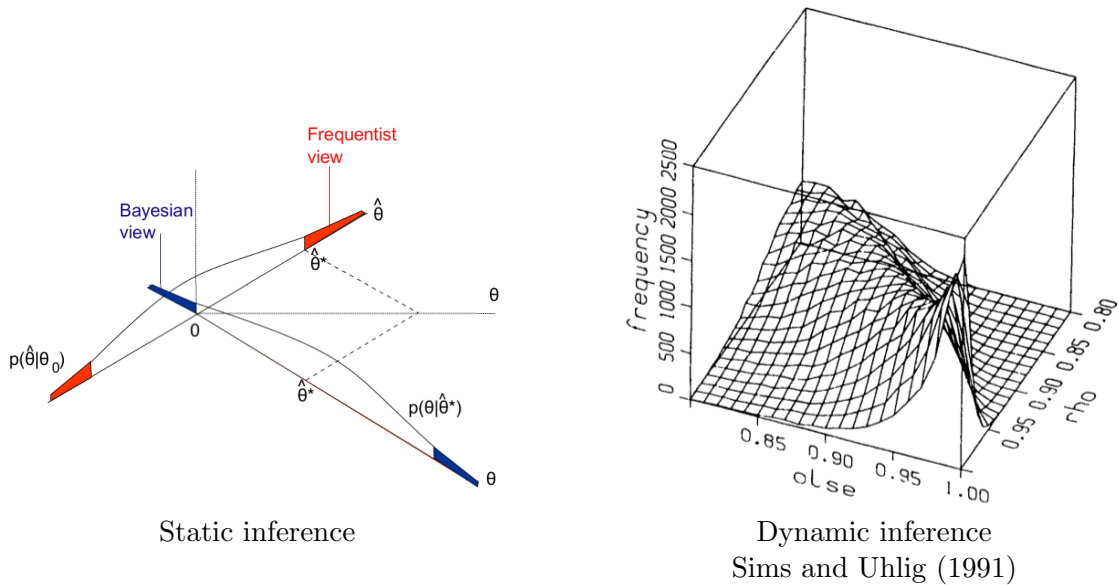
Sims and Uhlig (1991) and Schotman and Van Dijk (1991a,b) suggest that Bayesian inference for models with a unit root is more sensible, as well as much easier to handle analytically, than the fre-

⁷Part of this paragraph is taken from Van Dijk (2013a).

Table 1: Duality and differences between Bayesian and frequentist inference

| Frequentist inference | Bayesian inference |
|--|---|
| Parameters β are fixed unknown constants | Parameters β are stochastic variables. One defines a prior distribution on parameter space. |
| Data y are used to estimate β and check validity of postulated model, by comparing data with (infinitely large, hypothetical) data set from model. | Data y are used as evidence to update state of the mind: data transform prior into posterior distribution using the likelihood. |
| Frequentist concept of probability: Probability is the fraction of occurrences when the process is repeated infinitely often. | Subjective concept of probability: Probability is a degree of belief that an event occurs. |
| One can use the maximum likelihood estimator as an estimator of β . | One uses Bayes' theorem to obtain the posterior distribution of β . One can use $E(\beta y)$ or minimize a loss function to estimate θ . |
| R^2 is used to compare models. | Model comparison is carried out by using posterior odds ratio. |

Figure 8: Frequentist versus Bayesian econometrics



quentist confidence statements. Even under the assumptions of linearity and Gaussian disturbances, and even if conditioning on initial conditions is maintained, frequentist small-sample distribution theory for autoregressions is complicated. Frequentist asymptotic theory breaks discontinuously at the boundary of the stationary region. Therefore, the usual simple normal asymptotic approximations are not available. The likelihood function, however, is well known to be the same in autoregressions and non-dynamic regressions, assuming independence of disturbances from lagged dependent variables. Thus inference satisfying the likelihood principle has the same character in autoregressions whether or not the data may be non-stationary. The illustration of the likelihood for this autoregressive model in Sims and Uhlig (1991) is given in the right panel of Figure 8. Phillips (1991) stresses the fragility of Bayesian inference to the specification of the prior and warns against the mechanical use of a flat prior. Schotman and Van Dijk (1991b) approach this problem in a different more natural parameterizations of this model. Further, a unit root is not a testing problem in economics but a choice problem on the relative weights of two states of nature: the stationary and the nonstationary case. Schotman and Van Dijk (1991a) suggest to use posterior odds test (for the choice between a unit root model and an AR(1) stationary model) in this context. One can use these weights in evaluating forecasts and impulse response functions, see De Pooter, Ravazzolo, Segers and Van Dijk (2009).

The difference between Bayesian and frequentist econometrics is even more apparent for the case of multivariate dynamic models with possible nonstationarity. Kleibergen and Van Dijk (1993, 1994) propose a Bayesian procedure to inference and show that by using flat priors, the shape of the likelihood and marginal posteriors of the cointegration vectors are irregular behaved when certain parameters become weakly and in the limit non-identified. This problem also plagues standard frequentist inference. Kleibergen and Van Dijk (1994) analyze this problem by proposing the Information Matrix or Jeffreys' prior. Apart from the discontinuous asymptotic theory another key problem with the frequentist approach is that a sequential testing procedure is used to determine the number of stationary and the number of nonstationary relations, in other words in a sequential maneer the number of stable and unstable relations is determined while strictly speaking only falsification is possible in this testing approach. In the Bayesian approach weights can be evaluated for each member of the set of stable and unstable relations. Forecasts can be made and impulse responses evaluated with a weighted average of such relations using marginal and predictive likelihoods, see Wright (2008) and Strachan and Van Dijk (2013) for examples. We also refer to Koop (1991) and Koop and Korobilis (2013) for interesting applications.

Bayesian analysis has become a dominant forecasting and counterfactual analysis tool in recent decades, see Geweke and Whiteman (2006). There are four main reasons for this phenomenon. First, many of the complex, otherwise non-estimable, models can be estimated using simulation based

Bayesian methodology. Perhaps, the most important example of these models includes the class of structural micro founded macroeconomic models, such as Dynamic Stochastic General Equilibrium models, that are used both for policy analysis and for forecasting, see for example Smets and Wouters (2003, 2007), An and Schorfheide (2007). Currently, many of the central banks employ such models to obtain short and long term projections of the economy. An advantage of the Bayesian methodology is that it provides a solid statistical ground for efficient analysis using these structural models. As Bayesian inference provides the distribution of many key parameters that play a crucial role in economic analysis it is often used as a tool for counterfactual analysis. For instance, the questions such as ‘if quantitative easing were not conducted in US, would the course of the recession differ?’ could be answered by estimating relevant structural models. Bayesian analysis provides a statistically coherent tool for employing counterfactual analysis by forecasting under counterfactuals.

Second, prior distributions can play an integral part of the forecasting especially for the overparametrized models. Vector Auto Regression models (VAR) are major examples where Bayesian inference facilitates forecasting using the prior distributions for shrinking the parameters towards zero and thereby decreasing the dimensionality of the models. Decreasing the dimension of the overparametrized models using clever prior distributions has proved to be very useful in many applications. Prominent examples of this approach constitute Doan et al. (1984), Kadiyala and Karlsson (1997), Banbura et al. (2010). In macroeconomic forecasting the priors proposed by Doan et al. (1984) have become a standard tool among econometricians in academia and in other institutions such as central banks. In more general cases, many tailored priors are used for shrinkage of the model parameters towards zero and therefore they are efficiently used in variable selection when there are many candidate variables to select from. Prominent examples include, George and McCulloch (1993), Ishwaran and Rao (2005) and Park and Casella (2008) among others.

Third, Bayesian methodology takes the parameter uncertainty into account which may be of crucial importance in many applications. This enables researchers to obtain the entire predictive distribution rather than point forecasts based on the mode of the parameter distribution as in the frequentist approach. An important advantage of this feature is that different parts of the predictive distribution can be analyzed easily. This yields an obvious advantage for the analysis of various types of risk in finance and macroeconomics. A recent example is given in Basturk, Cakmakli, Ceyhan and Van Dijk (2013a) where the probability of deflation is evaluated for the US.

Fourth, the Bayesian methodology provides a natural and statistically solid way to take model uncertainty into account and to combine models to increase the predictive ability of many competing models. Bayesian model averaging technique provides one elegant way to do so, see for example Min and Zellner (1993), Fernández et al. (2001), Avramov (2002), Cremers (2002) and Wright (2008). Recent

advances in Bayesian model combination also allows to combine models where the model space is not complete implying that none of the competing models might be the true model. In that case, optimal combinations are proposed in Geweke and Amisano (2011, 2012); Durham and Geweke (2014). For a recent example where Sequential Monte Carlo is used to obtain density combinations from different model structures we refer to Billio, Casarin, Ravazzolo and Van Dijk (2013).

Vector autoregressive versus structural modeling

Given that in the early 1970s oil price shocks and high inflation affected macroeconomic time series both in levels and volatility, the existing classes of econometric models often based on Keynesian structures did not fit and forecast well. In the words of Christopher Sims these models were not ‘realistic’. In his 1980 paper in *Econometrica*, Sims advocated the use of Vector Autoregressive Models (VAR) to describe better the time series patterns in the data. One may characterize his work as: Sims against the ‘Econometric Establishment’. However pragmatic the VAR approach was, there quickly were discussions on the fact that the unrestricted VAR had the curse of parameter dimensionality or otherwise stated an over-parametrization danger. Several approaches to overcome this criticism were developed. One approach is to make use of shrinkage priors that were of great help in forecasting. This class of priors became known as the Minnesota prior from Doan et al. (1984). A useful alternative is the dummy based observation prior due to amongst others Kadiyala and Karlsson (1997) and Sims and Zha (1998), see Sims (2005) for a review. In the late 1990s structural economic priors Del Negro and Schorfheide (2004) came into existence parallel to the use of more structural VAR models like the DSGE model from Smets and Wouters (2007) and many other Structural VAR’s. This latter topic has been discussed before. Nowadays structural VARs with informative priors are used everywhere in macroeconomics in academia and professional organizations both for forecasting and policy analysis. Given the recent economic crisis it is clear that this class of models needs to be developed further to include financial sectors. A recent interesting paper on sign restrictions and structural vector autoregressions is Baumeister and Hamilton (2014).

7 Internal debates and influence of non-Bayesian papers

In this section, we group three topics that gave rise to internal and external debates in the field of Bayesian econometrics and we discuss briefly the links between them. First, the internal debate on the relative merits of objective versus subjective econometrics, next a description of communication mechanisms between statistics and econometrics and we end with indicating a few non-Bayesian papers which have had a substantial influence on the development of Bayesian econometrics in the past 30

Table 2: Objective versus subjective econometrics

| Objective | Subjective |
|--|--|
| Let the model speak: analyze the shape of the likelihood | Everything is personal |
| Scientific evidence dominates | Personal probabilities should be solicited |
| Experts opinion may fail | Experts opinions matter |
| Reach a large public | |

years.

Objective versus subjective econometrics

Probabilities are not physical quantities like weight, distance, or heat that one can measure using physical instruments. As De Finetti (1989) suggested, ‘Probabilities are a state of the mind’. Therefore, in general it follows that Bayesians are subjectivists and probabilities are personal statements. The objective part of Bayesian analysis is the use of Bayes rule as an information processing technique or more specifically as ‘a set of rules for transforming an initial distribution into an updated distribution conditional upon observations’, see Sims (2007). However, some Bayesian econometricians are more subjectivist/personal than others and this may affect their preference in reporting results of empirical analysis. As Hildreth (1963) argued: ‘Reporting the shape of the likelihood and its properties is an important task for a Bayesian econometrician.’ For examples of nontrivial and nonelliptical shapes of the likelihood function of a set of econometric models we refer to Kleibergen and Van Dijk (1994), Hoogerheide, Opschoor and Van Dijk (2012) and Basturk, Hoogerheide and Van Dijk (2013c) and the examples in Figure 1. The more subjectivist Bayesians will usually argue that very personal prior information is available and should be part of the the analysis.

This emphasis on the shape of the likelihood indicates that such researchers usually make use of diffuse priors and they are often referred to as ‘objectivist’. Their viewpoint is based on the idea that experts opinions may fail and/or that the complexity of modern econometric models make a subjective prior too difficult to specific. Their aim is to ‘let the model speak’ in reporting scientific evidence. This viewpoint, due to the more limited inclusion of personal or expert statements, reaches a large public. The more subjective viewpoint, on the other hand, argues that when experts opinions fail or when these are very different from the likelihood information, the likelihood will show this feature.

A brief summary of the differences between the more subjectivist and the more objectivist viewpoints is provided in Table 2. This classification of subjectivists and objectivists is clearly subjective. First, the argument that scientific evidence dominates results in the objective viewpoint assumes that data are collected objectively, which does not hold usually, see Press and Tanur (2012, ch. 1) for an example

on subjective influences on data collection and data interpretation. Second, the choice of a specific model is inevitably subjective, hence inference based on the shape of the likelihood is affected by this subjective information. Third, when the data in the posterior indicate that an adjustment of an initial hypothesis has to be made, then usually a new hypothesis is defined, a new model is chosen or new data are collected. The posterior belief in the early experiment becomes the ‘subjective’ understanding of the process in the next experiment and this learning process is inevitably partly subjective, see Press and Tanur (2012, ch. 10).

Thus the distinction between subjectivism and objectivism is rather based on the degree of subjectivity related to the approach how results of the analysis are reported since pure objectivity is hardly feasible.

One may classify Bayesian econometricians in three groups: there exist ‘true’ Bayesian econometricians who belong in the right hand column of Table 2; ‘instrumental’ and ‘pragmatic’ Bayesian econometricians who belong more in the left column of Table 2; and finally ‘closet’ Bayesian econometricians, who use regression outcomes and talk about ‘strongly significant’ t -values and ‘accept’ the null hypothesis. One major conclusion is that almost all researchers in empirical econometrics apply Bayesian techniques explicitly or implicitly nowadays!

Communication between Statistics and Econometrics

Statistics and Econometrics have had a difficult relationship, with several switches in the past 50 years, partly due to the fact that econometric models are high-dimensional while early statisticians often preferred a few parameters which are directly interpretable, like survival probability. Clearly, the recent advances in financial statistics require the analysis of multidimensional model structures. Early statistics was applied to economic time series while recent statistics is applied more to biology and finance and is becoming very computational. Econometrics is more model-oriented with a large number of parameters.

There have been attempts to construct bridges between statistics and econometrics. Among these are the Seminar on Bayesian Inference in Econometrics and Statistics (SBIES) that was pioneered by Arnold Zellner from 1970 onwards and now actively steered by Siddharta Chib⁸, the European Seminar in Bayesian Econometrics (ESOB) that started in 2010 by Herman K. van Dijk⁹ and the Economics, Finance and Business Section (EFAB) of the International Society of Bayesian Analysis (ISBA) that was started in 2013 by Mike West.¹⁰

Apart from these workshops that bridge the gap between statistics and econometrics, there exist

⁸apps.olin.wustl.edu/conf/sbies/Home/

⁹www.esobe.org/

¹⁰bayesian.org/sections/EFaB/bylaws

also several initiatives by Bayesian econometricians to bridge the gap between theory and practice. One series of meetings which take place on a regular basis is the series of workshops on ‘Methods and Applications for Dynamic Stochastic General Equilibrium Models’, at the Federal Reserve Bank of Philadelphia organized by Jesús Fernández-Villaverde, Giorgio Primiceri, Frank Schorfheide and Keith Sill, in the NBER work group on ‘Empirical Methods and Applications for DSGE Models’.

Influence of non-Bayesian papers

Bayesian and non-Bayesian econometrics are naturally connected through methodological issues as well as application areas. It is beyond the scope of this paper to cover this interconnection in detail, but we provide a summary of four topics where the influence of key non-Bayesian papers on the advances in Bayesian econometrics has been substantial.

The first topic concerns high-dimensional macroeconomic models with flexible model structures which in turn require parameter restrictions in order to facilitate inference. The importance of flexible model structures, replacing the earlier tightly parameterized Keynesian models was convincingly argued by Sims (1980) in his influential paper: ‘Macroeconomics and reality’. The large number of parameters in the proposed class of VAR models brought out the need to use informative priors and impose some structure on its parameters, see Sims (2008) for a discussion. In recent years, extended versions of standard macro models, such as DSGE models or other structural VAR models, were shown to track and forecast macroeconomic data, see Smets and Wouters (2003); Christiano et al. (2005). Inference, forecasting and policy analysis based on these models has become an important area in Bayesian econometrics partly due to several non-Bayesian papers of which Sims (1980) is the most influential one.

The second topic is the methodological connection between Bayesian and non-Bayesian inference in latent variables or unobserved components models. These models and their estimation using filtering methods, like the Gaussian or Markov filter, have been influential in both Bayesian and non-Bayesian approaches due to the possibility of incorporating time-varying parameters, measurement errors and missing observations through state space models and filtering of the states. Hamilton (1989) is a leading paper, applying Markov-switching models to macroeconomic series in order to describe nonstationarity and business cycle features. It may be classified as a semi-Bayesian paper. The topic of this paper drew particular attention in macroeconomic and financial analysis due to the possibility of estimating time-varying levels and trends in economic data. Given the filtered state variables, one can obtain point estimates of the unobserved state or the full conditional distribution of the state conditional on the data and model parameters. This conditional distribution enables straightforward Bayesian inference of these models, e.g. using the Gibbs sampler. More details are nowadays given in many books nowadays

of which we mention Hamilton (1994).

A third topic, which is related to the second one, concerns relaxation of the linearity and normality assumptions in standard state space models. These assumptions have been considered as severe restrictions in econometric analysis, particularly for applications in financial risk management, option pricing and financial econometrics, dealing with high-frequency data. The restrictions were relaxed by the contributions that appeared in several Bayesian and non-Bayesian papers, see e.g. West and Harrison (1997, ch. 13 and 15), Doucet (2004), Jazwinski (2007, ch. 9) and Durbin and Koopman (2012). A recent methodology to relax these assumptions is particle filtering, see Gordon et al. (1993); Kim et al. (1998), which relies on the approximation of the non-standard filtering densities using Importance Sampling or Metropolis Hastings methods, also see Creal (2012) for a review. A key non-Bayesian paper in this context is Pitt and Shephard (1999), which has had substantial influential on Bayesian econometrics. The sampling algorithm relies on Bayesian updates and a full conditional density of the unobserved states is obtained using the methodologies proposed.

The final link that we establish is built on particular data properties and flexible model structures in marketing choice models and treatment effect models in microeconomics. These models make often use of panel data consisting of large cross sections and small time series. Chamberlain's very complete survey of panel data models and its properties generated much subsequent research in this area, see Chamberlain (1984). Several other non-Bayesian papers stated important estimation issues in these models, particularly due to reliance on standard asymptotic properties, or problems with separability between observables and unobservables, see, for instance, Angrist and Hahn (2004). Due to the reported limitations in dealing with these data, several Bayesian econometric papers attempted to overcome them. One approach using different forms of prior information ranging from a propensity score to unobservables with instrumental variables is presented in Chamberlain (2011). Furthermore, non-Bayesian papers establishing a connection between these models and unobserved components models for treatment effects have been influential on the Bayesian econometrics literature, see Heckman and Vytlacil (1999); Abbring and Heckman (2007). This influence is partly due to the possibility to extend Bayesian inference methods to deal with these specific unobserved components models. We note that Bayesian analysis of treatment effects models is only relatively recently emerging literature, see the discussions in Heckman et al. (2014).

Bayesian analysis of marketing choice models, with applications to consumer demand analysis, are often used for decision making. This field has recently gained substantial attention and generated many interesting papers, see Rossi et al. (2005) for a comprehensive treatment of Bayesian methods in marketing. A major advance in this field is the use of flexible priors and model structure using finite and infinite mixtures of several classes of distributions. basic papers are due to Ferguson (1973) and

Blackwell and MacQueen (1973). Substantial advances in simulation based Bayesian methods in this field have recently been obtained. We refer the Frühwirth-Schnatter (2006), Geweke et al. (2011, ch. 4) and references cited there.

8 Decisions, Treatment and Policy Analysis

‘Knowledge is useful if it helps to make the best decisions’, Marschak (1953). This phrase from Marschak is eminently suitable to describe the potential of Bayesian econometrics to combine different sources of information on an econometric model as well as to trace consequences of estimation and model uncertainty to decision fields that are known as policy scenario analysis in macroeconomics, finance and marketing and treatment effects in microeconomics.

Apart from the world of academia, decision making under uncertainty is highly relevant in advanced professional organizations like Central Banks. In this context Alan Greenspan, former Chairman of the U.S. Federal Reserve System stated in 2004 that ‘In essence, the risk management approach to policy making is an application of Bayesian decision-making’, see Greenspan (2004) and for a detailed discussion of Greenspan’s statement, we refer to Zellner (2009).

Accurate and timely macroeconomic policy decisions need detailed understanding of the dynamics of the economy. Thus, estimation of economic dynamics using probabilistic models is crucial for policy decisions. This was already recognized by Tinbergen (1939) and formalized by Haavelmo (1943a,b, 1944). Economists and econometricians were in the 1970s only able to determine the impact of estimation uncertainty on policy in simple models. Brainard (1967) studied the distribution of the multiplier in a simple Keynesian model in order to determine the effectiveness of policy analysis in a situation of uncertainty about parameter values. His paper is clearly one of the earliest on that topic. An other example is presented in Van Dijk and Kloek (1980) where prior uncertainty about structural parameters in a simple Keynesian model was simulated through to the implied prior predictive value of the multiplier and the implied prior period of oscillation of the business cycle and in a next step this prior predictive information was combined, using simulation based Bayesian methods, with the likelihood information to obtain sensible posterior predictive results. Drèze in his 1970 presidential address, see Drèze (1972), asked attention from econometric researchers for a complete research program on econometric decision analysis, see also Rothenberg (1973, ch. 6). In the field of statistics for business and economics there was the well-known paper by Pratt et al. (1964). Some more details in this topic are provided in Zellner (2009).

There exist fundamental issues in this context. As mentioned in section 2, Sims (2012) listed unclear formal treatment of policy interventions as a major weakness in Haavelmo’s research program. Here

we discuss Sims' arguments. Since model parameters are treated as non-random (as opposed to their estimators), uncertainty about them cannot carry over to predictive distributions, limiting substantially decision making under model uncertainty. Secondly, policy behavior was not incorporated in Haavelmo's and subsequent Cowles Commission's econometric models. While for a policy maker, a random policy behavior equation in the model may seem odd, since policy decisions are known to the policy maker, it is random for the private sector and the econometrician.

Bayesian inference provides clear solutions to these issues. From a Bayesian point of view, uncertainty, i.e. what is 'random' and what is 'non-random' depends on what is 'observed' and what is not 'not-observed'. Model parameters, being unknown, do have a probability distribution that is updated by data information using Bayes' rule. This provides a clear probabilistic tool for efficient and real-time decision making under uncertainty. Further, Bayesian interpretation eliminates also the paradox of random policy behavior in economic modeling. That is, for a policy maker, who observes her own policy decisions, policy behavior is not a source of uncertainty but for the private sector or the econometrician, policy behavior is not observed and clearly contributes to the uncertainty she faces. These two notions can jointly exist from a Bayesian point of view. We also refer to Geweke (2005, ch. 1) for an introduction to the basic components of Bayesian decision making.

An important aspect of state-of-the-art policy analysis is that policy behavior has to be accommodated by an economic model where economic agents' decision making is specified. This modeling process is different in different fields of economics and depends on the typical decision that need to be taken.

In macroeconomics, the incorporation of consumer preferences and technology features is important apart from several other sources of constraints. This is required in order to analyze economic agents' decisions in response to changing government policies also known as the Lucas' critique, Lucas (1976). Early macroeconomic models were real business cycle (RBC) models, see Kydland and Prescott (1982). Fully articulated models, however, require more parameters ranging from those that relate to policy decisions to ones that refer to economic agents' decisions. An early inferential method for RBC models, called 'calibration', implies plugging some plausible values for parameters. Once the parameter values have been set up, artificial data are generated using the model and the model fit can be evaluated by checking whether the artificial data capture stylized real data facts. These models have evolved to the so-called Dynamic Stochastic General Equilibrium (DSGE) models incorporating several mechanisms, rather than only exogenous shocks, for generating business cycle fluctuations, see Christiano et al. (2005) and Smets and Wouters (2002, 2007). Bayesian inference techniques provide nowadays tools for estimation rather than calibration.

Some recent policy issues in the field of macroeconomics are: relative importance of monetary

versus fiscal policy, inflation targeting, possibly ineffective monetary policy at the boundary of near zero interest rates, regime shift effects and finally how to determine whether ‘luck or policy’ is relevant. For more details and references we refer to the chapter on macroeconometrics in the Handbook.

In the field of finance behavioral assumptions of economic agents are naturally linked to how to deal with predictability, in particular that of uncertainty and risk. The connection between forecasting and decision making has always been a central topic in finance. More specifically we mention: economic relevance of prediction of asset returns and its effect on portfolio management, measurement of volatilities for Value-at-Risk, option pricing and modern algorithmic trading as recent topics of simulation based Bayesian econometric research. In this context the issue of efficient model combinations that are treated in the same way as dynamic portfolio analysis becomes more and more important, see below and we refer to Chapter 9 on Bayesian Methods in Finance in the the Handbook.

In the field of microeconomics, decision making based on marketing choice models is an important area where the above listed advantages of the Bayesian approach are applicable. The relatively recent availability of large data sets and the proposal of flexible model structures to capture heterogeneity in the behavior of consumers give researchers and practitioners a stimulus to take additional advantage of the Bayesian approach, namely the ease of computation using advanced computational methods, see Rossi et al. (2005). Several Bayesian papers now propose methods relating decisions on product design to optimization of profitability. Here we refer to Chapter 8 on Bayesian Applications in Marketing in the Handbook.

Despite their potential applicability, Bayesian methods have, until recently, not been widely used in the field of microeconometrics, dealing with structural models for labor, health and education issues, see Zellner (2009) for a background analysis. The Bayesian approach does provide a clear solution to incorporating parameter and model uncertainty in applications focusing on decision making based on predictions in panel data models dealing with treatment effects. This is recently becoming an important subject of research and for an early paper we refer to Rubin (1978). Several more recent papers, e.g. Imbens and Rubin (1997); Li et al. (2004) among others, provide important contributions on the Bayesian analysis of treatment effects models. For a very recent study on Bayesian estimation of mean treatment effects, including a summary of literature on Bayesian estimation of treatment effects, we refer to Heckman et al. (2014). For more details, see also Chapter 6 on Bayesian Methods in Microeconometrics in the Handbook.

In some cases, the predictions of the agents may be revealed not by means of an econometric models, but in the form of prediction markets, where contracts based on the predictions of economic agents are traded. Berg et al. (2010) derives a probability distribution using Bayesian nonparametric methods for the specific event of US presidential elections that is consistent with the prediction market prices of

these contracts. As predictive distributions are the key elements of the formal decision making, this also provides a tool to summarize the information in the prediction markets to use in decision making.

In summary, systematic analysis of including parameter uncertainty and model uncertainty into decision analysis requires both clear economic structure containing behavior of economic agents and sophisticated stochastic simulation algorithms in conjunction with recent hardware developments. The latter make computations feasible in reasonable time. This topic of specifying relevant economic structure and novel simulation based Bayesian decision analysis is very much in development and our list of topics and papers is surely incomplete.

Finally, we discuss the topic of model selection and model combination as inputs for decision making. Standard Bayesian analysis makes use of the Bayes factor, defined as the ratios of marginal likelihoods of two competing models. It is well known that this is sensitive to the choice of prior distributions. An attractive more robust alternative is the predictive likelihood, see Geweke and Amisano (2010, 2011). If the data is indeed generated by a specific model then this approach gives an guidance towards the ‘best’ model. More realistic appears to be the model combination concept with information from different sources.

However, decision making can also be conveyed by combining competing models rather than choosing the best. While Bayesian model averaging indicates a statistically basic framework of model combination, still the efficient combination of models is an active area of research. We sketch a few lines of research here by summarizing recent papers. Durham and Geweke (2014) propose to use optimal prediction pools proposed in Geweke and Amisano (2011) to combine predictive models of asset returns which is closely related to decision under uncertainty for a risk averse decision maker. Moreover, such a combination also admits that all of the models considered may be false, which is not the case for example in the standard Bayesian model averaging. An and Schorfheide (2007) use different but similar DSGE models and focus on the potential difficulties as model misspecification, identification and multimodality of the parameter distributions. Such comparisons reveal whether the restrictions imposed are correctly specified. The authors combine DSGE model based priors with VAR’s to form the so-called DSGE-VAR models. Del Negro and Schorfheide (2009) focus on monetary policy analysis using potentially misspecified DSGE models. Leeper et al. (1996) argue that models for ‘policy analysis’ and ‘forecasting’ are not sharply distinct. These authors show that size effects attributed to shifts in monetary policy vary across specifications of economic behavior. A robust conclusion, common across several models, is that a large fraction of the variation in monetary policy instruments is attributable to systematic reaction by policy authorities to the state of the economy.

As final case study relating to a model combination where individual models are false, we mention Billio et al. (2013) where a financial experiment is performed on investing in risky and risk free assets

using a model combination that constitutes of a random walk model on stock returns and a forecasting model from professional forecasters. Professional investors in the stock market would have encountered substantial losses when they had followed the forecasting model of only the professional forecasters while investors that used a model combination of survey forecasters and a random walk would be obtained much better results. Yet, the authors show evidence of incompleteness of the model combinations by looking at posterior residual analysis. The topic of model incompleteness, be it a single model or a model combination, is a challenging area of research, see Geweke (2010) for more details.

9 Personal predictions

Inspired by the path that Bayesian econometrics has followed for the last half century, we end this paper by presenting the authors' personal expectations of important subjects for the future of Bayesian econometrics in the 21st Century. One prediction is that 'the second computational simulation revolution' where efficient information distillation from 'big data' with sophisticated Bayesian methodology using parallelization techniques is going to play an important role. Another topic that is predicted to gain popularity is complex economic structures with nonlinearities and complete predictive densities. A third topic that is expected to have importance in the future is the analysis of implied model features, such as risk or instability due to diverging ratios, and decision analysis. Finally, model incompleteness, which refrains from the assumption that the true data generating process is in the defined model set, is predicted to be important topics in Bayesian econometrics, see Geweke (2010).

Besides focusing on important topics in Bayesian econometrics, we further predict that the influence of the Bayesian approach in the field of econometrics will continue to increase over time. This final prediction is in line with the statement 'Econometrics should always and everywhere be Bayesian' in Sims (2007). We refer to Sims (2007) for a detailed discussion on this topic and on how Bayesian approaches might become more prevalent in several areas of economic applications

As a final statement, we play a bit of a game. The citation numbers we analyzed for Bayesian papers can be related to the h-index of authors, a conventional measure for the impact of published work by scholars, see Hirsch (2005). Nowadays, the h-index is sometimes used in the career path and promotion stages of young researchers. We employ a simple simulation study to assess the expected h-index of a 'random Bayesian' publishing a predefined number of papers in the leading journals we consider. In order to assess this expected h-index, we consider a random sample of size J from the set of papers in our database and calculate the h-index. The average h-index for 1000 such random samples is used to approximate the expected h-index for an author with J publications in leading journals. For a young Bayesian econometrician who is the author of 5 such publications, we find that the h-index

is approximately 4, i.e. very high compared to the total number of publications of this author, and the expected number of citations for this author's papers is 334. For an author, coming up for tenure, with 12 publications in leading journals the expected h-index is approximately 9 with an expected number of citations of 765. For a very established author with 50 publications in these journals, the expected h-index is approximately 25, with an expected number of citations of 1644. The papers and the journals considered therefore have a considerable impact in the field, according to the calculated h-indexes. Conditional upon our data set we conclude that young Bayesian econometricians have a very good chance to follow an academic career successfully.

References

- Abbring JH, Heckman JJ. 2007. Econometric evaluation of social programs, part III: Distributional treatment effects, dynamic treatment effects, dynamic discrete choice, and general equilibrium policy evaluation. In Heckman JJ, Leamer EE (eds.) *Handbook of Econometrics*, volume 6, chapter 72. North Holland, Amsterdam: Elsevier, 5145–5303.
- Albert JH, Chib S. 1993. Bayesian analysis of binary and polychotomous response data. *Journal of the American statistical Association* **88**: 669–679.
- Aldrich EM, Fernández-Villaverde J, Ronald Gallant A, Rubio-Ramrez JF. 2011. Tapping the super-computer under your desk: Solving dynamic equilibrium models with graphics processors. *Journal of Economic Dynamics and Control* **35**: 386–393.
- Aldrich J. 1995. R. A. Fisher and the making of maximum likelihood 1912–22. Discussion Paper Series In Economics And Econometrics 9504, Economics Division, School of Social Sciences, University of Southampton.
- An S, Schorfheide F. 2007. Bayesian analysis of DSGE models. *Econometric Reviews* **26**: 113–172.
- Anderson TW. 1947. A note on a maximum-likelihood estimate. *Econometrica* **15**: 241–244.
- Anderson TW, Rubin H. 1949. Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annals of Mathematical Statistics* **20**: 46–63.
- Andrieu C, Doucet A. 2002. Particle filtering for partially observed Gaussian state space models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **64**: 827–836.
- Andrieu C, Doucet A, Holenstein R. 2010. Particle Markov chain Monte Carlo methods. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **72**: 269–342.
- Angrist J, Hahn J. 2004. When to control for covariates? panel asymptotics for estimates of treatment effects. *Review of Economics and Statistics* **86**: 58–72.
- Angrist JD, Krueger AB. 1991. Does compulsory school attendance affect schooling and earnings? *The Quarterly Journal of Economics* **106**: 979–1014.
- Antoniak CE. 1974. Mixtures of Dirichlet processes with applications to Bayesian nonparametric problems. *The Annals of Statistics* **2**: 1152–1174.
- Avramov D. 2002. Stock return predictability and model uncertainty. *Journal of Financial Economics* **64**: 423–458.
- Banbura M, Giannone D, Reichlin L. 2010. Large Bayesian vector auto regressions. *Journal of Applied Econometrics* **25**: 71–92.
- Bartlett MS. 1957. Comment on ‘a statistical paradox’ by D.V. Lindley. *Biometrika* **44**: 533–534.
- Basturk N, Cakmakli C, Ceyhan P, Van Dijk HK. 2013a. Posterior-predictive evidence on US inflation using extended Phillips curve models with non-filtered data. Forthcoming in *Journal of Applied Econometrics*.
- Basturk N, Cakmakli C, Ceyhan SP, Van Dijk HK. 2013b. Historical developments in Bayesian econometrics after Cowles Foundation monographs 10, 14. Tinbergen Institute Discussion Papers 13–191/III, Tinbergen Institute.
- Basturk N, Hoogerheide LF, Van Dijk HK. 2013c. Measuring returns to education: Bayesian analysis using weak or invalid instrumental variables. Unpublished manuscript.
- Baumeister C, Hamilton JD. 2014. Sign restrictions, structural vector autoregressions, and useful prior information. Unpublished manuscript.

- Bauwens L. 1991. The ‘pathology’ of the natural conjugate prior density in the regression model. *Annales d’Economie et de Statistique* **23**: 49–64.
- Bauwens L, Van Dijk HK. 1990. Bayesian limited information analysis revisited. In Gabszewicz J, Richard J, Wolsey L (eds.) *Economic Decision-Making: Games, Econometrics and Optimisation: Contributions in Honour of Jacques H. Drèze*, chapter 18. Amsterdam: North Holland, 385–424.
- Berg JE, Geweke J, Rietz TA. 2010. Memoirs of an indifferent trader: Estimating forecast distributions from prediction markets. *Quantitative Economics* **1**: 163–186.
- Berger JO, Pericchi LR. 1996. The intrinsic Bayes factor for linear models. In Bayarri MJ, Berger JO, Bernardo JM, Dawid AP, Heckerman D, Smith AFM, West M (eds.) *Bayesian Statistics 5*. London: Oxford University Press, 25–44.
- Billio M, Casarin R, Ravazzolo F, Van Dijk HK. 2013. Time-varying combinations of predictive densities using nonlinear filtering. *Journal of Econometrics* **177**: 213–232.
- Blackwell D, MacQueen JB. 1973. Ferguson distributions via Pólya urn schemes. *The Annals of Statistics* **1**: 353–355.
- Bos CS, Mahieu RJ, Van Dijk HK. 2000. Daily exchange rate behaviour and hedging of currency risk. *Journal of Applied Econometrics* **15**: 671–696.
- Brainard WC. 1967. Uncertainty and the effectiveness of policy. *The American Economic Review* **57**: 411–425.
- Cappé O, Douc R, Guillin A, Marin J, Robert CP. 2008. Adaptive importance sampling in general mixture classes. *Statistics and Computing* **18**: 447–459.
- Carlin BP, Polson NG, Stoffer DS. 1992. A Monte Carlo approach to nonnormal and nonlinear state-space modeling. *Journal of the American Statistical Association* **87**: 493–500.
- Carter CK, Kohn R. 1994. On Gibbs sampling for state space models. *Biometrika* **81**: 541–553.
- Carter CK, Kohn R. 1996. Markov chain Monte Carlo in conditionally Gaussian state space models. *Biometrika* **83**: 589–601.
- Casarin R, Grassi S, Ravazzolo F, Van Dijk HK. 2013. Parallel sequential Monte Carlo for efficient density combination: The Deco Matlab toolbox. Tinbergen Institute Discussion Papers 13-055/III, Tinbergen Institute.
- Chamberlain G. 1984. Panel data. In Griliches Z, Intriligator MD (eds.) *Handbook of Econometrics*, volume 2 of *Handbook of Econometrics*, chapter 22. North Holland, Amsterdam: Elsevier, 1247–1318.
- Chamberlain G. 2011. Bayesian aspects of treatment choice. In Geweke J, Koop G, Van Dijk HK (eds.) *The Oxford Handbook of Bayesian Econometrics*, chapter 1. New York, NY: Oxford University Press, 11–39.
- Chernoff H. 1954. Rational selection of decision functions. Cowles Foundation paper 91, Cowles Commission for Research in Economics. Reprinted in *Econometrica*, 1954, 22 (4): 422–443.
- Chernoff H, Divinsky N. 1953. The computation of maximum likelihood estimates of linear structural equations. In Hood WC, Koopmans TC (eds.) *Studies in Econometric Method*. New Haven: Yale University Press, 236–302. Cowles Commission Monograph 14, chapter X.
- Chib S. 1992. Bayes inference in the Tobit censored regression model. *Journal of Econometrics* **51**: 79–99.
- Chib S, Greenberg E. 1996. Markov chain Monte Carlo simulation methods in econometrics. *Econometric theory* **12**: 409–431.

- Chib S, Hamilton BH. 2002. Semiparametric Bayes analysis of longitudinal data treatment models. *Journal of Econometrics* **110**: 67–89.
- Christiano LJ, Eichenbaum ME, Evans CL. 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* **113**: 1–45.
- Conley TG, Hansen CB, McCulloch RE, Rossi PE. 2008. A semi-parametric Bayesian approach to the instrumental variable problem. *Journal of Econometrics* **144**: 276–305.
- Creal D. 2012. A survey of sequential Monte Carlo methods for economics and finance. *Econometric Reviews* **31**: 245–296.
- Cremers KJM. 2002. Stock return predictability: A Bayesian model selection perspective. *Review of Financial Studies* **15**: 1223–1249.
- De Finetti B. 1989. Probabilism: A critical essay on the theory of probability and the value of science. *Erkenntnis* **31**: 169–223.
- De Jong P, Shephard N. 1995. The simulation smoother for time series models. *Biometrika* **82**: 339–350.
- De Pooter M, Ravazzolo F, Segers R, Van Dijk HK. 2009. Bayesian near-boundary analysis in basic macroeconomic time-series models. In Chib S, Koop G, Griffiths W, Terrell D (eds.) *Advances in Econometrics (Bayesian Econometrics)*, volume 23. Bingley: JAI press, 331–402.
- Del Negro M, Schorfheide F. 2004. Priors from general equilibrium models for VARs. *International Economic Review* **45**: 643–673.
- Del Negro M, Schorfheide F. 2008. Forming priors for DSGE models (and how it affects the assessment of nominal rigidities). *Journal of Monetary Economics* **55**: 1191–1208.
- Del Negro M, Schorfheide F. 2009. Monetary policy analysis with potentially misspecified models. *American Economic Review* **99**: 1415–1450.
- Doan T, Litterman R, Sims C. 1984. Forecasting and conditional projection using realistic prior distributions. *Econometric Reviews* **3**: 1–100.
- Doucet A. 2004. Sequential Monte Carlo methods. In Kotz S, Read CB, Balakrishnan N, Vidakovic B (eds.) *Encyclopedia of Statistical Sciences*, volume 12. New York, NY: John Wiley & Sons.
- Drèze JH. 1962. The Bayesian approach to simultaneous equations estimation. Technical report, The Technological Institute, Northwestern University. ONR Research Memorandum 67.
- Drèze JH. 1972. Econometrics and decision theory. *Econometrica* **40**: 1–18.
- Drèze JH. 1976. Bayesian limited information analysis of the simultaneous equations model. *Econometrica* **44**: 1045–1075.
- Drèze JH, Richard J. 1983. Bayesian analysis of simultaneous equation systems. In Griliches Z, Intriligator M (eds.) *Handbook of Econometrics*, volume 1 of *Handbook of Econometrics*, chapter 9. North Holland, Amsterdam: Elsevier, 517–598.
- Durbin J, Koopman SJ. 2012. *Time Series Analysis by State Space Methods*. 38. Oxford: Oxford University Press.
- Durham G, Geweke J. 2013. Adaptive sequential posterior simulators for massively parallel computing environments. Working Paper Series 9, Economics Discipline Group, UTS Business School, University of Technology, Sydney.
- Durham G, Geweke J. 2014. Improving asset price prediction when all models are false. *Journal of Financial Econometrics* **12**: 278–306.

- Escobar MD, West M. 1995. Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association* **90**: 577–588.
- Ferguson TS. 1973. A Bayesian analysis of some nonparametric problems. *The Annals of Statistics* **1**: 209–230.
- Fernández C, Ley E, Steel MFJ. 2001. Model uncertainty in cross-country growth regressions. *Journal of Applied Econometrics* **16**: 563–576.
- Fernández-Villaverde J, Rubio-Ramírez JF. 2007. Estimating macroeconomic models: A likelihood approach. *The Review of Economic Studies* **74**: 1059–1087.
- Fernández-Villaverde J, Rubio-Ramírez JF. 2008. How structural are structural parameters? In Daron Acemoglu KR, Woodford M (eds.) *NBER Macroeconomics Annual 2007*, volume 22, chapter 2. University of Chicago Press, 83–137.
- Fisher RA. 1912. On an absolute criterion for fitting frequency curves. *Messenger of Mathematics* **41**: 155–160.
- Fisher RA. 1922. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society A*: 309–368.
- Fisher RA. 1973. *Statistical Methods and Scientific Inference*. New York, NY: Hafner Press.
- Florens JP, Simoni A. 2012. Nonparametric estimation of an instrumental regression: A quasi-Bayesian approach based on regularized posterior. *Journal of Econometrics* **170**: 458–475.
- Frühwirth-Schnatter S. 1994. Data augmentation and dynamic linear models. *Journal of Time Series Analysis* **15**: 183–202.
- Frühwirth-Schnatter S. 2006. *Finite Mixture and Markov Switching Models: Modeling and Applications to Random Processes*. Springer Series in Statistics. New York, NY: Springer.
- Frühwirth-Schnatter S, Tuchler R, Otter T. 2004. Bayesian analysis of the heterogeneity model. *Journal of Business & Economic Statistics* **22**: 2–15.
- Gelfand AE, Dey DK. 1994. Bayesian model choice: Asymptotics and exact calculations. *Journal of the Royal Statistical Society. Series B (Methodological)* **56**: 501–514.
- Gelfand AE, Smith AFM. 1990. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association* **85**: 398–409.
- Geman S, Geman D. 1984. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence PAMI-6*: 721–741.
- George EI, McCulloch RE. 1993. Variable selection via Gibbs sampling. *Journal of the American Statistical Association* **88**: 881–889.
- Gerlach R, Carter C, Kohn R. 2000. Efficient Bayesian inference for dynamic mixture models. *Journal of the American Statistical Association* **95**: 819–828.
- Geweke J. 1988. The secular and cyclical behavior of real GDP in 19 OECD countries, 1957–1983. *Journal of Business & Economic Statistics* **6**: 479–486.
- Geweke J. 1989. Bayesian inference in econometric models using Monte Carlo integration. *Econometrica* **57**: 1317–39.
- Geweke J. 2005. *Contemporary Bayesian Econometrics and Statistics*. New York, NY: Wiley.
- Geweke J. 2010. *Complete and Incomplete Econometric Models*. Princeton, NJ: Princeton University Press.

- Geweke J, Amisano G. 2010. Comparing and evaluating Bayesian predictive distributions of asset returns. *International Journal of Forecasting* **26**: 216–230.
- Geweke J, Amisano G. 2011. Optimal prediction pools. *Journal of Econometrics* **164**: 130–141.
- Geweke J, Amisano G. 2012. Prediction with misspecified models. *American Economic Review* **102**: 482–486.
- Geweke J, Whiteman C. 2006. Bayesian forecasting. In Elliot G, Granger CWJ, Timmermann A (eds.) *Handbook of Economic Forecasting*, chapter 1. North Holland, Amsterdam: Elsevier, 3–80.
- Geweke JF, Koop G, Van Dijk HK (eds.) . 2011. *The Oxford Handbook of Bayesian Econometrics*. Oxford: Oxford University Press.
- Gilbert CL, Qin D. 2005. The first fifty years of modern econometrics. Working Papers 544, Queen Mary, University of London, School of Economics and Finance.
- Giordani P, Kohn R. 2008. Efficient Bayesian inference for multiple change-point and mixture innovation models. *Journal of Business & Economic Statistics* **26**: 66–77.
- Goldberger AS. 1972. Structural equation methods in the social sciences. *Econometrica* **40**: 979–1001.
- Gordon NJ, Salmond DJ, Smith AFM. 1993. Novel approach to nonlinear/non-Gaussian Bayesian state estimation. In *IEE Proceedings F (Radar and Signal Processing)*, volume 140. IET, 107–113.
- Greenspan A. 2004. Risk and uncertainty in monetary policy. *American Economic Review* **94**: 33–40.
- Griffin JE, Steel MFJ. 2004. Semiparametric Bayesian inference for stochastic frontier models. *Journal of Econometrics* **123**: 121–152.
- Haavelmo T. 1943a. The statistical implications of a system of simultaneous equations. *Econometrica* **11**: 1–12.
- Haavelmo T. 1943b. Statistical testing of business-cycle theories. *The Review of Economics and Statistics* **25**: 13–18.
- Haavelmo T. 1944. The probability approach in econometrics. *Econometrica* **12**: 1–115.
- Hamilton JD. 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* **57**: 357–384.
- Hamilton JD. 1994. *Time Series Analysis*, volume 2. Princeton, NJ: Princeton University Press.
- Hammersley JM, Handscomb DC. 1964. *Monte Carlo Methods*. London: Chapman & Hall.
- Hansen K, Singh V, Chintagunta P. 2006. Understanding store-brand purchase behavior across categories. *Marketing Science* **25**: 75–90.
- Harvey AC. 1990. *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.
- Harvey AC, Trimbur TM, Van Dijk HK. 2007. Trends and cycles in economic time series: A Bayesian approach. *Journal of Econometrics* **140**: 618–649.
- Hastings WK. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika* **57**: 97–109.
- Heckman JJ, Lopes HF, Piatek R. 2014. Treatment effects: A Bayesian perspective. *Econometric Reviews* **33**: 36–67.
- Heckman JJ, Robb RJ. 1985. Alternative methods for evaluating the impact of interventions: An overview. *Journal of Econometrics* **30**: 239–267.

- Heckman JJ, Vytlacil EJ. 1999. Local instrumental variances and latent variable models for identifying and bounding treatment effects. *Proceedings of the National Academy of Sciences* **96**: 4730–4734.
- Hildreth C. 1963. Bayesian statisticians and remote clients. *Econometrica* **31**: 422–438.
- Hirano K. 2002. Semiparametric Bayesian inference in autoregressive panel data models. *Econometrica* **70**: 781–799.
- Hirsch JE. 2005. An index to quantify an individual’s scientific research output. *Proceedings of the National Academy of Sciences of the United States of America* **102**: 165–169.
- Hood WC, Koopmans TC (eds.) . 1953. *Studies in Econometric Method*. New York, NY: John Wiley & Sons. Cowles Commission for Research in Economics, Monograph No. 14.
- Hoogerheide L, Opschoor A, Van Dijk HK. 2012. A class of adaptive importance sampling weighted EM algorithms for efficient and robust posterior and predictive simulation. *Journal of Econometrics* **171**: 101–120.
- Hoogerheide LF, Kaashoek JF, Van Dijk HK. 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**: 154–180.
- Hurwicz L. 1950. Bayes and minimax interpretation of the maximum likelihood estimation criterion. Cowles Commission Discussion Paper Economics 352, Cowles Commission for Research in Economics.
- Imbens GW, Rubin DB. 1997. Bayesian inference for causal effects in randomized experiments with noncompliance. *The Annals of Statistics* **25**: 305–327.
- Ishwaran H, Rao JS. 2005. Spike and slab variable selection: Frequentist and Bayesian strategies. *The Annals of Statistics* **33**: 730–773.
- Jacquier E, Polson NG, Rossi PE. 1994. Bayesian analysis of stochastic volatility models. *Journal of Business & Economic Statistics* **12**: 371–389.
- Jazwinski AH. 2007. *Stochastic Processes and Filtering Theory*. New York, NY: Courier Dover Publications.
- Jensen MJ. 2004. Semiparametric Bayesian inference of long-memory stochastic volatility models. *Journal of Time Series Analysis* **25**: 895–922.
- Jensen MJ, Maheu JM. 2010. Bayesian semiparametric stochastic volatility modeling. *Journal of Econometrics* **157**: 306–316.
- Kadiyala KR, Karlsson S. 1997. Numerical methods for estimation and inference in Bayesian VAR-models. *Journal of Applied Econometrics* **12**: 99–132.
- Keynes JM. 1939. Professor Tinbergen’s method. *Economic Journal* **49**.
- Keynes JM. 1940. Comment. *Economic Journal* **50**: 154–156.
- Kim S, Shephard N, Chib S. 1998. Stochastic volatility: Likelihood inference and comparison with ARCH models. *The Review of Economic Studies* **65**: 361–393.
- Kleibergen F, Van Dijk HK. 1993. Non-stationarity in GARCH models: A Bayesian analysis. *Journal of Applied Econometrics* **8**: 41–61.
- Kleibergen F, Van Dijk HK. 1994. On the shape of the likelihood/posterior in cointegration models. *Econometric Theory* **10**: 514–551.
- Kleibergen F, Van Dijk HK. 1998. Bayesian simultaneous equations analysis using reduced rank structures. *Econometric Theory* **14**: 701–743.

- Kloek T, Van Dijk HK. 1975. Bayesian estimates of equation system parameters: An unorthodox application of Monte Carlo. Econometric Institute Report 7511, Erasmus University Rotterdam.
- Kloek T, Van Dijk HK. 1978. Bayesian estimates of equation system parameters: An application of integration by Monte Carlo. *Econometrica* **46**: 1–19.
- Koop G. 1991. Cointegration tests in present value relationships: A Bayesian look at the bivariate properties of stock prices and dividends. *Journal of Econometrics* **49**: 105–139.
- Koop G. 1994. Recent progress in applied Bayesian econometrics. *Journal of Economic Surveys* **8**: 1–34.
- Koop G, Korobilis D. 2013. Large time-varying parameter VARs. *Journal of Econometrics* **177**: 185–198.
- Koop G, Poirier DJ, Tobias JL. 2007. *Bayesian Econometric Methods*, volume 7. Cambridge: Cambridge University Press.
- Koopman SJ, Durbin J. 2000. Fast filtering and smoothing for multivariate state space models. *Journal of Time Series Analysis* **21**: 281–296.
- Koopmans TC. 1945. Statistical estimation of simultaneous economic relations. *Journal of the American Statistical Association* **40**: 448–466.
- Koopmans TC (ed.) . 1950. *Statistical Inference in Dynamic Economic Models*. New York, NY: John Wiley & Sons. Cowles Commission for Research in Economics, Monograph No. 10.
- Kydland FE, Prescott EC. 1982. Time to build and aggregate fluctuations. *Econometrica: Journal of the Econometric Society* : 1345–1370.
- Lancaster T. 2000. The incidental parameter problem since 1948. *Journal of Econometrics* **95**: 391–413. ISSN 0304-4076.
- Lancaster T. 2002. Orthogonal parameters and panel data. *The Review of Economic Studies* **69**: 647–666.
- Lancaster T. 2004. *An Introduction to Modern Bayesian Econometrics*. Oxford: Blackwell.
- Leamer EE. 1973. Multicollinearity: A Bayesian interpretation. *The Review of Economics and Statistics* **55**: 371–380.
- Leamer EE. 1974. False models and post-data model construction. *Journal of the American Statistical Association* **69**: 122–131.
- Leamer EE. 1978. *Specification Searches: Ad Hoc Inference with Nonexperimental Data*. New York, NY: Wiley.
- Leamer EE. 1983. Let’s take the con out of econometrics. *American Economic Review* **73**: 31–43.
- Leamer EE. 1985. Sensitivity analyses would help. *American Economic Review* **75**: 308–313.
- Leeper EM, Sims CA, Zha T. 1996. What does monetary policy do? *Brookings Papers on Economic Activity* **27**: 1–78.
- Li M, Poirier DJ, Tobias JL. 2004. Do dropouts suffer from dropping out? estimation and prediction of outcome gains in generalized selection models. *Journal of Applied Econometrics* **19**: 203–225.
- Lindley DV. 1957. A statistical paradox. *Biometrika* **44**: 187–192.
- Lucas RJ. 1976. Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy* **1**: 19–46.

- Marsaglia G, Bray TA. 1964. A convenient method for generating normal variables. *SIAM Review* **6**: 260–264.
- Marschak J. 1953. Economic measurements for policy and prediction. In Hood WC, Koopmans TC (eds.) *Studies in Econometric Method*. New York, NY: John Wiley & Sons, 1–26. Cowles Commission Monograph 14, chapter I.
- Martin GM, Martin VL. 2000. Bayesian inference in the triangular cointegration model using a Jeffreys prior. *Communications in Statistics-Theory and Methods* **29**: 1759–1785.
- McCloskey DN, Ziliak ST. 1996. The standard error of regressions. *Journal of Economic Literature* **34**: 97–114.
- McCulloch RE, Tsay RS. 1994. Bayesian inference of trend- and difference-stationarity. *Econometric Theory* **10**: 596–608.
- Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. 1953. Equation of state calculations by fast computing machines. *The Journal of Chemical Physics* **21**: 1087–1092.
- Min Ck, Zellner A. 1993. Bayesian and non-Bayesian methods for combining models and forecasts with applications to forecasting international growth rates. *Journal of Econometrics* **56**: 89–118.
- Neal RM. 2000. Markov chain sampling methods for Dirichlet process mixture models. *Journal of Computational and Graphical Statistics* **9**: 249–265.
- O’Hagan A. 1995. Fractional Bayes factors for model comparison. *Journal of the Royal Statistical Society. Series B (Methodological)* **57**: 99–138.
- Omori Y, Chib S, Shephard N, Nakajima J. 2007. Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics* **140**: 425–449.
- Paap R, Van Dijk HK. 2003. Bayes estimates of Markov trends in possibly cointegrated series: An application to U.S. consumption and income. *Journal of Business & Economic Statistics* **21**: 547–563.
- Pagan A. 1987. Three econometric methodologies: A critical appraisal. *Journal of Economic Surveys* **1**: 3–24.
- Pagan A. 1995. Three econometric methodologies: An update. In George DAR, Roberts CL, Sayer S (eds.) *Surveys in Econometrics*. Oxford: Blackwell, 30–41.
- Park T, Casella G. 2008. The Bayesian Lasso. *Journal of the American Statistical Association* **103**: 681–686.
- Phillips PCB. 1991. To criticize the critics: An objective Bayesian analysis of stochastic trends. *Journal of Applied Econometrics* **6**: 333–364.
- Pitt MK, Shephard N. 1999. Filtering via simulation: Auxiliary particle filters. *Journal of the American Statistical Association* **94**: 590–599.
- Poirier DJ. 1989. A report from the battlefield. *Journal of Business & Economic Statistics* **7**: 137–139.
- Poirier DJ. 1992. A return to the battlefield. *Journal of Business & Economic Statistics* **10**: 473–474.
- Poirier DJ. 2006. The growth of Bayesian methods in statistics and economics since 1970. *Bayesian Analysis* **1**: 969–979.
- Pratt JW, Raiffa H, Schlaifer R. 1964. The foundations of decision under uncertainty: An elementary exposition. *Journal of the American Statistical Association* **59**: 353–375.
- Pratt JW, Raiffa H, Schlaifer R. 1995. *Introduction to Statistical Decision Theory*. Cambridge: MIT press.

- Press SJ, Tanur JM. 2012. *The Subjectivity of Scientists and the Bayesian Approach*, volume 775. New York, NY: John Wiley & Sons.
- Qin D. 1996. Bayesian econometrics: The first twenty years. *Econometric Theory* **12**: 500–516.
- Qin D. 2013. *A History of Econometrics: the Reformation from the 1970s*. Oxford: Oxford University Press.
- Raiffa H, Schlaifer R. 1961. *Applied Statistical Decision Theory*. Amsterdam: Harvard University Press.
- Richard JF. 1973. *Posterior and Predictive Densities of Simultaneous Equation Models*. Berlin: Springer Verlag.
- Robert CP, Casella G. 2004. *Monte Carlo Statistical Methods*. New York, NY: Springer Verlaag.
- Rossi PE, Allenby GM, McCulloch RE. 2005. *Bayesian Statistics and Marketing*. New York, NY: Wiley Series in Probability and Statistics.
- Rossi PE, McCulloch RE, Allenby GM. 1996. The value of purchase history data in target marketing. *Marketing Science* **15**: 321–340.
- Rothenberg TJ. 1963. A Bayesian analysis of simultaneous equation system. Econometric Institute Report 6315, Erasmus University Rotterdam.
- Rothenberg TJ. 1973. *Efficient Estimation with a Priori Information*. New York, NY: Yale University Press. Cowles Foundation Monograph No. 23.
- Rubin DB. 1978. Bayesian inference for causal effects: The role of randomization. *The Annals of Statistics* **6**: 34–58.
- Savage LJ. 1961. The subjective basis of statistical practice. Technical report.
- Schlaifer R. 1959. *Probability and Statistics for Business Decisions: An Introduction to Managerial Economics under Uncertainty*. New York, NY: McGraw-Hill.
- Schotman P, Van Dijk HK. 1991a. A Bayesian analysis of the unit root in real exchange rates. *Journal of Econometrics* **49**: 195–238.
- Schotman PC, Van Dijk HK. 1991b. On Bayesian routes to unit roots. *Journal of Applied Econometrics* **6**: 387–401.
- Sethuraman J. 1994. A constructive definition of Dirichlet priors. *Statistica Sinica* **4**: 639–650.
- Sims CA. 1980. Macroeconomics and reality. *Econometrica* **48**: 1–48.
- Sims CA. 2005. Dummy observation priors revisited. Technical report, Princeton University.
- Sims CA. 2007. Bayesian methods in applied econometrics, or, why econometrics should always and everywhere be Bayesian. Hotelling lecture, presented June 29, 2007 at Duke University.
- Sims CA. 2008. Making macro models behave reasonably. Technical report, Princeton University.
- Sims CA. 2012. Statistical modeling of monetary policy and its effects. *American Economic Review* **102**: 1187–1205.
- Sims CA, Uhlig H. 1991. Understanding unit rooters: A helicopter tour. *Econometrica* **59**: 1591–1599.
- Sims CA, Zha T. 1998. Bayesian methods for dynamic multivariate models. *International Economic Review* **39**: 949–968.
- Sims CA, Zha T. 2006. Were there regime switches in U.S. monetary policy? *American Economic Review* **96**: 54–81.

- Smets F, Wouters R. 2002. Openness, imperfect exchange rate pass-through and monetary policy. *Journal of Monetary Economics* **49**: 947–981.
- Smets F, Wouters R. 2003. An estimated dynamic stochastic general equilibrium model of the Euro area. *Journal of the European Economic Association* **1**: 1123–1175.
- Smets F, Wouters R. 2007. Shocks and frictions in US business cycles: A Bayesian DSGE approach. *American Economic Review* **97**: 586–606.
- Stock JH, Trebbi F. 2003. Retrospectives who invented instrumental variable regression? *The Journal of Economic Perspectives* **17**: 177–194. ISSN 08953309.
- Strachan RW, Van Dijk HK. 2013. Evidence on features of a DSGE business cycle model from Bayesian model averaging. *International Economic Review* **54**: 385–402.
- Tanner MA, Wong WH. 1987. The calculation of posterior distributions by data augmentation. *Journal of the American Statistical Association* **82**: 528–540.
- The Economist. 2004. Signifying nothing? *Economics Focus* January 31st, p. 63.
- Tierney L. 1994. Markov chains for exploring posterior distributions. *The Annals of Statistics* **22**: 1701–1728.
- Tinbergen J. 1939. *Statistical Testing of Business Cycle Theories. I: A Method and Its Application to Investment Activity. II: Business Cycles in the United States of America, 1919–1932*. Geneva: League of Nations.
- Tinbergen J. 1940. On a method of statistical business-cycle research; a reply (to Keynes). *The Economic Journal* **50**: 141–154.
- Van Dijk HK. 2013a. Bridging two key issues in Bayesian inference: The relationship between the Lindley paradox and non-elliptical credible sets. In Singpurwalla N, Dawid P, O’Hagan A (eds.) *Festschrift for Dennis Lindley’s Ninetieth Birthday*, volume 2. Blurb publishers, 511–530.
- Van Dijk HK. 2013b. The Keynes-Tinbergen debate on the relevance of estimating econometric models. *TSEconomist* **4**: 8–10.
- Van Dijk HK, Klok T. 1980. Further experience in Bayesian analysis using Monte Carlo integration. *Journal of Econometrics* **14**: 307–328.
- Van Dijk HK, Klok T. 1985. Experiments with some alternatives for simple importance sampling in Monte Carlo integration. In Bernardo JM, Degroot M, Lindley D, Smith AFM (eds.) *Bayesian Statistics*, volume 2. Amsterdam: North Holland, 511–530.
- Van Eck NJ, Waltman L. 2010. Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics* **84**: 523–538.
- von Neumann J. 1951. Various techniques used in connection with random digits. *Journal of Research of the National Bureau of Standards, Appl. Math. Series* **3**: 36–38.
- Walker SG. 2007. Sampling the Dirichlet mixture model with slices. *Communications in Statistics—Simulation and Computation* **36**: 45–54.
- Waltman L, Van Eck NJ, Noyons ECM. 2010. A unified approach to mapping and clustering of bibliometric networks. *Journal of Informetrics* **4**: 629–635.
- West M, Harrison J. 1997. *Bayesian Forecasting and Dynamic Models*. New York, NY: Springer-Verlag.
- Wright JH. 2008. Bayesian model averaging and exchange rate forecasts. *Journal of Econometrics* **146**: 329–341. Honoring the research contributions of Charles R. Nelson.
- Wright PG. 1928. *The Tariff on Animal and Vegetable Oils*. New York, NY: Macmillan.

- Wright S. 1934. The method of path coefficients. *Annals of Mathematical Statistics* **5**: 161–215.
- Zellner A. 1971. *An Introduction to Bayesian Inference in Econometrics*. New York, NY: Wiley.
- Zellner A. 2009. Bayesian econometrics: Past, present, and future. *Advances in Econometrics* **23**: 11–60.
- Zellner A, Ando T, Baştürk N, Hoogerheide L, Van Dijk HK. 2014. Bayesian analysis of instrumental variable models: Acceptance-rejection within direct Monte Carlo. *Econometric Reviews* **33**: 3–35.
- Zellner A, Bauwens L, Van Dijk HK. 1988. Bayesian specification analysis and estimation of simultaneous equation models using Monte Carlo methods. *Journal of Econometrics* **38**: 39–72.

Appendix

A Bayesian Papers in Leading Journals

Table A.1: Percentages of pages devoted to Bayesian papers in leading journals

| Years | RES | ReStat | Ectra | IER | ET | MS | ER | JE | JAE | JBES |
|-----------|------|--------|-------|------|-------|-------|-------|-------|-------|-------|
| 1978 | 0.00 | 0.00 | 2.92 | 0.00 | — | — | — | 6.27 | — | — |
| 1979 | 0.00 | 0.00 | 0.25 | 1.77 | — | — | — | 4.76 | — | — |
| 1980 | 0.00 | 0.00 | 0.00 | 5.12 | — | — | — | 7.17 | — | — |
| 1981 | 1.32 | 0.00 | 0.00 | 1.87 | — | — | — | 5.70 | — | — |
| 1982 | 0.00 | 0.97 | 0.75 | 1.60 | — | 0.00 | 12.46 | 10.69 | — | — |
| 1983 | 0.00 | 0.00 | 0.00 | 0.00 | — | 10.00 | 1.50 | 12.31 | — | 11.14 |
| 1984 | 0.00 | 1.81 | 1.40 | 0.00 | — | 0.00 | 44.93 | 10.97 | — | 4.32 |
| 1985 | 0.00 | 1.52 | 1.11 | 3.17 | 0.00 | 5.61 | 29.86 | 15.64 | — | 6.95 |
| 1986 | 5.54 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 4.89 | 4.08 | 15.80 |
| 1987 | 0.00 | 1.47 | 0.00 | 0.00 | 0.65 | 0.00 | 0.00 | 0.00 | 5.40 | 15.47 |
| 1988 | 0.00 | 0.00 | 0.00 | 2.38 | 0.00 | 0.00 | 0.00 | 13.57 | 0.00 | 5.52 |
| 1989 | 0.00 | 2.83 | 2.73 | 1.58 | 0.00 | 0.00 | 0.00 | 5.21 | 3.84 | 3.17 |
| 1990 | 0.00 | 1.10 | 0.00 | 0.00 | 0.00 | 4.66 | 0.00 | 3.66 | 0.00 | 5.39 |
| 1991 | 1.24 | 0.00 | 0.98 | 2.41 | 0.00 | 0.00 | 0.00 | 24.91 | 29.92 | 27.73 |
| 1992 | 0.00 | 0.00 | 1.63 | 0.00 | 2.92 | 0.00 | 3.49 | 7.95 | 5.47 | 18.62 |
| 1993 | 0.00 | 0.00 | 1.83 | 0.00 | 0.00 | 0.00 | 0.00 | 20.65 | 11.79 | 7.82 |
| 1994 | 0.00 | 5.07 | 1.76 | 0.00 | 33.13 | 0.00 | 10.75 | 9.47 | 3.10 | 24.60 |
| 1995 | 0.00 | 5.83 | 0.00 | 2.36 | 0.00 | 1.57 | 0.00 | 14.88 | 0.00 | 4.87 |
| 1996 | 0.00 | 0.65 | 3.35 | 0.00 | 8.87 | 18.53 | 0.00 | 13.34 | 5.34 | 8.60 |
| 1997 | 0.00 | 0.00 | 0.00 | 1.41 | 0.00 | 5.88 | 0.00 | 10.91 | 15.75 | 9.07 |
| 1998 | 3.79 | 8.01 | 0.00 | 3.32 | 8.74 | 3.76 | 5.48 | 13.04 | 10.09 | 11.37 |
| 1999 | 0.00 | 2.98 | 2.90 | 0.00 | 0.00 | 14.38 | 32.67 | 5.98 | 10.45 | 12.82 |
| 2000 | 0.00 | 1.41 | 0.00 | 0.00 | 2.11 | 12.56 | 0.00 | 18.05 | 16.40 | 26.41 |
| 2001 | 0.00 | 2.69 | 4.19 | 6.88 | 0.69 | 0.00 | 8.04 | 4.40 | 7.17 | 13.59 |
| 2002 | 1.96 | 0.00 | 0.75 | 0.00 | 4.25 | 19.71 | 4.03 | 9.19 | 8.81 | 4.20 |
| 2003 | 0.00 | 0.00 | 3.70 | 4.41 | 0.00 | 11.76 | 2.92 | 15.65 | 10.49 | 20.93 |
| 2004 | 0.00 | 0.00 | 5.94 | 4.61 | 0.00 | 12.68 | 0.00 | 16.57 | 6.73 | 21.22 |
| 2005 | 5.15 | 1.65 | 0.00 | 0.00 | 2.72 | 7.25 | 0.00 | 11.92 | 16.23 | 11.67 |
| 2006 | 2.47 | 0.00 | 1.56 | 0.00 | 0.00 | 15.03 | 31.43 | 6.14 | 12.37 | 11.44 |
| 2007 | 4.14 | 0.00 | 4.35 | 2.27 | 2.78 | 16.74 | 56.83 | 12.45 | 1.57 | 14.40 |
| 2008 | 0.00 | 0.00 | 0.00 | 0.00 | 3.07 | 23.15 | 9.87 | 10.62 | 22.73 | 20.32 |
| 2009 | 0.00 | 0.00 | 4.98 | 8.54 | 1.06 | 0.00 | 4.13 | 7.28 | 24.65 | 10.18 |
| 2010 | 5.85 | 0.00 | 0.90 | 0.00 | 0.86 | 22.16 | 9.44 | 7.06 | 24.22 | 21.97 |
| 2011 | 1.91 | 1.30 | 0.00 | 5.35 | 1.75 | 35.17 | 14.29 | 6.75 | 13.57 | 4.82 |
| 2012 | 0.00 | 4.49 | 2.10 | 4.82 | 0.00 | 13.65 | 6.70 | 13.44 | 18.32 | 13.42 |
| 2013 | 3.51 | 3.96 | 1.74 | 2.18 | 0.00 | 12.56 | 0.00 | 5.43 | 21.56 | 4.97 |
| 2014 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 28.34 | 40.65 | 5.78 | 22.61 | 6.62 |
| 1978-2014 | 0.96 | 1.18 | 1.40 | 1.70 | 2.45 | 8.94 | 9.98 | 9.91 | 11.18 | 12.48 |

The table presents the percentages of pages devoted to Bayesian papers in the journals for each year and average percentages for the period 1978-2014. The table is an extension of Table 2 in Poirier (1992). The numbers in red correspond to years with special issues. *Econometric Reviews*, *Econometric Theory*, *Journal of Applied Econometrics* and *Marketing Science* did not exist before 1982, 1985, 1986 and 1982, respectively. Average numbers of Bayesian pages only include years for which the journal existed. Journal abbreviations are as in Figure 2.

Table A.2: Citation information for papers in leading journals

| | Total | RES | ReStat | IER | Ecetra | ET | MS | ER | JAE | JE | JBES |
|---|-------|------|--------|------|--------|------|------|------|------|-------|-------|
| <i>All papers</i> | | | | | | | | | | | |
| Number of citations | 67891 | 3143 | 3104 | 2351 | 6000 | 1799 | 7590 | 3592 | 8103 | 21432 | 10777 |
| Number of cited papers | 969 | 16 | 28 | 29 | 34 | 30 | 121 | 61 | 125 | 329 | 178 |
| Number of Bayesian papers | 1020 | 16 | 29 | 29 | 34 | 31 | 129 | 82 | 130 | 336 | 206 |
| <i>Papers with at least 100 citations</i> | | | | | | | | | | | |
| Number of citations | 46895 | 2748 | 2392 | 1581 | 5147 | 1164 | 4922 | 2711 | 5340 | 13361 | 7293 |
| Number of Bayesian papers | 170 | 6 | 9 | 6 | 13 | 7 | 23 | 4 | 20 | 55 | 26 |
| <i>Papers with at least 400 citations</i> | | | | | | | | | | | |
| Number of citations | 20460 | 2086 | 808 | 408 | 3442 | 425 | 944 | 2603 | 1806 | 4376 | 3562 |
| Number of Bayesian papers | 31 | 2 | 1 | 1 | 5 | 1 | 2 | 3 | 3 | 8 | 5 |

The table presents citation information for each leading journal and for the leading journals jointly. Total number of cited papers is based on papers which are cited at least once. Journal abbreviations are as in Figure 2.

Table A.3: Average number of citations for papers in leading journals for 5-year intervals

| Years | RES | ReStat | Ecetra | IER | ET | MS | ER | JAE | JBES | JE |
|---|-----|--------|--------|-----|-----|-----|-----|-----|------|-----|
| <i>All papers</i> | | | | | | | | | | |
| 1978-1982 | 2 | 0 | 134 | 10 | — | 0 | 6 | — | 279 | — |
| 1983-1987 | 12 | 50 | 22 | 1 | 0 | 108 | 214 | 163 | 181 | 360 |
| 1988-1992 | 0 | 91 | 330 | 28 | 0 | 23 | 1 | 145 | 349 | 117 |
| 1993-1997 | 0 | 113 | 122 | 18 | 252 | 235 | 7 | 199 | 956 | 727 |
| 1998-2002 | 315 | 286 | 274 | 134 | 78 | 440 | 166 | 412 | 1011 | 411 |
| 2003-2007 | 229 | 26 | 205 | 137 | 12 | 406 | 292 | 258 | 1071 | 342 |
| 2008-2014 | 51 | 39 | 80 | 102 | 12 | 219 | 26 | 386 | 315 | 141 |
| 1978-2014 | 85 | 84 | 162 | 64 | 60 | 230 | 109 | 279 | 579 | 337 |
| <i>Papers with at least 100 citations</i> | | | | | | | | | | |
| 1978-1982 | 0 | 0 | 124 | 0 | — | 0 | 0 | — | 116 | — |
| 1983-1987 | 0 | 35 | 0 | 0 | 0 | 94 | 204 | 121 | 79 | 291 |
| 1988-1992 | 0 | 75 | 324 | 0 | 0 | 22 | 0 | 88 | 197 | 54 |
| 1993-1997 | 0 | 87 | 116 | 0 | 182 | 207 | 0 | 163 | 660 | 618 |
| 1998-2002 | 315 | 255 | 259 | 108 | 51 | 381 | 142 | 322 | 746 | 236 |
| 2003-2007 | 212 | 26 | 132 | 129 | 0 | 235 | 197 | 166 | 698 | 217 |
| 2008-2014 | 16 | 0 | 53 | 57 | 0 | 33 | 0 | 200 | 126 | 30 |
| 1978-2014 | 74 | 65 | 139 | 43 | 39 | 149 | 82 | 184 | 361 | 228 |

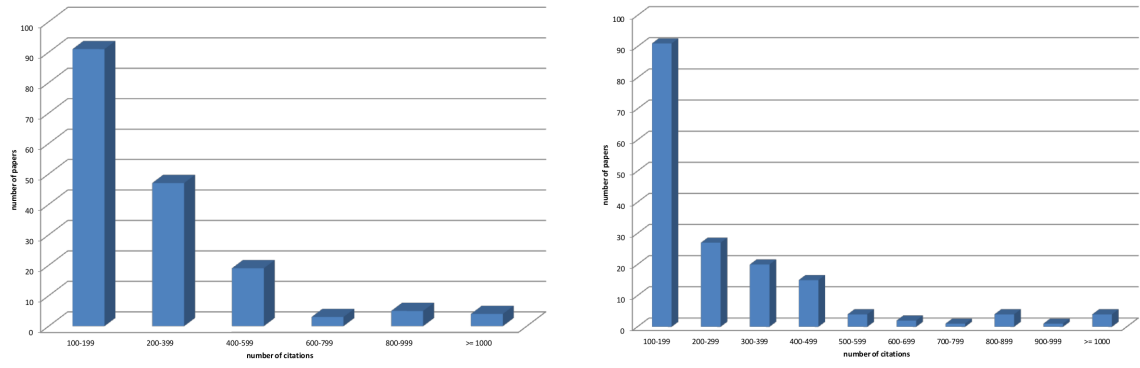
The table presents the average citation numbers of the Bayesian papers in the journals for the 5 year periods for all papers in leading journals (top panel) and a subset of papers with at least 100 citations (bottom panel). The table is an extension of Table 2 in Poirier (1992). Note that the period of observation is different for the following journals: Econometric Theory did not exist before 1985. Therefore the mean for the period 1983-1987 is taken over the periods 1985-1987. Journal of Applied Econometrics did not exist before 1986. Therefore, the mean over the years 1983-1987 is equal to the mean for 1986-1987. Econometric Reviews did not exist before 1982. Therefore, the mean for the period 1978-1982 is equal to the value in 1982. Marketing Science did not exist before 1982. So, the mean for the period 1978-1982 is equal to the value in 1982. Journal abbreviations are as in Figure 2.

Table A.4: Papers in leading journals with at least 100 citations

| Paper | Cite'd | Paper | Cite'd | Paper | Cite'd | Paper | Cite'd | Paper | Cite'd | Paper | Cite'd |
|-------------------------------|--------|--------------------------------|--------|------------------------------------|--------|---|--------|---|--------|-------|--------|
| Kloek and Van Dijk (1978) | 495 | Goodwin (1993) | 163 | Filardo and Gordon (1998) | 196 | Eraker (2001) | 394 | Rust and Verhoef (2005) | 394 | | 125 |
| Akaike (1981) | 580 | Albert and Chib (1993) | 446 | Phillips (1998) | 241 | Isionas (2002) | 241 | Ainslie et al. (2005) | 134 | | 146 |
| Leamer (1982) | 126 | Uhlig (1994) | 105 | Chib (1998) | 370 | Engle (2002) | 382 | Primiceri (2005) | 382 | | 627 |
| Monahan (1983) | 107 | Elliott and Stock (1994) | 118 | Allenby and Rossi (1998) | 508 | Chib et al. (2002) | 416 | Yu and Meyer (2006) | 416 | | 108 |
| Gersch and Kitagawa (1983) | 128 | Kleibergen and Van Dijk (1994) | 120 | So et al. (1998) | 174 | Moe and Fader (2002) | 109 | Chib et al. (2006) | 109 | | 198 |
| Hausser and Shugan (1983) | 469 | Phillips and Ploberger (1994) | 141 | Kim and Nelson (1998) | 344 | Hofstede et al. (2002) | 121 | Hansen et al. (2006) | 121 | | 100 |
| Doan et al. (1984) | 1020 | Van den Broeck et al. (1994) | 235 | Ainslie and Rossi (1998) | 227 | Kim et al. (2002) | 204 | Pesaran et al. (2006) | 170 | | 170 |
| Zellner and Rossi (1984) | 179 | Chib and Greenberg (1994) | 333 | Otok and Whiteman (1998) | 130 | Talukdar et al. (2002) | 231 | Das et al. (2007) | 540 | | 540 |
| Connolly and Hirschey (1984) | 173 | McCulloch and Rossi (1994) | 466 | Sims and Zha (1998) | 408 | Lancaster (2002) | 392 | An and Schorfheide (2007) | 875 | | 875 |
| Geweke (1986) | 242 | Andrews and Chen (1994) | 380 | Kim et al. (1998) | 1459 | Geweke et al. (2003) | 120 | Omori et al. (2007) | 168 | | 168 |
| Hamilton (1986) | 110 | Jacquier et al. (1994) | 1347 | Sims and Zha (1999) | 661 | Chernozhukov and Hong (2003) | 284 | Amisano and Giacomini (2007) | 193 | | 193 |
| Ohta and Griliches (1986) | 105 | Geweke et al. (1994) | 332 | Geweke (1999) | 708 | Jones (2003) | 355 | Del Negro et al. (2007) | 450 | | 450 |
| Clemen and Winkler (1986) | 230 | Chib and Greenberg (1995) | 170 | Zha (1999) | 112 | Hamilton (2003) | 964 | Heerde et al. (2007) | 110 | | 110 |
| Litterman (1986) | 859 | Chern et al. (1995) | 104 | Koop and Potter (1999a) | 114 | Jedidi et al. (2003) | 108 | Fernández-Villaverde and Rubio-Ramirez (2007) | 263 | | 263 |
| García-Ferrer et al. (1987) | 135 | Phillips (1996) | 187 | Koop and Potter (1999b) | 137 | Canova and Ciccarelli (2004) | 100 | Masanjala and Papageorgiou (2008) | 113 | | 113 |
| Geweke (1988) | 185 | Chib and Greenberg (1996) | 425 | Waggoner and Zha (1999) | 125 | Fernández-Villaverde and Rubio-Ramirez (2004) | 213 | Diebold et al. (2008) | 115 | | 115 |
| Geweke (1989a) | 1303 | Terrell (1996) | 139 | Kim and Nelson (1999) | 808 | Hansen et al. (2004) | 251 | Wright (2008) | 123 | | 123 |
| Zellner and Hong (1989) | 120 | Pfann et al. (1996) | 113 | Neelamegham and Chintagunta (1999) | 237 | Jacquier et al. (2004) | 339 | Mol et al. (2008) | 179 | | 179 |
| Geweke (1989b) | 144 | Geweke (1996) | 126 | Manchanda et al. (1999) | 281 | Berg et al. (2004) | 147 | Lee and Card (2008) | 355 | | 355 |
| Karp and Perloff (1989) | 111 | Chib (1996) | 368 | Rust et al. (1999) | 309 | Kim et al. (2004) | 181 | Chauvet and Piger (2008) | 110 | | 110 |
| Koray and Lastrapes (1989) | 265 | Smith and Kohn (1996) | 494 | Martin (2000) | 133 | Gilbride and Allenby (2004) | 238 | Manchanda et al. (2008) | 115 | | 115 |
| Lenk and Rao (1990) | 109 | Dejong et al. (1996) | 106 | Schorfheide (2000) | 481 | Montgomery et al. (2004) | 240 | Netzer et al. (2008) | 116 | | 116 |
| Sims and Uhlig (1991) | 317 | Lenk et al. (1996) | 352 | Nakatsuma (2000) | 115 | Canova (2004) | 274 | Imai et al. (2009) | 110 | | 110 |
| Schotman and Van Dijk (1991b) | 158 | Rossi et al. (1996) | 475 | Geweke and Keane (2000) | 126 | Negro and Schorfheide (2004) | 373 | Ley and Steel (2009) | 154 | | 154 |
| Phillips (1991) | 280 | Kadiyala and Karlsson (1997) | 300 | Dejong et al. (2000) | 189 | Fernández-Villaverde and Rubio-Ramirez (2005) | 137 | Mountford and Uhlig (2009) | 845 | | 845 |
| Zellner et al. (1991) | 131 | Fernández et al. (1997) | 121 | McCulloch et al. (2000) | 212 | Peersman (2005) | 189 | Faust and Wright (2009) | 101 | | 101 |
| Schotman and Van Dijk (1991a) | 158 | Koop et al. (1997) | 196 | Lancaster (2000) | 392 | Canova (2005) | 213 | Canova and Ciccarelli (2009) | 119 | | 119 |
| Hamilton (1991) | 159 | Maddala et al. (1997) | 189 | Aguilar and West (2000) | 333 | Smets and Wouters (2005) | 292 | Justiniano and Preston (2010) | 100 | | 100 |
| Chib (1992) | 246 | Faust and Leeper (1997) | 460 | Wedel and Pieters (2000) | 185 | Dehejia (2005) | 119 | Nbura et al. (2010) | 190 | | 190 |
| Shea (1992) | 110 | Montgomery (1997) | 208 | Hamilton (2001) | 190 | O'Donnell and Coelli (2005) | 122 | Leeper et al. (2010) | 109 | | 109 |
| Andrews (1993) | 395 | Kleibergen and Van Dijk (1998) | 118 | Elerian et al. (2001) | 443 | Imai and van Dyk (2005) | 133 | Rubio-Ramirez et al. (2010) | 114 | | 114 |
| Geweke (1993) | 378 | Poirier (1998) | 137 | Fernandez et al. (2001b) | 480 | Yu (2005) | 245 | Schmitt-Grohe and Uribe (2012) | 260 | | 260 |
| Min and Zellner (1993) | 232 | Chib et al. (1998) | 102 | Fernandez et al. (2001a) | 501 | Owyang et al. (2005) | 130 | Köse et al. (2012) | 277 | | 277 |
| Chib (1993) | 447 | Chiang et al. (1998) | 134 | Chib and Winkelmann (2001) | 143 | Borje et al. (2005) | 107 | | | | |

Note: The table gives the number of citations for papers which are published in the ten journals we consider and are cited at least 100 times. The citation numbers are collected using Google Scholar on 13–14 April, 2014.

Figure A.1: Average number of citations for papers in leading journals for 5-year intervals



Note: The figures show the number of Bayesian papers in the journals classified according to the number of citations. Data definition is as in Table A.2.