DEUTSCHE
BUNDESBANK
EUROSYSTEM

Empirical Bayesian density forecasting in Iowa and shrinkage for the Monte Carlo era

Kurt F. Lewis
(University of Iowa)

Charles H. Whiteman
(University of Iowa)

Discussion Paper
Series 1: Economic Studies
No 28/2006

Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main,
Postfach 10 06 02, 60006 Frankfurt am Main

Tel +49 69 9566-1
Telex within Germany 41227, telex from abroad 414431, fax +49 69 5601071

Please address all orders in writing to: Deutsche Bundesbank,
Press and Public Relations Division, at the above address or via fax +49 69 9566-3077

ISBN 3–86558–191–9 (Printversion)
ISBN 3–86558–192–7 (Internetversion)

This paper was presented at the 8th Bundesbank spring conference (May 2006) on “New Developments in Economic Forecasting”. The views expressed in the paper are those of the authors and not necessarily those of the Bundesbank.

New Developments in Economic Forecasting
8th Bundesbank Spring Conference

(Preliminary programme as at 10 March 2006)

Organisers: Michael Binder, Jörg Breitung, Heinz Herrmann

Eltville, 05/06 May 2006

Friday, 5 May 2006

9:00 - 9:30	**Introductory remarks by** Axel A Weber *(Deutsche Bundesbank)*	
9:30 - 10:45	**Multiple structural breaks, forecasting and present value calculations** Hashem Pesaran *(University of Cambridge)*	
	Discussant:	Uwe Hassler *(Johann Wolfgang Goethe-University)* Siem Jan Koopman *(Free University Amsterdam)*
10:45 – 11:00	*Coffee break*	
11:00 - 12:15	**Testing predictive ability using estimated co integration relationships: Business as usual?** Lutz Kilian *(University of Michigan)*	
	Discussant:	Hans-Eggert Reimers *(Hochschule Wismar)* Andreas Beyer *(European Central Bank)*

12:15 – 14:00	*Lunch*		
14:00 – 15:00	**Global yield curve dynamics and interactions** Francis X Diebold *(University of Pennsylvania)*		
	Discussant:	Stefan Mittnik *(Ludwig-Maximilians-University Munich)* Joachim Grammig *(Eberhard-Karls-University Tübingen)*	
15:00 - 15:15	*Coffee break*		
15:15 - 16:30	**A benchmark for models of growth and inflation** Massimiliano Marcellino *(Università Bocconi)*		
	Discussant:	George Kapetanios *(University of London)* Todd Clark *(Federal Reserve Bank of Kansas City)*	
16:30 - 17:45	**Forecasting with panel data** Badi Baltagi *(Texas A&M University)*		
	Discussant:	Helmut Herwartz *(Christian-Albrechts-University Kiel)* Jean-Pierre Urbain *(University Maastricht)*	
20:00	*Dinner* hosted by Axel A Weber *(Deutsche Bundesbank)*		
	Speaker:	Zdenek Tuma *(Czech National Bank)*	

Saturday, 6 May 2006

9:30 - 10:45	**Did the ECB make a difference for Euro area business cycle?** Fabio Canova* *(Universitat Pompeu Fabra)* Matteo Ciccarelli *(European Central Bank)* Eva Ortega *(Bank of Spain)*	
	Discussant:	Sandra Eickmeier *(Deutsche Bundesbank)* Domenico Giannone *(Université Libre de Bruxelles)*
10:45 - 11:00	*Coffee break*	
11:00 - 12:15	**Bayesian density forecasting with best performing priors determined by entropic tilting** Charles Whiteman* (University of Iowa) Kurt Lewis (University of Iowa)	
	Discussant:	John Geweke *(University of Iowa)* Stephane Adjemian *(CEPREMAP)*
12:15 - 13:30	*Lunch*	
13:30 - 14:45	**Shrinkage methods for forecasting using many predictors** Mark Watson* *(Princeton University)* James Stock *(Harvard University)*	
	Discussant:	Gary Koop *(University of Leicester)* Carlo Favero *(Universitat Pompeu Fabra)*

14:45 - 16:00	**Forecasting using a large number of predictors** Lucrezia Reichlin *(European Central Bank)*
	Discussant: Christian Schumacher *(Deutsche Bundesbank)* Peter Vlaar *(Dutch Central Bank)*
16:00	**Concluding remarks by** Hermann Remsperger *(Deutsche Bundesbank)*

Deutsche Bundesbank
Protocol and Conferences, K 100
Tel: +49 69 9566-4066/-2864, Fax: +49 69 9566 4027
Mobile: +49 177 7382251, E-mail: conferences@bundesbank.de

Non-technical summary

The track record of a sixteen-year history of density forecasts of state tax revenue in Iowa is studied, and potential improvements sought through a search for better performing "priors" similar to that conducted two decades ago for point forecasts by Doan, Litterman, and Sims (*Econometric Reviews*, 1984). Comparisons of the point- and density-forecasts produced under the flat prior are made to those produced by the traditional (mixed estimation) "Bayesian VAR" methods of Doan, Litterman, and Sims, as well as to fully Bayesian, "Minnesota Prior" forecasts. The actual record, and to a somewhat lesser extent, the record of the alternative procedures studied in pseudo-real-time forecasting experiments, share a characteristic: subsequently realized revenues are in the lower tails of the predicted distributions "too often". An alternative empirically-based prior is found by working directly on the probability distribution for the VAR parameters, seeking a betterperforming entropically tilted prior that minimizes in-sample mean-squared-error subject to a Kullback-Leibler divergence constraint that the new prior not differ "too much" from the original. We also study the closely related topic of robust prediction appropriate for situations of ambiguity. Robust "priors" are competitive in out-of-sample forecasting; despite the freedom afforded the entropically tilted prior, it does not perform better than the simple alternatives.

Nicht-technische Zusammenfassung

Es wird die Treffsicherheit der seit sechzehn Jahren durchgeführten Dichteprognosen zu den Steuereinnahmen des US-Bundesstaates Iowa analysiert und nach möglichen Verbesserungen gesucht; dies geschieht durch die Suche nach besseren „Priors“ ähnlich wie vor zwei Jahrzehnten bei den Punktprognosen von Doan, Litterman und Sims (*Econometric Reviews*, 1984). Die Punkt- und Dichteprognosen auf der Grundlage des flachen Priors werden mit jenen der traditionellen Bayes'schen VAR-Methoden nach Doan, Litterman und Sims sowie mit den reinen Bayesianischen „Minnesota Prior“-Prognosen, verglichen. Das tatsächliche Ergebnis und – in etwas geringerem Umfang – auch jenes der alternativen Verfahren, die anhand von Experimenten in Pseudoechtzeit untersucht werden, haben eines gemeinsam: Die tatsächlich erzielten Einnahmen liegen „zu oft“ am unteren Rand der vorausgesagten Verteilung. Ein alternativer, empirisch-basierter Prior lässt sich ermitteln, indem die Wahrscheinlichkeitsverteilung für die VAR-Parameter direkt verwendet wird und ein besserer Prior gesucht wird, der den quadrierten mittleren „in-sample“-Fehler minimiert, und zwar unter der Bedingung, dass der neue Prior nicht „zu stark“ vom Original abweicht. Wir untersuchen auch das eng verwandte Thema einer stabilen Vorhersage, die für nicht eindeutige Situationen zweckmäßig ist. Stabile „Priors“ sind in „out-of-sample“-Prognosen kompetitiv; trotz des Freiraums dieses echteren Priors schneidet er nicht besser ab als die einfachen Alternativen.

EMPIRICAL BAYESIAN DENSITY FORECASTING IN IOWA AND SHRINKAGE FOR THE MONTE CARLO ERA

KURT F. LEWIS AND CHARLES H. WHITEMAN†

Abstract. The track record of a sixteen-year history of density forecasts of state tax revenue in Iowa is studied, and potential improvements sought through a search for better performing "priors" similar to that conducted two decades ago for point forecasts by Doan, Litterman, and Sims (*Econometric Reviews*, 1984). Comparisons of the point- and density-forecasts produced under the flat prior are made to those produced by the traditional (mixed estimation) "Bayesian VAR" methods of Doan, Litterman, and Sims, as well as to fully Bayesian, "Minnesota Prior" forecasts. The actual record, and to a somewhat lesser extent, the record of the alternative procedures studied in pseudo-real-time forecasting experiments, share a characteristic: subsequently realized revenues are in the lower tails of the predicted distributions "too often". An alternative empirically-based prior is found by working directly on the probability distribution for the VAR parameters, seeking a better-performing entropically tilted prior that minimizes in-sample mean-squared-error subject to a Kullback-Leibler divergence constraint that the new prior not differ "too much" from the original. We also study the closely related topic of robust prediction appropriate for situations of ambiguity. Robust "priors" are competitive in out-of-sample forecasting; despite the freedom afforded the entropically tilted prior, it does not perform better than the simple alternatives.

All models are wrong but some are useful.

G.E.P. Box

1. Introduction

The Institute for Economic Research at The University of Iowa has, since 1990, produced Bayesian density forecasts for state tax revenue growth using vector autoregressions (VARs) under noninformative prior distributions. In the beginning, these were met by state officials

Date: April, 2006. PRELIMINARY AND INCOMPLETE.

JEL codes: C11, C15, C32.

† We thank Chris Sims, who, by objecting to some aspects of it, sharpened our thinking about the approach of this paper.

Prepared for the 2006 Bundesbank Conference "New Developments in Economic Forecasting".

with incredulity ("the range of reasonable predictions seems quite wide") or disdain (an official long since departed "didn't like those inverted umbrellas" surrounding the median predicted time path). But after an early success in prediction (that would have saved the state–had it only listened–from multiple arbitration proceedings with its largest labor unions and from subsequent very painful layoffs), the relevant decision makers did indeed realize that they in fact wanted the best possible advice regarding the future, and a reasonable measure of the uncertainties they face.

This paper provides a description of the density forecast effort, and an assessment of the record. The upshot of this is that there have been more occasions (though perhaps not statistically significantly so) than one would prefer in which subsequently realized revenue growth occurred in the lower left tails of the predictive distributions. This suggests that the model used and the predictive densities produced imparted a greater sense of security to the policy makers than was warrented. Thus we conduct a search for better predictive densities by considering alternative models (within the class of vector autoregressions–VARs), and "better performing" prior distributions in the spirit of the search conducted by Doan, Litterman, and Sims (1984) for point forecasts two decades ago.

One alternative to our procedure would involve leaving the class of fixed-specification VAR models. While this is a subject for future research and development, we were motivated by Box's statement to try to determine how to proceed in forecasting given a universe of models (albeit a small one) that is unlikely to contain the "one true model". Thus we consider two procedures that can be interpreted as pragmatic alternatives to Bayesian updating to be used when the model generating the forecasts is potentially misspecified, there are no obviously correct alternative specifications, and the objective is to produce accurate density forecasts at multi-step horizons.

Both procedures work directly on the subjective probabilities of what is taken to be unknown in our exercise–the parameters of the models we consider. The first of these, which we call "entropic tilting", can be interpreted as a nonparametric (or "super parametric", depending on perspective) alternative to the search over hyperparameters of a hierarchically specified prior (as in Doan, Litterman, and Sims (1984)), in which the objective is better out-of-sample forecast performance. Given an initial posterior distribution for the unknown parameters, the entropic tilt finds a new distribution that achieves a forecasting objective not employed in the construction of the original distribution, subject to a penalty for Kullback-Leibler divergence between the two.

The second, closely related procedure, involves a reasonable (and recently axiomatically justified) response to the possibility that the model might be misspecified. The resulting "robust" prediction involves a new density which is an entropically tilted version of the original.

The results of the paper indicate that much of the forecasting benefit of the Monte Carlo era is in the production of density forecasts rather than point forecasts. Even a flat-prior posterior predictive density does well in comparison to various feasible shrinkage schemes. That is, density forecasts used to produce point predictions under reasonable loss seem to improve over direct point forecasts using fixed parameterizations, and this improvement is about as great as what can be achieved employing best shinkage practices. Moreover, the results of the entropic tilting exercises are largely negative. Robust predictions provide a little more realism, but the effort to design a better performing prior by pursuing this goal directly did not improve out-of-sample forecasts appreciably.

2. The Iowa Experience

The Institute for Economic Research at The University of Iowa began producing density forecasts of state tax revenues in 1990. This section describes the nature of that effort, provides an assessment of the forecast record, and makes note of some lessons learned from that assessment. Other aspects of the Institute's forecast effort have been examined in Whiteman (1996) and Otrok and Whiteman (1998).

A typical budget cycle in the State of Iowa begins in December, the sixth month of the fiscal year, with the Official Revenue Estimates. The estimates are of revenues to be received during the current and ensuing fiscal years, and must be made, in accord with statute, by December 15. Subsequently, the Governor makes spending proposals in early January, in conjunction with the "Condition of the State" speech that opens the legislative session. The legislature meets during the spring; bills passed by both houses of the legislature are sent to the Governor for signature or veto. Most spending bills sent to the Governor during the session are for expenditures to take place during the next fiscal year, though sometimes "supplemental" spending bills for the current fiscal year are passed. Both the Governor's spending recommendations and all spending bills must hold spending to 99% of revenues estimated to be available during the fiscal year of the expenditure.

The Official Revenue Estimates are set by the Revenue Estimating Conference (REC), comprising three individuals appointed by the Governor. Two of the appointees are from state government. Typically, the Governor has appointed the Directors of the Legislative Service Bureau (the state agency analogous to the U.S. Congressional Budget Office) and the Department of Management (the state analogue to the U.S. Office of Management and Budget). The third appointment goes to an individual not associated with state government, but the third appointee must be agreed to by the other two.

Resources available to the REC include analyses made by staff members in the Legislative Service Bureau, the Department of Management, and the Department of Revenue (which is responsible for collecting the revenue). In addition, the Conference has available the report of the Governor's Council of Economic Advisors (CEA), which meets prior to the REC. The Council comprises a dozen or so business and academic economists, including the Director of the University of Iowa Institute for Economic Research. The Institute's revenue forecasts (and forecasts of economic conditions) form an important part of the basis of discussions during quarterly CEA meetings, and are made available ahead of time to the CEA, the REC, staff members of the relevant state agencies, and the Legislature.

2.1. **The Model.** When the Institute's modern forecast effort began in February of 1990, revenue data were available monthly from July 1982. Earlier data were collected under quite different accounting conventions, meaning that additional data comparable to the more recent observations were not available.[1] The paucity of data dictated consideration of small models (or incorporation of substantial prior information, which was not readily available).

Given the success of forecasts made using Vector Autoregression (VARs) at the Federal Reserve Bank of Minneapolis (Litterman (1986)), effort was undertaken to develop a VAR forecasting model for Iowa revenues.[2] The lack of data suggested a small, low-order VAR,

[1] More precisely, a Research Assistant was unable to find data that could be sensibly spliced to the 1982 data despite a month spent in the state capital attempting to do so.

[2] Prior to 1990, and under the administration of a different director, the Institute employed a commercial consulting firm to design and periodically reestimate a standard (for the time) simultaneous equations forecasting model for economic conditions, but not tax revenues. When informed of the decision to move the Institute's focus to locally maintained Vector Autoregression models, the firm dispatched a senior sales

and experimentation with 2- and 3-variable VARs with 1-4 lags and time series representing components of employment (e.g., employment in services, employment in retail trade, employment in durable goods manufacturing, etc.) and income (e.g., wage and salary disbursements, property income, etc.) led to a two-lag (later changed to four) quarterly specification involving total revenues and total nonfarm personal income. Deterministic variables included a constant and seasonal dummies. The experimentation involved pseudo-real time forecasting experiments like those described in Litterman (1979) and Doan, Litterman, and Sims (1984) and in Section 5 below. Briefly, this involved specification of a holdout sample beginning in period $T + 1$ (prior to the end of the sample), estimation of the parameters of the equations by ordinary least squares, equation by equation (which coincides in this simple case with maximum likelihood) using data from the beginning of the sample until period T, forecasting through $T+h$ using the estimated parameters and the chain rule, recording of the resulting "out-of-sample" forecast accuracy, and finally incrementing T and repeating the process until all the data had been exhausted. The best performing model in this exercise was the one chosen. This model choice was revisited in 2004, when over 500 two- and three-variable models involving permutations of variables to be included alongside tax revenue were pitted against one-another in pseudo-real time forecasting over the period 1999-2004. No model was found to be measurably superior to a 4-lag, 2-variable VAR involving tax revenue and personal income.

This procedure for choosing the model to be used of course violates the likelihood principle (Berger and Wolpert (1988)). By the time the exercise is completed, all the data have been used, but the evidence regarding the unknown parameters is summarized not exclusively by the likelihood function (i.e., in one-step-ahead prediction), but rather by performance in multi-step forecasting. The procedure is essentially the same one used by Doan, Litterman, and Sims (1984), and is a pragmatic approach to model choice when there is suspicion that the model is not correct. It is quite unlikely that the process generating income, revenue, and employment data is literally a VAR or anything remotely like it. Indeed, the motivation for using VARs, stemming from Sims's original work (Sims (1980) is that they merely *represent* stationary time series in the sense of reproducing the first and second moments of the time series. Moreover, following Box's dictum, it is unlikely that the universe of models that could

executive, whose last ditch appeal to keep the contract involved the statement "we don't care about accuracy, we care about consistency." The appeal failed.

be considered without massive research efforts every three months would in fact include the true model. Yet the false model (here, the VAR) is useful for producing forecasts, and it seems natural to base aspects of its specification on how useful it is in producing them.

The specification employed was

$$\mathbf{y}_t = \mathbf{B}_D D_t + \mathbf{B}_1 \mathbf{y}_{t-1} + \mathbf{B}_2 \mathbf{y}_{t-2} + ... + \mathbf{B}_m \mathbf{y}_{t-m} + \varepsilon_t \tag{1}$$

where $\mathbf{y}_t$ is the 2×1 vector including the (log of) revenues and personal income, D_t is the deterministic component, and $\varepsilon_t \overset{iid}{\sim} N(0, \mathbf{\Psi})$. It is convenient to group the parameters together in the vector θ:

$$\theta = (\mathbf{B}_D, \mathbf{B}_1, ..., \mathbf{B}_m, \mathbf{\Psi}).$$

The model selected in this fashion was to be used in a fully Bayesian fashion to produce predictive densities for revenue growth for the current and one ensuing fiscal year. To understand this procedure, it is helpful to fix some notation (which follows Geweke and Whiteman (2006) closely). First, denote by $\mathbf{Y}_t$ the history $\{\mathbf{y}_s\}_{s=1}^{t}$. Then the probability density function for $\mathbf{y}_t$ is given by

$$p(\mathbf{y}_t \mid \mathbf{Y}_{t-1}, \theta) \tag{2}$$

Thus (2) compactly encodes the structure and distributional assumptions of (1). It should be understood as a function involving the random vector $\mathbf{y}_t$. The density for the entire history $\mathbf{Y}_T = \{\mathbf{y}_s\}_{s=1}^{T}$ is given by

$$p(\mathbf{Y}_T \mid \theta) = \prod_{t=1}^{T} p(\mathbf{y}_t \mid \mathbf{Y}_{t-1}, \theta). \tag{3}$$

Using the superscript o to denote the observed value of a random vector, the likelihood function associated with the observed sample is

$$L(\theta; \mathbf{Y}_T^o) \propto p(\mathbf{Y}_T^o \mid \theta). \tag{4}$$

The prior distribution, representing pre-sample subjective beliefs regarding the unobservable parameter vector θ is $p(\theta)$, and the posterior distribution for θ after observing the sample $\mathbf{Y}_T$ is given by

$$p(\theta \mid \mathbf{Y}_T^o) \propto p(\theta)\, p(\mathbf{Y}_T^o \mid \theta) \tag{5}$$

The predictive density for the next h values of $\mathbf{y}$ is obtained by integrating the product of the conditional predictive given θ and the posterior:

$$p\left(\mathbf{y}_{T+1},\ldots,\mathbf{y}_{T+h} \mid \mathbf{Y}_T^o\right) = \int_\Theta p\left(\theta \mid \mathbf{Y}_T^o\right) p\left(\mathbf{y}_{T+1},\ldots,\mathbf{y}_{T+h} \mid \mathbf{Y}_T^o, \theta\right) d\theta. \tag{6}$$

We shall seek to generate a Monte Carlo sample from this predictive density. This is accomplished by first sampling from the posterior, $\theta^{(i)} \stackrel{iid}{\sim} p\left(\theta \mid \mathbf{Y}_T^o\right)$, and then for each i, performing a dynamic simulation from the VAR by drawing h $k-$variate noise values $\epsilon_{T+1}, \epsilon_{T+2}, \ldots, \epsilon_{T+h}$ from the distribution $N(0, \mathbf{\Psi})$ and computing $\mathbf{y}_{T+1}, \mathbf{y}_{T+2}, \ldots, \mathbf{y}_{T+h}$ sequentially using (1).

To operationalize this, note that the k right-hand-side variables of each of the equations in (1) are the same, and collect the T observations on these variables in a $T \times k$ matrix $\mathbf{X}_T$. Then the VAR can be written in the form

$$vec \mathbf{y}_T = (\mathbf{I} \otimes \mathbf{X}_T)\mathbf{b} + \mathbf{U}_T \tag{7}$$

where now $\mathbf{b}$ contains the elements of the $\mathbf{B}_i$'s in (1), and $\mathbf{U}_T$ is the $T \times 2$ matrix of values of the ϵ's in (1), $\mathbf{U}_T \sim N(0, \mathbf{\Psi} \otimes \mathbf{I}_T)$. Following the development in Zellner (1971), under the "noninformative" prior

$$p(\mathbf{\Psi}) \propto |\mathbf{\Psi}|^{-(p+1)/2}$$

where p is the number of equations, the posterior distribution is given by

$$\mathbf{b}|\mathbf{\Psi}, \mathbf{y}_T, \sim N(\bar{\mathbf{b}}, \mathbf{\Psi} \otimes (\mathbf{X}_T^{'}\mathbf{X}_T)^{-1}) \tag{8}$$

$$\mathbf{\Psi}|\mathbf{y}_T \sim IW(\mathbf{S}, T-k)$$

where $\bar{\mathbf{b}}$ denotes the OLS estimates of the slope parameters, and IW denotes the inverted Wishart distribution with "shape" parameter $\mathbf{S}$ (given by the OLS residual variance-covariance matrix) and degrees of freedom $T-k$.

This form for the posterior distribution means that it is straightforward to generate a Monte Carlo sample $\theta^{(i)} \stackrel{iid}{\sim} p\left(\theta \mid \mathbf{Y}_T^o\right)$. The random sample from the predictive distribution is used to create a density forecast like the one in Figure 1, taken from the Institute's December 2005 forecast. During the first few quarters the Institute produced such forecasts, state officials seemed to interpret the densities as granting them latitude in making predictions. However, after making several forecast errors larger than what would have occurred at the

Institute's mean by embracing that latitude, they began to see the spread in the predictive density as an indication of the degree of uncertainty characterizing the decision making environment. This recognition by state officials–that many things can happen, some more likely than others–was perhaps the most important consequence of the decision to produce density forecasts.

The predictive density summarizes possible outcomes and associated probabilities for the vector $\mathbf{y}$ at some future date, say $T+h$. We denote this random vector by $\tilde{\mathbf{y}}_{T+h}$. If a point forecast is needed (as is required of the REC in Iowa), it is necessary to specify a loss function describing the costs associated with making a point prediction, say $\mathbf{y}^F_{T+h}$, of the random vector $\tilde{\mathbf{y}}_{T+h}$: $L(\mathbf{y}^F_{T+h}, \tilde{\mathbf{y}}_{T+h})$. Then the loss minimizing forecast $\mathbf{y}^F_{T+h}$ minimizes

$$E\left[L\left(\mathbf{y}^F_{T+h}, \tilde{\mathbf{y}}_{T+h}\right) \mid \mathbf{Y}^o_T\right] = \int L\left(\mathbf{y}^F_{T+h}, \tilde{\mathbf{y}}_{T+h}\right) p\left(\tilde{\mathbf{y}}_{T+h} \mid \mathbf{Y}^o_T\right) d\tilde{\mathbf{y}}_{T+h}. \tag{9}$$

Since the early 1990's, the Institute has reported loss-minimizing point predictions under the loss functions

$$L\left(y^F_{T+h}, \tilde{y}_{T+h}\right) = d\cdot\left(y^F_{T+h} - \tilde{y}_{T+h}\right) I_{\left(-\infty, y^F_{T+h}\right)}\left(\tilde{y}_{T+h}\right) + \left(\tilde{y}_{T+h} - y^F_{T+h}\right) I_{\left(y^F_{T+h}, \infty\right)}\left(\tilde{y}_{T+h}\right), \tag{10}$$

where y_{T+h} denotes the (scalar) annual growth rate of revenue computed from the relevant element of $\mathbf{y}_{T+h}$, and $d = 1, 2, \ldots, 10$. This loss function represents the view that revenue shortfalls are d times as costly as equal-sized surpluses. The solution to the minimization problem for given d is to set y_{T+h} equal to the $1/(1+d)$'th quantile of the predictive distribution of $\tilde{y}_{T+h}$.

An example of this sort of prediction for annual revenue growth is given in Table 1. Thus the 3-1 loss-ratio forecast for fiscal year 2007 is 2% growth. Although there is no necessary connection between the Institute's density forecasts and the official REC forecast, and the REC does not indicate anything about how it arrives at its estimate, the estimates have been coincidentally similar to the 3-1 loss ratio forecast. The next subsection assesses the forecast record.

Revenue Growth (%) Forecasts			
Loss Factor	FY06	FY07	FY08
1	3.5	3.0	4.2
2	2.9	2.3	3.4
3	2.6	2.0	3.0
4	2.4	1.6	2.7
5	2.3	1.3	2.5
6	2.1	1.1	2.3
7	2.1	1.0	2.2
8	2.0	1.0	2.1
9	1.9	0.9	2.0
10	1.8	0.8	1.8

Table 1. Loss Ratio Forecasts for Iowa Revenue Growth Made December 2005

2.2. **The Record.** The density forecast record is summarized in Table 2 and Figures 2 and 3. Table 3 provides the complete record of annual revenue growth realizations, together with the predictive means and standard deviations, and subsequently realized quantiles of the one- and two-year forecasts made at the end of the calendar year just prior to the December meeting of the REC. The root mean squared error over the entire period for the (most) important one-step forecast is 2.6%. This is perhaps comforting, given the remark by former state comptroller Marvin Seldon (quoted in Geweke and Whiteman (2006)) at the inception of this effort: "if you can find a revenue forecaster who can get you within 3 percent, keep him".

But the focus on the root mean squared error or the mean absolute error of the forecast obscures how well the density forecasts performed. The column labeled "Percentile" indicates the quantile of the predictive distribution at which the realized value occurred. Under the correct specification, these percentiles should be distributed uniformly (Section 3 of Diebold, Gunther, and Tay (1998)). The omnibus Pearson χ^2 tests for uniformity using quartiles have p-values 0.27 and 0.29 for the one- and two-step forecasts, but this test has little power with so few observations–15 and 14, respectively. A more powerful procedure involves comparison of the empirical and population cumulative distributions. These are displayed in Figure 2. The Kolmogorov-Smirnov statistic (the maximum absolute deviation between the empirical and population CDF's) for testing uniformity is 0.34 for the one-step forecast, which is at

	Fiscal Year 1				Fiscal Year 2			
Forecast	Actual	Mean	Std Dev	Percentile	Actual	Mean	Std Dev	Percentile
Dec-90	4.7	4.3	1.7	0.59	5.9	4	4.3	0.67
Dec-91	5.9	6.7	1.5	0.30	10.3	4.7	3.7	0.93
Dec-92	10.3	11.5	1.8	0.25	6.9	5.1	4.2	0.67
Dec-93	6.9	8.6	2.3	0.23	5.7	5	4.6	0.56
Dec-94	5.7	7.9	1.4	0.06	7.1	4.8	2.2	0.85
Dec-95	7.1	7.3	1.3	0.44	3.4	5.5	2	0.15
Nov-96	3.4	6.4	1.3	0.01	5	6.3	2	0.26
Nov-97	5	4.2	1.2	0.75	1	3.1	2.1	0.16
Nov-98	1	3.3	1.3	0.04	4.5	6.3	2.3	0.22
Nov-99	4.5	3.3	1.6	0.77	0.4	4.7	3.2	0.09
Nov-00	0.4	5.9	1.3	0.00	-2.1	6	2.3	0.00
Nov-01	-2.1	3.7	1.4	0.00	0.8	3.6	3.1	0.18
Nov-02	0.8	3.2	1.2	0.02	4.2	2	2	0.86
Dec-03	4.2	3.3	1.3	0.76	5.25	0.8	2	0.99
Dec-04	5.25	2.8	1.3	0.97	NA	2.4	1.9	NA

	MSE	MAE		MSE	MAE
90-04	2.6	2.1	90-04	3.5	2.9
90-97	1.6	1.3	90-97	2.6	2.2
98-04	3.4	2.9	98-04	4.2	3.7

Table 2. Real-Time Performance Statistics, Iowa Revenue Forecasts

the 5% point of the sampling distribution for a sample size of 15, but the p-value for the two-step statistic 0.24 exceeds 20%.

Of course not being able to reject uniformity may be small consolation in practice. Indeed, there are several realized values in the $0-5\%$ left tails of the one-step predictive distributions, and such lower-tail values can be especially problematic for state budgeting. Moreover, there is evidence of serial correlation in the realizations depicted in Figure 3: Ljung-Box Q-Statistics have p-values of 0.04 (fiscal year 1) and 0.09 (fiscal year 2) based on three autocorrelations. The next section begins a study of whether it might be possible to improve on the density forecast record.

3. Conventional Alternative Procedures for Producing Density Forecasts

In this section we outline several alternative VAR structures whose performance is studied in Section 5. As benchmarks, we first consider flat-prior and "Minnesota Prior" VARs treated in customary historical fashion–estimated using equation-by-equation ordinary least squares and mixed estimation procedures, and used to produce forecasts at the point estimates of the parameters. The accuracy of these forecasts is assessed conventionally. We also consider fully Bayesian flat-prior and Minnesota prior VARs analyzed by posterior importance samplers. When analyzed in fully Bayesian fashion, it is straightforward to produce density forecasts from these models, and we study the perfomance of such forecasts in pseudo-real time experiments.

The classes of models are:

(1) VAR (equation (1)) estimated by ordinary least squares,
(2) "Shrinkage" (mixed) equation-by-equation estimation of (1) with a Minnesota prior,
(3) VAR (equation (1)) analyzed by fully Bayesian Monte Carlo methods (the model of Section 2),
(4) VAR (equation (1)) analyzed by fully Bayesian Monte Carlo methods with a Minnesota prior.

By the early 1980's, conventional wisdom based on experience (e.g., the evidence in Fair (1979)) and James and Stein (1961)-like shrinkage arguments had it that OLS estimation of VARs, while straightforward, led to poor out-of-sample forecasting performance, and that "shrinkage" of OLS estimates toward virtually *anything* else could improve performance. This spurred Litterman (1979) and later Doan, Litterman, and Sims (1984) to develop a class of shrinkage targets intended to improve forecasting performance. Collectively, this class is known as the Minnesota prior.

The Minnesota prior embodies the notion that the collection of time series can be approximated as a set of univariate random walks. In terms of the parameters of (1), the simplest version of the prior is silent on the deterministic components in $\mathbf{B}_d$, takes the mean of the first lag of variable i in equation i (the "own first lag coefficient") to be unity, and sets the mean value of the other slope parameters to zero. Prior certainty (that the parameters are zero) increases with the lag. Specifically, if δ_{ij}^l is the prior standard deviation of the coefficient on lag l of variable j in equation i,

$$\delta_{ij}^l = \begin{cases} \frac{\lambda}{l^{\gamma_1}} & \text{if } i = j \\ \\ \frac{\lambda\gamma_2\hat{\sigma}_i}{l^{\gamma_1}\hat{\sigma}_j} & \text{if } i \neq j \end{cases} \tag{11}$$

where γ_1 (the "decay rate") is nonnegative, γ_2 ("others' weight") and λ ("overall tightness") are scale factors, and $\hat{\sigma}_i$ and $\hat{\sigma}_j$ are the estimated residual standard deviations in unrestricted ordinary least squares estimates of equations i and j of the system. (This is the formulation in Litterman (1979); subsequently, Litterman (1986) took these scale factors from univariate autoregressions.) As others' weight is reduced, the specification shrinks toward a univariate autoregression. For small γ_2 (others' weight), as the decay rate increases, the specification shrinks toward a random walk. Finally, the prior exerts a greater influence on the results the smaller is λ (overall tightness). We will refer to the collection of hyperparameters as Λ.

The Minnesota prior can be understood as a set of stochastic (the RATS manual (Doan (2004)) uses the term "fuzzy") linear restrictions on the coefficients $\mathbf{B}_j$ of equation (1) of the form

$$\mathbf{r}_i = \mathbf{R}_i\beta_i + \mathbf{v}_i; \qquad \mathbf{v}_i \sim N(0, \lambda^2\mathbf{I}) \tag{12}$$

where β_i represents the lag coefficients in equation i (the i^{th} row of $\mathbf{B}_1, \mathbf{B}_{2,}..., \mathbf{B}_m$ in equation (1)), R_i is a diagonal matrix with elements λ/δ^l_{ij} corresponding to the l^{th} lag of variable j, and r_i is a vector of zeros except for a one corresponding to the first lag of variable i.

As ordinarily implemented, the prior is applied equation by equation. To see how this is carried out, let $\mathbf{Y}_{it}$ denote the $T \times 1$ vector of observations on the i'th element of $\mathbf{y}_t$, and write equation i as

$$\mathbf{Y}_{iT} = \mathbf{X}_T\beta_i + \mathbf{u}_{iT}. \tag{13}$$

where $\mathbf{u}_{iT} \sim N(0, \sigma_i^2 I)$. A fully Bayesian, single-equation analysis would work with the product of the Gaussian likelihood associated with (13) and the Gaussian prior distribution in (12), which is in turn Gaussian with mean determined by completion of the square. Alternatively, the more customary "shrinkage" estimator analysis works with a regression interpretation of (12) and (13) by stacking the left-and right-hand side variables as

$$\begin{bmatrix} \mathbf{Y}_{iT} \\ (\sigma_i/\lambda)\mathbf{r}_i \end{bmatrix} \qquad \begin{bmatrix} \mathbf{X}_T \\ (\sigma_i/\lambda)\mathbf{R}_i \end{bmatrix}$$

and applying OLS to the stacked system, yielding a mixed estimator (Theil (1963))

$$\hat{\beta}_i = (\mathbf{X}_T'\mathbf{X}_T + (\frac{\sigma}{\lambda})^2\mathbf{R}_i'\mathbf{R}_i)^{-1}(\mathbf{X}_T'\mathbf{Y}_{iT} + (\frac{\sigma}{\lambda})^2\mathbf{R}_i'\mathbf{r}_i). \tag{14}$$

Equation-by-equation estimators like these are then collected together and used in (1) to forecast future values of $\mathbf{y}_{T+1}, \mathbf{y}_{T+2}, \ldots, \mathbf{y}_{T+h}$. Doan, Litterman, and Sims (1984) conditioned on values of the hyperparameters (e.g., overall tightness, others' weight, and the decay rate) and studied out-of-sample forecast performance as these quantities were varied systematically.

Fully Bayesian analysis under the Minnesota prior is typically not carried out (though see Kadiyala and Karlsson (1993) for an exception) because the prior cannot be written in conjugate form. Indeed, the conjugate prior for the model in (1) is the "Normal-Wishart" prior (see Kadiyala and Karlsson (1993))

$$\mathbf{b}|\mathbf{\Psi} \sim N(\underline{\mathbf{b}}, \mathbf{\Psi} \otimes \underline{\mathbf{\Omega}}) \tag{15}$$

$$\mathbf{\Psi} \sim IW(\underline{\mathbf{\Psi}}, \alpha)$$

in which case the posterior is of the same form,

$$\mathbf{b}|\mathbf{\Psi}, \mathbf{Y}_T \sim N(\bar{\mathbf{b}}, \mathbf{\Psi} \otimes \bar{\mathbf{\Omega}})$$

$$\mathbf{\Psi}|\mathbf{Y}_T \sim IW(\bar{\mathbf{\Psi}}, T + \alpha)$$

where the posterior parameters $\bar{\mathbf{b}}$, $\bar{\mathbf{\Omega}}$, and $\bar{\mathbf{\Psi}}$ are functions of the data and the prior parameters $\underline{\mathbf{b}}$, $\underline{\mathbf{\Omega}}$, and $\underline{\mathbf{\Psi}}$. As noted by Rothenberg (1963) and detailed by Geweke and Whiteman (2006), this prior encodes prior views that are strongly at odds with the Minnesota prior. In particular, the Kroneker structure of the prior covariance matrix implies that the ratio of the prior standard deviation of the own lag coefficient in the revenue equation to the prior standard deviation of the first lag of revenue in the income equation is the *same* as the ratio of the standard deviation of the coefficient on lagged income in the revenue equation to the standard deviation of the own first lag in the income equation. This is inconsistent with the Minnesota prior notion that more is known about own lags than about others' lags.

To maintain the spirit of the Minnesota prior in a fully Bayesian analysis, it is necessary to abandon the comfort of conjugate priors and instead work with the more general formulation.

Even with independent Minnesota priors for each equation in the form (12) and a diffuse prior for Ψ (as in (2.1)), the posterior for the $\mathbf{B}_j$'s and Ψ is the product of the flat-prior posterior (8) and a factor $p_M(\theta)$ coming from the Minnesota prior. This product form, $p_M(\theta)|\mathbf{\Psi}|^{-(p+1)/2}L\,(\theta;\mathbf{Y}_T^o)$ is not amenable to (much) analytic work, and is not recognizable as another density from which it is possible to sample directly.

Fortunately, it is straightforward to sample from the relevant distribution *indirectly* via importance sampling (Geweke (1989)).[3] We seek a sample

$$\theta^{(i)} \sim p(\theta|\mathbf{Y}_T) \propto p_M(\theta)|\mathbf{\Psi}|^{-(p+1)/2}L\,(\theta;\mathbf{Y}_T^o)\,. \tag{16}$$

We employ the flat-prior posterior as our importance function, $I(\theta) = |\mathbf{\Psi}|^{-(p+1)/2}L\,(\theta;\mathbf{Y}_T^o)$, sample $\theta^{(i)}$ from it, and assign to the drawing $\theta^{(i)}$ a weight $w(\theta^{(i)})$ equal to the ratio $p(\theta^{(i)}|\mathbf{Y}_T)/I(\theta^{(i)})$, which in our case is just $p_M(\theta^{(i)})$. Then under regularity conditions given by Geweke (1989), weighted averages of the form $\bar{g} = \Sigma_m w(\theta^{(i)})g(\theta^{(i)})/\Sigma_m w(\theta^{(i)})$ will approximate expectations $\int g(\theta)p(\theta|\mathbf{Y}_T)d\theta$. The quality of the approximation is influenced by variation in the weights $w(\theta^{(i)})$–many very small weights are associated with inaccurate estimates. The quality of the approximation is assessed via Geweke's measure of *relative numerical efficiency* (RNE), the ratio of the variance of $g(\theta)$ to the asymptotic (in i) variance of the estimator $\bar{g}$. This number is unity for a random sample from the posterior itself; when importance sampling it is typically smaller than unity, and can be interpreted as the effective sample size as a fraction of the Monte Carlo sample.

It should be clear that the hyperparameters of the Minnesota prior will influence the $RMSE$, MAE, and $K-S$ statistics by shifting and reshaping the predictive distributions. Indeed, Doan, Litterman, and Sims (1984) proceeding in hierarchical fashion, specified their prior for the VAR parameters in terms of the hyperparameters and imagined integrating the VAR parameters out, leaving a marginal distribution for the data given the hyperparameters. They then searched over values of the hyperparameters to find high values of this marginal distribution (actually, a stand-in involving the determinant of the covariance matrix of forecast errors). As they put it,

> This looks like a process of "estimating" our "prior" from the data, a most un-Bayesian notion. But it should be clear that in fact we mean this search

[3]See Kadiyala and Karlsson (1997) for an alternative based on the Gibbs sampler.

> as an informal numerical integration over [the hyperparameters]. Conclusions are meant to be averaged across "priors" determined by different [hyperparameters], with weights given by [the marginal density], the measure of fit. (p.5)

As Doan, Litterman, and Sims (1984) implemented this search, it was indirect in two ways. First, they used the not-yet-fully Bayesian mixed estimation procedure described above. Thus their "prior" represented a point in the VAR parameter space toward which standard estimators would be shrunk, as well as the degree of shrinkage. Thus in place of integrating (averaging) over VAR parameters, they focussed on the mixed estimates, which are maximizing values of the product of the likelihood function and the prior conditional on the hyperparameters. Second, even if they had integrated over the VAR parameters, the effort to improve forecast accuracy by giving large weight to those values of the hyperparameters that forecast well upweights or downweights regions of the VAR parameter space only through the conditional distribution of the VAR parameters given the hyperparameters. The next section explores a way of searching *directly* in the space of θ's to achieve the same goal.

4. Best Performing Priors Determined By Entropic Tilting

The approach that we take in this section stems from working with importance samples from posterior distributions for the (VAR) parameters of interest. Suppose, for example, that we have a random sample of size N from the posterior distribution, $\{\theta^{(i)}\}_{i=1}^{N}$. This could be a random sample from the flat-prior posterior $|\mathbf{\Psi}|^{-(p+1)/2}L\left(\theta;\mathbf{Y}_T^o\right)$, with each drawing having weight $1/N$. To create a sample from the posterior distribution associated with the Minnesota prior $p_M(\theta|\Lambda)$, we simply assign each flat-prior-posterior drawing $\theta^{(i)}$ the weight $p_M(\theta^{(i)}|\Lambda) = \pi_i$. A search in the style of Doan, Litterman, and Sims for a prior that improves forecast accuracy would operate on these weights by varying Λ. But with an importance sample $\{\theta^{(i)}\}_{i=1}^{N}$ and associated weights $\{\pi_{(i)}\}_{i=1}^{N}$ (which are normalized to sum to unity and will generally be equal in our applications) in hand, it is natural to seek a better performing "prior" directly by finding a new set of weights $\{\pi^*_{(i)}\}_{i=1}^{N}$. Just as Doan, Litterman, and Sims (1984), we will need to be wary of overfitting; they did this by varying only a small number of hyperparameters and operating through the hierarchical prior structure, while we will do this by penalizing Kullback-Leibler divergence of $\{\pi^*_{(i)}\}_{i=1}^N$ from $\{\pi_{(i)}\}_{i=1}^N$. [4]

4.1. **Entropic Tilting.** We now wish to study how to find a set of $\{\pi^*_{(i)}\}_{i=1}^N$ weights that would have improved (say) $RMSE$'s of forecasts within the sample period. That is, just like Doan, Litterman, and Sims (1984), *after* assembling the posterior via the product of the likelihood and any–possibly diffuse–prior information about the parameters, we wish to *revisit* the sample as if we were forcasting within that time frame, and study the performance of alternative specifications. Since we wish to work with moments and quantiles of predictive *densities* in performance assessment, this process will be somewhat more involved for us than it was for Doan, Litterman, and Sims two decades ago.

Denote by $y^i_{t,t+4}$ draw i of the time-t pseudo-predictive density for revenue at time $t+4$. We focus on the 4-step ahead forecast for illustrative purposes only; it should be clear that we could be working with a vector of forecasts. Here t is understood to be prior to the end of the data sample (of length T) used to generate the posterior drawings $\{\theta^{(i)}\}_{i=1}^N$. Observations subsequent to T will be used in assessing the performance of the procedure in true out-of-sample forecasting. The drawing $y^i_{t,t+4}$ is generated using the VAR, equation (1), the drawing $\theta^{(i)}$ from the posterior, data through t, and samples from a $N(0,\Psi^{(i)})$ distribution to perform a dynamic simulation from the VAR. The predictive distribution $\{y^i_{t,t+4}\}_{i=1}^N$ is the one that would have been produced at time t had the posterior sample $\{\theta^{(i)}\}_{i=1}^N$ been available. Now the squared error of the predictive mean under the distribution $\{\pi^*_{(i)}\}_{i=1}^N$ is given by $(\Sigma_{i=1}^N \pi^*_i y^i_{t,t+4} - y_{t+4})^2$. Like Doan, Litterman, and Sims, we will look for good performance over several periods; that is, small values of $\Sigma_{t=1}^{T-4}\frac{1}{2}(\Sigma_{i=1}^N \pi^*_i y^i_{t,t+4} - y_{t+4})^2$. Generally these will not difficult to find if we have complete freedom to move the $\{\pi^*_{(i)}\}_{i=1}^N$ distribution away from $\{\pi_{(i)}\}_{i=1}^N$. Thus we will penalize such movement (prevent overfitting) in proportion with the Kullback-Leibler divergence $\Sigma_{i=1}^N \pi^*_i ln(\frac{\pi^*_i}{\pi_i})$. Thus the choice problem ("shrinkage for the Monte Carlo era") is

$$\min_{\{\pi^*_{(i)}\}_{i=1}^N} \Sigma_{t=1}^{T-4}\frac{1}{2}(\Sigma_{i=1}^N \pi^*_i y^i_{t,t+4} - y_{t+4})^2 + \nu\Sigma_{i=1}^N \pi^*_i ln(\frac{\pi^*_i}{\pi_i}) - \mu(\Sigma_{i=1}^N \pi^*_i - 1) \tag{17}$$

[4]See Robertson, Tallman, and Whiteman (2005) for a related way of modifying forecasts to reflect moment conditions not used directly in their production.

where ν is a (prespecified) penalty factor and μ is a Lagrange multiplier ensuring that the new weights sum to unity. Rearrangement of the first order conditions yields

$$\pi_i^* = \frac{\pi_i \exp[\nu^{-1}(\Sigma_{j=1}^N \pi_j^* y_{t,t+4}^j - y_{t+4}) y_{t,t+4}^i]}{\Sigma_{i=1}^N \pi_i \exp[\nu^{-1}(\Sigma_{j=1}^N \pi_j^* y_{t,t+4}^j - y_{t+4}) y_{t,t+4}^i]}. \tag{18}$$

The appearance of π_i^* on both sides of this expression complicates solution. We have employed an iterative scheme that in practice convgerges so long as the penalty ν is not set too low.

Equation (18) makes clear that the π_i probabilities are exponentially or entropically *tilted* to produce the π_i^*'s. Of course, reweighting is exactly what transpires in importance sampling, and indeed, the entropic tilt can be interpreted as a reweighting ("tilt") from an initial importance sampler given by the π_i's. It is therefore possible to assess the degree of sample depletion associated with the tilt by examining RNE values for (say) predictive means.

As noted by Robertson, Tallman, and Whiteman (2005), there are several ways to interpret the entropic tilt and the information associated with it. Recall that the initial weights π_i embody information from the likelihood function and any prior information regarding the parameters θ. The new weights π_i^* comprise the product of the weights π_i and a factor reflecting the reweighting to improve forecasting performance. This additional factor can be thought of as different, *data-dependent* prior, as a modification to the original likelihood function, as an update of the posterior for the parameters, or as a modification to the predictive distribution itself. In what follows, we adopt the first of these, and interpret the entropic tilting as accomplishing *directly* the same thing–improvement of forecasts–that Doan, Litterman, and Sims (1984) accomplished indirectly by varying the hyperparameters of the Minnesota prior.

We study the performance of this procedure in practice by carrying it out in pseudo-real time by first using data through T to generate the initial (flat prior) posterior and reweighting to lower the 4-step $RMSE$ in sample. We then use the reweighted posterior to predict $y_{T+1}, y_{T+2}, \ldots, y_{T+h}$ and record results. Next, T is incremented by one time period and the process is repeated; we continue in this fashion until the data are exhausted.

4.2. **Practical Robust Forecasting.** Direct modification of the probabilities associated with drawings from the predictive distribution is very similar to what one would do as a robust response to a situation of ambiguity. Hansen and Sargent (2005) have argued that decision makers (e.g., revenue forecasters) should guard against the possibility that their model (e.g., the VAR, under whatever prior is in use) for the phenomena they face is incorrect by imagining that they are playing a game against an evil "nature" whose objective is to *minimize* the decision maker's welfare by choosing new event probabilities subject to the constraint that the evil probabilities not differ too much from the original ones in the Kullback-Leibler sense. Maccheroni, Marinacci, and Rustichini (2005) provide an axiomatic foundation for these so-called variational preferences by relaxing the independence axiom of expected utility theory.

In the context of forecasting, the "evil agent" game can be thought of as follows. Suppose the loss function to be used by the forecaster is quadratic, and suppose we have an initial sample from the joint distribution for $\mathbf{y}_{T+h}$ and θ,

$$p(\mathbf{y}_{T+h}, \theta \mid \mathbf{Y}_T^o) \propto p\,(\theta \mid \mathbf{Y}_T^o, \theta)\, p(\mathbf{y}_{T+h} \mid \theta, \mathbf{Y}_T^o). \tag{19}$$

Facing a benign nature, the forecaster chooses the forecast y_{T+h}^F to minimize expected loss:

$$\min_{y_{T+h}^F} \int\int \left(y_{T+h}^F - \tilde{y}_{T+h}\right)^2 p(\tilde{\mathbf{y}}_{T+h}, \theta \mid Y_T^o) d\theta d\tilde{\mathbf{y}}_{T+h} \tag{20}$$

but when nature is evil, she confronts the forecaster with a different probability distribution $q(\tilde{\mathbf{y}}_{T+h}, \theta \mid \mathbf{Y}_T^o)$ designed to *maximize* expected loss subject to the constraint that q not diverge too much from p and taking the forecast itself as given:

$$\max_{q(\tilde{\mathbf{y}}_{T+h}, \theta \mid \mathbf{Y}_T^o)} \int\int \left(y_{T+h}^F - \tilde{y}_{T+h}\right)^2 q(\tilde{\mathbf{y}}_{T+h}, \theta \mid \mathbf{Y}_T^o) d\theta d\tilde{\mathbf{y}}_{T+h} - \lambda KL(p:q) \tag{21}$$

where λ is the Lagrange multiplier associated with the constraint, and $KL(p:q)$ is the Kullback-Leibler divergence from p to q. Given the probabilities q, the forecaster minimizes expected (quadratic) loss, and of course chooses the mean. In terms of our sample from the predictive distribution for $\mathbf{y}_{T+h}$ with weights π_i, the problem of finding the Nash equilibrium of the game can be represented as

$$\min_{y_{T+h}^F} \max_{\{\pi_i^*\}_{i=1}^N} \Sigma_{i=1}^N \pi_i^* \left(y_{T+h}^F - y_{T+h}^{(i)}\right)^2 - \lambda \Sigma_{i=1}^N \pi_i^* ln(\frac{\pi_i^*}{\pi_i}) - \mu(\Sigma_{i=1}^N \pi_i^* - 1) \tag{22}$$

where the constraint with multiplier μ ensures adding up. The multiplier λ controls how much latitude the evil agent has in altering the probabilities; when λ is very large, little adjustment to the probabilities occurs, and when λ is small enough, the evil agent has great freedom to manipulate the probabilities delivered to the forecaster. [5] The solution to the problem is that the forecast is the predictive mean $\Sigma\pi_i^* y_{T+h}^{(i)}$ and the evil agent tilts the probabilities according to

$$\pi_i^* = \frac{\pi_i \exp[\lambda^{-1}(\Sigma_j \pi_j^* y_{T+h}^{(j)} - y_{T+h}^{(i)})^2]}{\Sigma_i \pi_i \exp[\lambda^{-1}(\Sigma_j \pi_j^* y_{T+h}^{(j)} - y_{T+h}^{(i)})^2]}. \tag{23}$$

As before, π_i^*'s appear on both sides, and this is in general a difficult problem to solve. Our experience is once again that iteration finds a solution quickly for moderate values of the penalty factor, but that the procedure does not converge for very small values.

5. Results

The forecasting procedures of the previous section are assessed in Table 3 and Figures 6 through 11. All results in this section are based on quarterly 1-8 step-ahead out-of-sample forecast performance over the period 1995:III-2005:III.[6] The specification search for the Minnesota prior was based on an initial estimate (in the mixed estimation case) or posterior analysis (for the fully Bayesian implementation) using data for the period 1982:III-1990:II. To choose the hyperparameter specification on which to focus, we fixed the decay rate at unity, and for a grid of values for overall tightness and others' weight, we made 1-8 step forecasts beginning in 1990:III, updated the data and repeated the process for forecasts beginning in 1990:IV, and so on until 1995:II data had been used. The fully Bayesian implementation of the Minnesota prior, and the entropically tilted and robust forecasts are based on importance samples from the flat prior posterior using data through 1995:II for the first forecast, through 1995:III for the second, and so on.

[5]In the setup of Hansen and Sargent (2005), there is a lower bound on λ, the "point of utter psychotic despair" (a term Hansen and Sargent attribute to Whittle) at which the evil agent's objective ceases to be concave, and the "forecaster" agent could be made unboundedly miserable. In our setup, the second order conditions do not change sign, so there does not seem to be a breakdown point.

[6]Recall that the forecast record reported for the real-time forecasts was for annual revenue growth. We used quarterly figures here to ensure a reasonably-sized sample of forecast errors. Predicting quarterly revenues is more difficult than predicting annual sums, so we expect the forecast error statistics to be larger for the quarterly predictions.

The best performing mixed-estimation Minnesota prior over this period involved overall tightness 0.3 and others' weight 1.0, while the best performing fully Bayesian implementation of the Minnesota prior involved overall tightness 0.8 and others' weight 1.0. Figure 4 represents the search for four horizons in the mixed estimation case. The vertical axes in each case are given by $-1 \times RMSE$ (so that higher is better). The base axes are settings of the hyperparameter values. Across the four horizons considered, local variation in the hyperparameters around the values selected for further analysis is relatively unimportant. Figures 5 and 8 summarize the hyperparameter search for the fully Bayesian procedure. For both the $RMSE$ and the Kolmogorov-Smirnov statistics, the best others' weight is about the same as for the mixed estimates forecasts, but overall tightness is somewhat higher. We believe this reflects the contrast between the single-equation focus of the mixed estimates and the system focus of the fully Bayesian approach: overall tightness applies simulaneously to all equations in the latter.

Once the hyperparameter settings were chosen in the holdout sample, they were not changed. At each date in the assessment period, as much data as were available at that date were used either in producing mixed estimates, fully Bayesian posterior distributions, or entropically reweighted predictives. The entropic reweighting used the previous 20 quarters of data at each forecast date. No "peeking" was allowed.

Table 3 reports results for the flat- and Minnesota-prior mixed estimation forecasts, as well as density forecasts from the flat and Minnesota priors, and the robust and "evil agent" forecasts. For the latter two, we used penalty factors that permitted as much freedom to tilt as we could allow and still obtain convergence of our iterative computational scheme. For each of the density forecast procedures, we report $RMSE$'s and MAE's of forecasts based on the predictive mean, median, and 25% quantile (the "LR" column, for (3-1) "loss ratio"). We also report Kolmogorov-Smirnov statistics at each horizon, and an example of RNE values of the predictive means at each horizon computed from the forecast made 2001:III. The empirical distributions of quantile realizations underlying the Kolmogorov-Smirnov statistics are detailed in Figures 7 through 10. The latter figures indicate that the quarterly pseudo-real-time experiments from 1995:III-2005:III yielded the same phenomenon that characterizes the real-time record: too many left tail values were realized for revenue predictions relative to the predictive densities.

The results in Table 3 are for predicting log revenue; the root mean squared error and mean absolute error entries can be interpreted (roughly) as percentage errors. Thus the 3-step $RMSE$ for the mixed-estimate Minnesota prior was just under 4.4% of the level of revenue. The statistics in the Table tend to confirm two long-standing suspicions and suggest two new ones.

The first of these suspicions was that there is much to be gained by implementing a fully Bayesian procedure that enables production of density forecasts, even if these are used to produce a natural point forecast under a reasonable loss function. For example, at horizons up to a year, the point forecasts produced from the fully Bayesian predictives under quadratic loss (the mean), 1-1 loss ratio (the median), and even the 3-1 loss ratio (the 25% point of the predictive) have lower $RMSE$ and MAE statistics than the forecasts produced with point estimates – whether shrinkage to the Minnesota prior is pursued or not. To the extent that overfitting drives poor out-of-sample forecasting performance, careful treatment of the uncertainty in the parametric specification as embodied in the fully Bayesian predictive distribution may compensate somewhat.

The second of these was that there not much was to be gained by adding shrinkage to a fully Bayesian flat prior forecast. Improvements of the best-performing Minnesota prior over the flat prior are modest at best, at least in terms of $RMSE$ and MAE. Comparing the quality of the entire density forecast via Kolmogorov-Smirnov statistics suggests that there may be some improvement in matching quantiles, but the p-values of these statistics generally exceed 20%, and in any case p-values do not represent strength of evidence.

The first new suggestion is that entropic tilting does not seem to improve out of sample forecasts. That is, despite the freedom to tinker with the prior permitted by the reweighting scheme, and despite the average 50% reduction in the in-sample $RMSE$ values, the out of sample forecasts from 1995:III on are little changed. It turns out that the tilt is accomplished with surprisingly little reweighting, and this is apparent in the "Lorenz" curve of weights in Figure 11. The figure depicts the cumulative fraction of the importance sample weight accounted for by draw i (sorted in ascending order); the 45-degree line characterizes the weights from a random sample; tilting and the importance-sampled Minnesota prior are associated with unequal weights, and this is represented by curves lying below the 45-degree line. The entropic tilt produces less inequality in the weights than does the Minnesota prior.

Put another way, the Minnesota prior adds *more* information than does the entropic tilt. Of course the Minnesota prior is doing something other than achieving a reduction in forecast error in sample, and this is reflected in the weight inequality.

The second new suggestion is that the spread in the forecast distribution accomplished by robustifying the predictions did not much improve out of sample forecast performance. The robust forecast does a little better than the flat-prior forecast, as is to be expected given that the latter produced too many left-tail values, and robustifying thickens *both* tails, thereby ensuring at least a slight increase in the quantiles of realized forecasts. But the effect of the tail thickening was slight.

Figures 12 and 13 provide some further insight into what the Minnesota prior and the tilt are doing. Figure 12 shows the posterior density of the own lag coefficient in the revenue equation (where "posterior" under the robust procedure and the entropic tilt means "reweighted flat-prior posterior"). As expected, the Minnesota prior pulls the own-lag coefficient distribution toward the prior mean value of unity, though this effect is not too dramatic. Robustification and entropic tilting spread the distribution a little, but not too much.

Figure 12 is typical, and shows the effect on the three-step-ahead prediction made at the end of the third quarter of 2001. The robust prediction is clearly a spread, though a small one. The other three predictives are very similar to each other; the flat prior is a little more concentrated about the mode than the other two, but the effect is minimal. That there is relatively little effect visible after reweighting suggests that at least the predictions are relatively insensitive to quite different prior views. Yet this is actually of some comfort, just as it was to Doan, Litterman, and Sims (1984):

> While our explorations are in some ways like fitting the parameters of a conventional model–we examine various points in a parameter space and check how well the resulting models fit the data–the motivation and implications of the results are different in important respects. Our ideal conclusion would be that the parameters are "ill determined"–that the fit is similar across a wide range of parameter settings having similar implications. (p.4)

In our case, it has indeed turned out that the result are insensitive, not just to hyperparameters of a Minnesota prior, but even to direct manipulation of prior probabilities in an unabashed attempt to improve forecast performance.

6. Denouement and Conclusion

This paper has reported on an attempt to improve forecast performance in a pragmatic, locally important setting–the prediction of state tax revenues in Iowa. Over a sixteen year period including two national recessions and a stock market boom, the record, as measured by root mean squared errors, is within the 3% error bounds set by the State Comptroller at the beginning of the period. More importantly, the production of density forecasts has helped educate state officials of the uncertainties of economic forecasting. Yet on occasion even this effort fell short: two forecasts made around the time of the most recent recession predicted that what actually happened over the ensuing nine months was essentially impossible. Thus it seems reasonable to expect that some improvement could be had by using alternative specifications of variables, lag lengths employed in VARs, prior specifications, etc. But we have found little, despite introducing two new tools for modifying prior distributions in direct attempts to improve forecast performance.

Ex-post, it is even difficult to reproduce the most serious overprediction. Figure 13 actually depicts the forecast made at about the time the largest real-time forecast error (four predictive standard deviations from the mean) was made. Yet the subsequently realized revenue value is not far from the mean of the predictive density in the figure. How can this be? We believe the answer lies in the quality of the data. We have carried out our improvement search using a fixed data set, the most recent one available to us. Because income data are revised periodically, and these revisions can be quite large in an agricultural state like Iowa, it would be useful to undertake a serious real-time-data study of the issues taken up in this paper. But the cost would be enormous, as such data archiving that did take place crossed multiple generations of media, from 5.25" floppy discs through tape backups, ZIP disks, etc. to external hard drives.

In place of a real-time analysis, we examined the minutiae of forecasts made around the time the largest forecast errors were made. One of the problems that characterized the time was that sales tax revenue data were "contaminated" by phantom revenues that were collected by the state and subsequently reverted to its constituent counties (subjurisdictions.) The amount of phantom revenues and other adjustments (e.g., for differences in the number of revenue processing days in successive years) were estimated by the Department of Revenue and removed from the data prior to the production of the forecasts. This turned out to be bad practice, as the Department's estimates were quite wide of the mark. For example, adjustments to revenues for July, August, and September revenues estimated at the time the November 2001 forecast overstated revenues by about 1.5% of *annual* revenues for the year. This much data error in a VAR, whose forecasts are heavily influenced by the most recent observations, has dramatic effect on the short-term forecasts it produces. Had the adjustments not been made, the realized value would have been within the interquartile range of the predictive distribution.

This anecdotal experience, and the work for the paper, serve as reminders that there is no substitute for better data or a better model. We're still looking for both.

References

Berger, J. O., and R. Wolpert (1988): *The Likelihood Principle*. Institute of Mathematical Statistics, Hayward, CA, second edn.

Diebold, F. X., T. A. Gunther, and A. S. Tay (1998): "Evaluating Density Forecasts with Applications to Financial Risk Management," *International Economic Review*, 39(4), 863–883.

Doan, T., R. B. Litterman, and C. A. Sims (1984): "Forecasting and Conditional Projection Using Realistic Prior Distributions," *Econometric Reviews*, 3, 1–100.

Doan, T. A. (2004): *RATS Version 6.0*. Estima, Evanston, IL.

Fair, R. C. (1979): "An Analysis of the Accuracy of Four Macroeconomic Models," *Journal of Political Economy*, 87, 701–718.

Geweke, J. (1989): "Bayesian Inference in Econometric Models Using Monte Carlo Integration," *Econometrica*, 57, 1317–1340.

Geweke, J. F., and C. H. Whiteman (2006): "Bayesian Forecasting," in *Handbook of Economic Forecasting*, ed. by G. Elliott, C. W. Granger, and A. Timmermann. Elsevier.

Hansen, L. P., and T. J. Sargent (2005): *Misspecification in Recursive Macroeconomic Theory*. Princeton University Press.

James, W., and C. Stein (1961): "Estimation with Quadratic Loss," in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability*, pp. 361–379. University of California Press, Berkeley.

Kadiyala, K. R., and S. Karlsson (1993): "Forecasting with generalized Bayesian vector autoregressions," *Journal of Forecasting*, 12, 365–378.

——— (1997): "Numerical Methods for Estimation and Inference in Bayesian VAR Models," *Journal of Applied Econometrics*, 12(2), 99–132.

Litterman, R. B. (1979): "Techniques of Forecasting Using Vector Autoregressions," Ph.D. thesis, University of Minnesota.

——— (1986): "Forecasting with Bayesian vector autoregressions - 5 years of experience," *The Journal of Business and Economic Statistics*, 4, 25–38.

Maccheroni, F., M. Marinacci, and A. Rustichini (2005): "Ambiguity Aversion, Malevolent Nature, and the Variational Representation of Preferences," unpublished manuscript, Universities of Bocconi, Torino, and Minnesota.

Otrok, C., and C. H. Whiteman (1998): "What to Do When the Crystal Ball is Cloudy: Conditional and Unconditional Forecasting in Iowa," *Proceedings of the National Tax Association, Ninetieth Annual Conference, 1997*, pp. 326–334.

Robertson, J. C., E. Tallman, and C. H. Whiteman (2005): "Forecasting Using Relative Entropy," *Journal of Money, Credit, and Banking*, 37(3), 383–401.

Rothenberg, T. J. (1963): "A Bayesian analysis of simultaneous equation systems," Report 6315, Econometric Institute, Netherlands School of Economics.

Sims, C. A. (1980): "Macroeconomics and Reality," *Econometrica*, 48(1).

Theil, H. (1963): "On the Use of Incomplete Prior Information in Regression Analysis," *Journal of the American Statistical Association*, 58, 401–414.

Whiteman, C. H. (1996): "Bayesian Prediction Under Asymmetric Linear Loss: Forecasting State Tax Revenues in Iowa," in *Modelling and Prediction Honoring Seymour Geisser*, ed. by J. Lee, W. Johnson, and A. Zellner, pp. 149–168. Springer, New York, NY.

Zellner, A. (1971): *An Introduction to Bayesian Inference in Econometrics*. Wiley, New York, NY.

	Mixed Estimation		Fully Bayesian Density Estimation											
	Flat	MN	Flat Prior			Minnesota Prior			Entropic Robustification			Modern DLS		
	Pt	*Pt*	*Mn*	*Md*	*LR*	*Mn*	*Md*	*LR*	*Mn*	*Md*	*LR*	*Mn*	*Md*	*LR*
RMSE														
1	0.0350	0.0421	0.0323	0.0324	0.0353	0.0323	0.0322	0.0350	0.0322	0.0322	0.0354	0.0318	0.0318	0.0350
2	0.0415	0.0523	0.0373	0.0373	0.0427	0.0377	0.0374	0.0418	0.0373	0.0371	0.0431	0.0369	0.0370	0.0424
3	0.0422	0.0437	0.0372	0.0372	0.0417	0.0368	0.0369	0.0391	0.0371	0.0372	0.0420	0.0375	0.0375	0.0416
4	0.0473	0.0422	0.0463	0.0463	0.0493	0.0461	0.0463	0.0476	0.0459	0.0459	0.0496	0.0467	0.0468	0.0501
5	0.0622	0.0568	0.0580	0.0581	0.0612	0.0575	0.0576	0.0595	0.0575	0.0577	0.0616	0.0580	0.0581	0.0616
6	0.0693	0.0657	0.0652	0.0652	0.0688	0.0642	0.0640	0.0660	0.0649	0.0650	0.0697	0.0656	0.0658	0.0700
7	0.0734	0.0651	0.0699	0.0699	0.0739	0.0687	0.0687	0.0702	0.0700	0.0697	0.0746	0.0708	0.0709	0.0753
8	0.0779	0.0670	0.0783	0.0779	0.0835	0.0771	0.0774	0.0798	0.0777	0.0774	0.0840	0.0795	0.0790	0.0853
MAE														
1	0.0261	0.0342	0.0252	0.0253	0.0265	0.0258	0.0256	0.0269	0.0249	0.0249	0.0265	0.0248	0.0248	0.0264
2	0.0333	0.0436	0.0306	0.0305	0.0322	0.0312	0.0309	0.0330	0.0304	0.0302	0.0324	0.0299	0.0300	0.0316
3	0.0336	0.0356	0.0297	0.0297	0.0313	0.0303	0.0302	0.0293	0.0294	0.0296	0.0316	0.0298	0.0297	0.0311
4	0.0384	0.0337	0.0366	0.0367	0.0375	0.0363	0.0363	0.0363	0.0363	0.0362	0.0378	0.0374	0.0376	0.0373
5	0.0503	0.0445	0.0467	0.0470	0.0472	0.0458	0.0464	0.0460	0.0464	0.0466	0.0472	0.0476	0.0477	0.0463
6	0.0562	0.0531	0.0528	0.0529	0.0517	0.0506	0.0507	0.0499	0.0527	0.0529	0.0526	0.0538	0.0538	0.0520
7	0.0586	0.0517	0.0561	0.0566	0.0554	0.0559	0.0560	0.0524	0.0558	0.0563	0.0561	0.0573	0.0575	0.0557
8	0.0639	0.0530	0.0629	0.0634	0.0583	0.0632	0.0637	0.0574	0.0621	0.0629	0.0585	0.0640	0.0642	0.0592
KS Stat														
1	——	——		0.1671			0.1552			0.1661			0.1455	
2	——	——		0.1756			0.1560			0.1616			0.1796	
3	——	——		0.1848			0.2366			0.1818			0.1813	
4	——	——		0.1724			0.1990			0.1659			0.2000	
5	——	——		0.2163			0.2423			0.2153			0.2225	
6	——	——		0.2135			0.2297			0.2121			0.2131	
7	——	——		0.2733			0.3167			0.2653			0.2707	
8	——	——		0.2903			0.3121			0.2887			0.3071	
RNE														
1	——	——		1.0000			0.2800			0.2393			0.4513	
2	——	——		1.0000			0.2253			0.7167			0.5064	
3	——	——		1.0000			0.2564			0.7108			0.4760	
4	——	——		1.0000			0.2964			0.1850			0.3750	
5	——	——		1.0000			0.3002			0.1579			0.4067	
6	——	——		1.0000			0.2413			0.4477			0.4991	
7	——	——		1.0000			0.2488			0.2232			0.4888	
8	——	——		1.0000			0.2846			0.1229			0.3526	

TABLE 3. Forecast Performance Statistics.

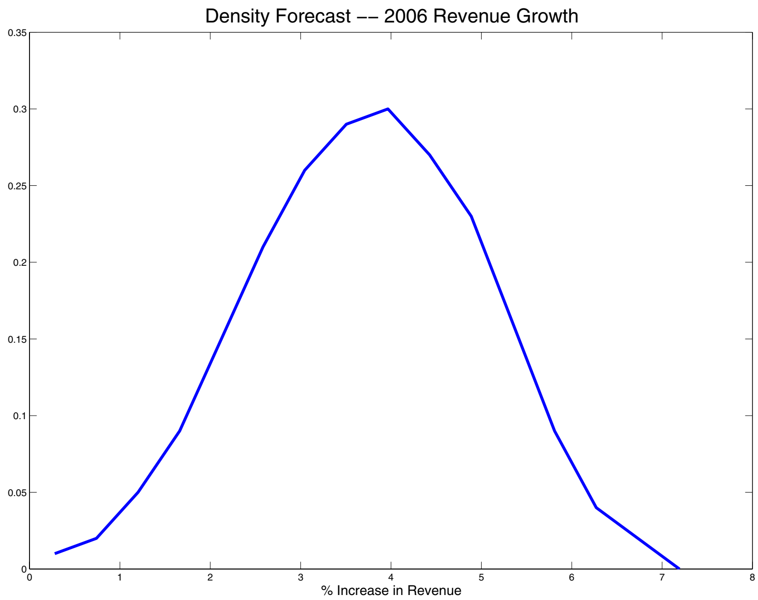

FIGURE 1. Density Forecast, Revenue Growth in Fiscal Year 2006

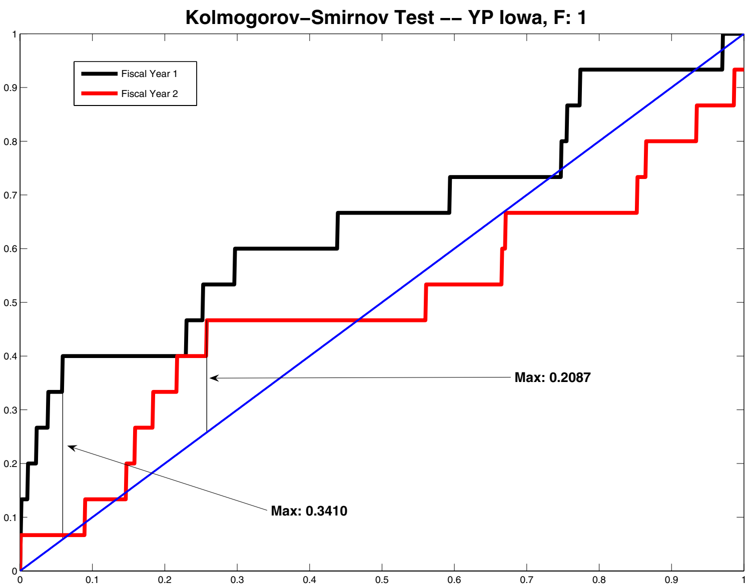

FIGURE 2. Realized Komolgorov-Smirnov Statistics

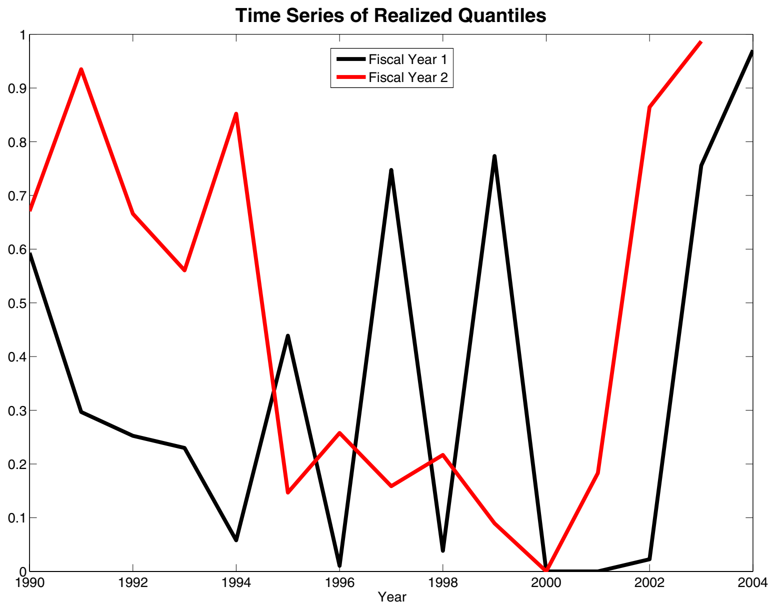

FIGURE 3. Time Series of Realized Quantiles

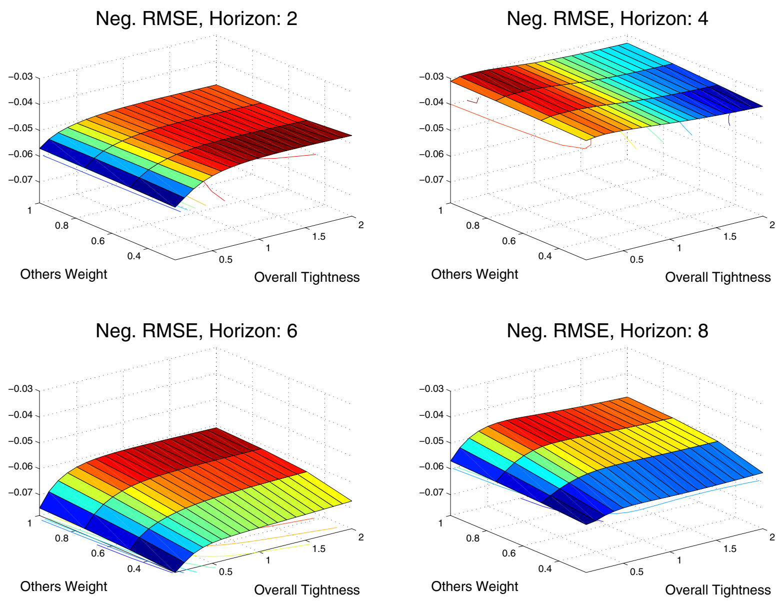

FIGURE 4. Negative RMSE and the MN Hyperparameters – Mixed Estimation.

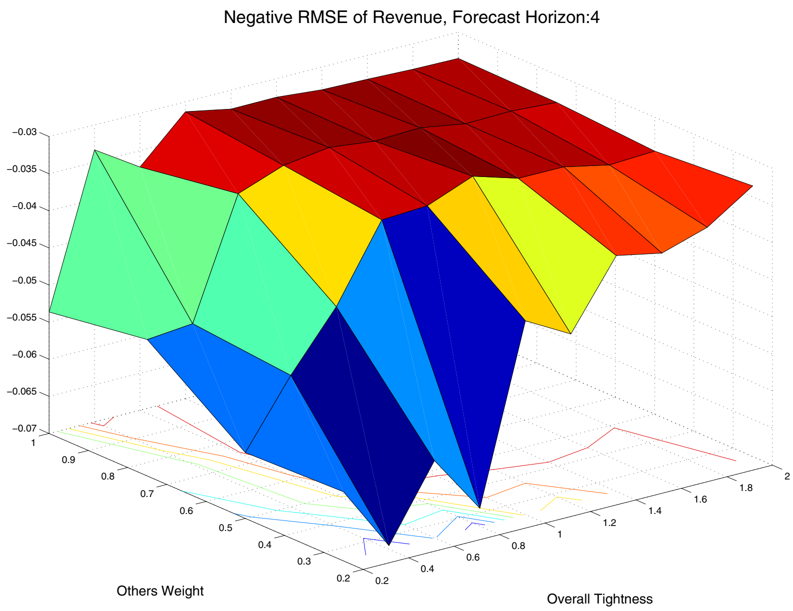

FIGURE 5. Negative RMSE and the MN Hyperparameters – Fully Bayesian.

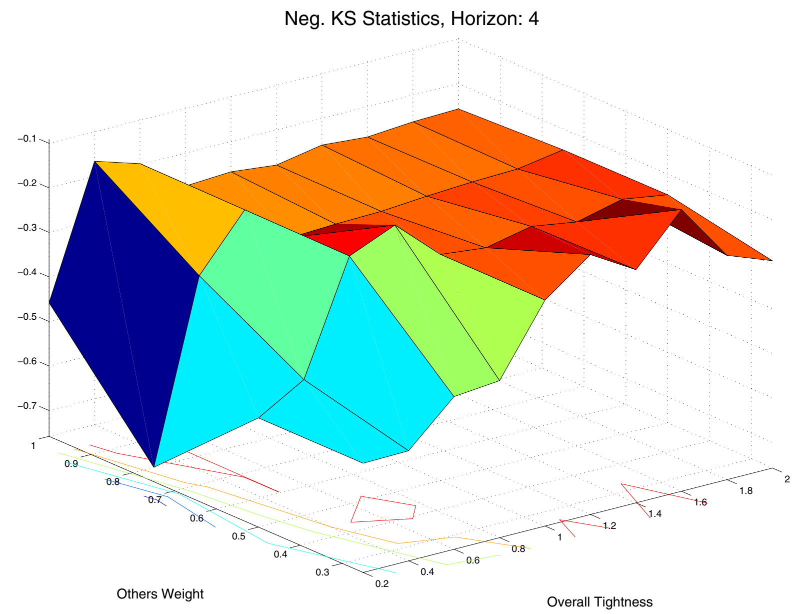

FIGURE 6. KS Surface for MN Prior Hyperparameters.

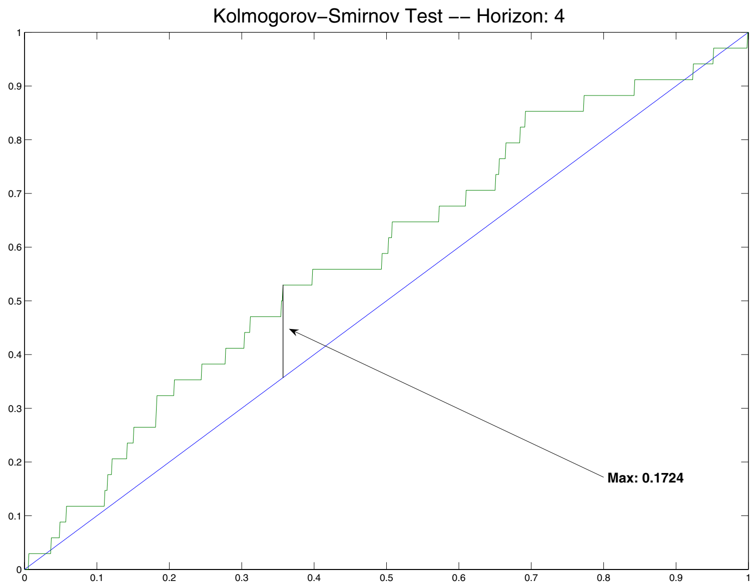

FIGURE 7. KS Plot for the Flat Prior.

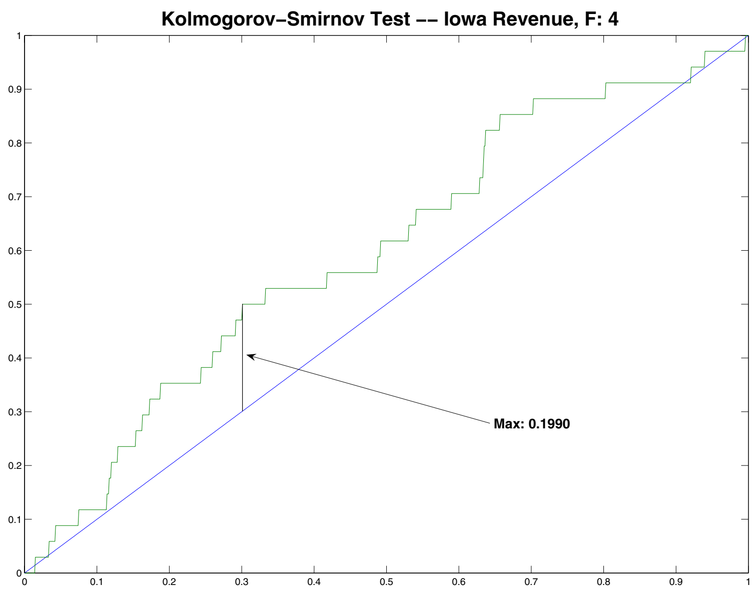

FIGURE 8. KS Plot for the MN Prior.

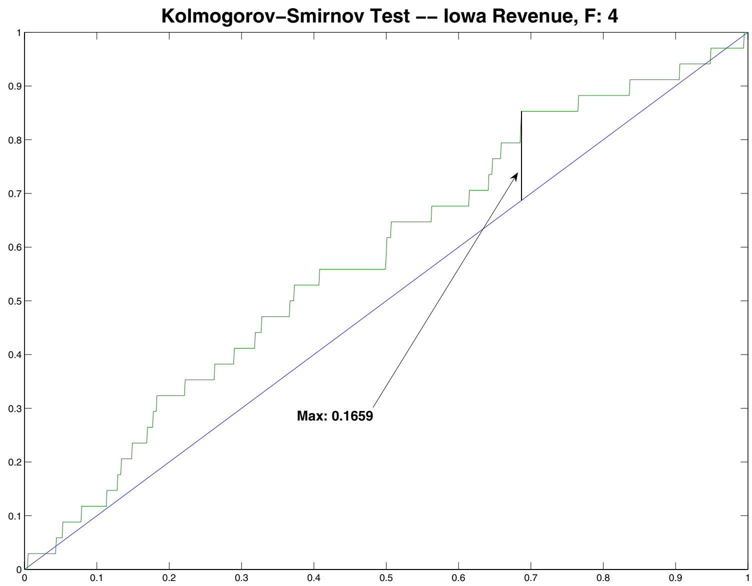

FIGURE 9. KS Plot for Robust Model.

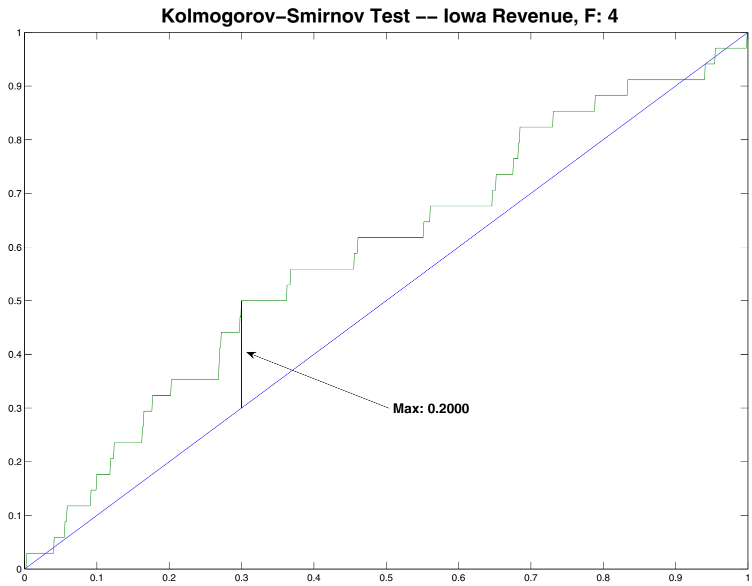

FIGURE 10. KS Plot for the Entropic Reweighting Algorithm.

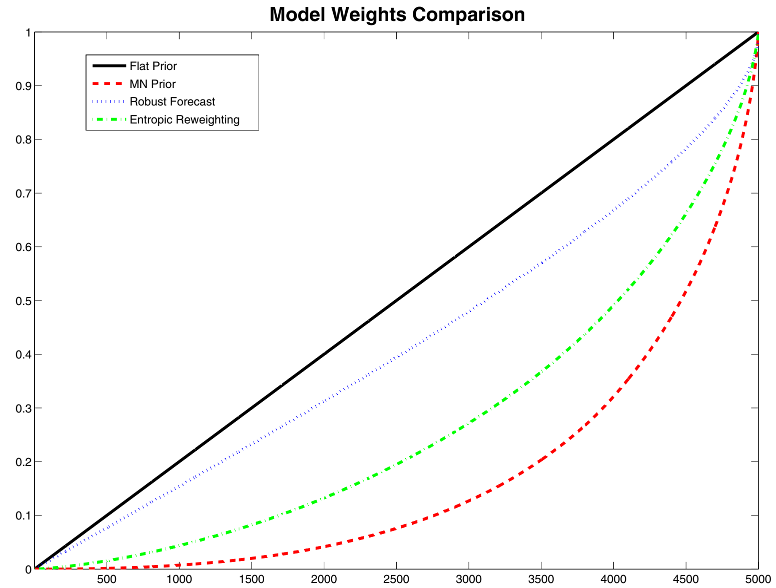

FIGURE 11. Comparison of Weights in Various Models

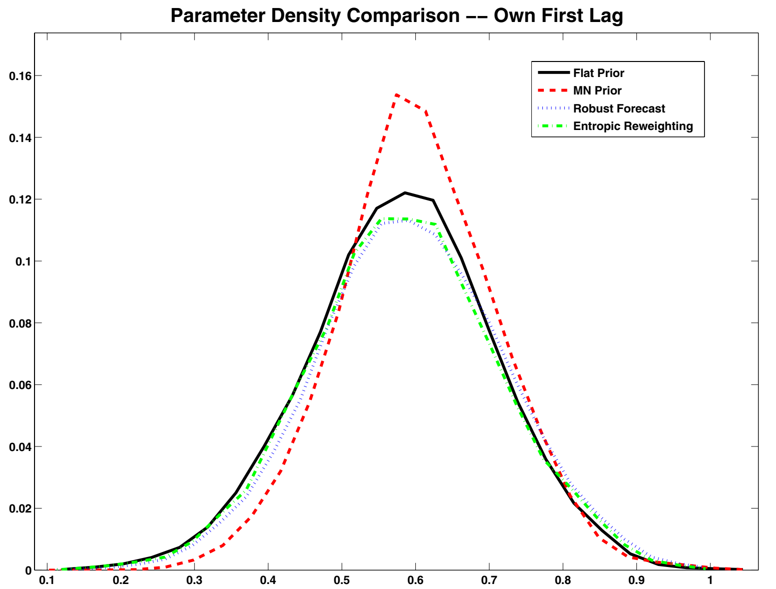

FIGURE 12. Comparison of the Density of the First Lag of Revenue in the Revenue Equation

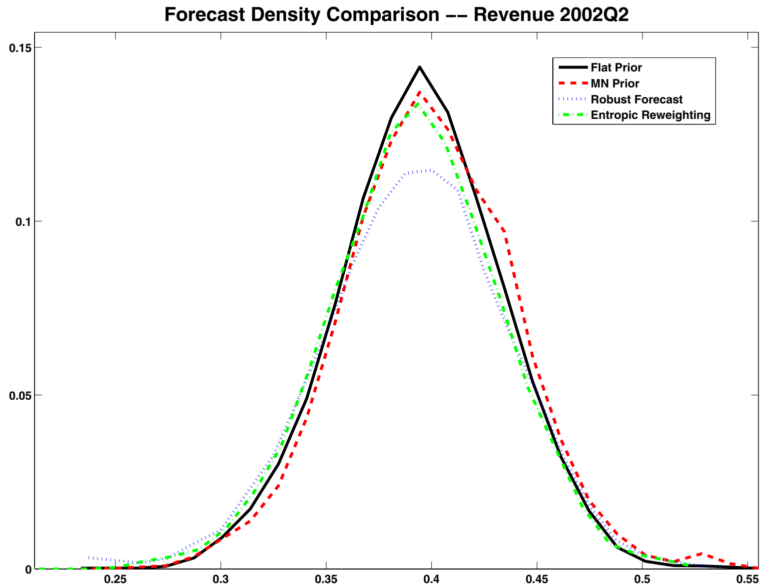

FIGURE 13. Comparison of the Predictive Density for Revenue in 2002 Quarter 2

The following Discussion Papers have been published since 2005:

Series 1: Economic Studies

1	2005	Financial constraints and capacity adjustment in the United Kingdom – Evidence from a large panel of survey data	Ulf von Kalckreuth Emma Murphy
2	2005	Common stationary and non-stationary factors in the euro area analyzed in a large-scale factor model	Sandra Eickmeier
3	2005	Financial intermediaries, markets, and growth	F. Fecht, K. Huang, A. Martin
4	2005	The New Keynesian Phillips Curve in Europe: does it fit or does it fail?	Peter Tillmann
5	2005	Taxes and the financial structure of German inward FDI	Fred Ramb A. J. Weichenrieder
6	2005	International diversification at home and abroad	Fang Cai Francis E. Warnock
7	2005	Multinational enterprises, international trade, and productivity growth: Firm-level evidence from the United States	Wolfgang Keller Steven R. Yeaple
8	2005	Location choice and employment decisions: a comparison of German and Swedish multinationals	S. O. Becker, K. Ekholm, R. Jäckle, M.-A. Muendler
9	2005	Business cycles and FDI: evidence from German sectoral data	Claudia M. Buch Alexander Lipponer
10	2005	Multinational firms, exclusivity, and the degree of backward linkages	Ping Lin Kamal Saggi

11	2005	Firm-level evidence on international stock market comovement	Robin Brooks Marco Del Negro
12	2005	The determinants of intra-firm trade: in search for export-import magnification effects	Peter Egger Michael Pfaffermayr
13	2005	Foreign direct investment, spillovers and absorptive capacity: evidence from quantile regressions	Sourafel Girma Holger Görg
14	2005	Learning on the quick and cheap: gains from trade through imported expertise	James R. Markusen Thomas F. Rutherford
15	2005	Discriminatory auctions with seller discretion: evidence from German treasury auctions	Jörg Rocholl
16	2005	Consumption, wealth and business cycles: why is Germany different?	B. Hamburg, M. Hoffmann, J. Keller
17	2005	Tax incentives and the location of FDI: evidence from a panel of German multinationals	Thiess Buettner Martin Ruf
18	2005	Monetary Disequilibria and the Euro/Dollar Exchange Rate	Dieter Nautz Karsten Ruth
19	2005	Berechnung trendbereinigter Indikatoren für Deutschland mit Hilfe von Filterverfahren	Stefan Stamfort
20	2005	How synchronized are central and east European economies with the euro area? Evidence from a structural factor model	Sandra Eickmeier Jörg Breitung
21	2005	Asymptotic distribution of linear unbiased estimators in the presence of heavy-tailed stochastic regressors and residuals	J.-R. Kurz-Kim S.T. Rachev G. Samorodnitsky

22	2005	The Role of Contracting Schemes for the Welfare Costs of Nominal Rigidities over the Business Cycle	Matthias Paustian
23	2005	The cross-sectional dynamics of German business cycles: a bird's eye view	J. Döpke, M. Funke S. Holly, S. Weber
24	2005	Forecasting German GDP using alternative factor models based on large datasets	Christian Schumacher
25	2005	Time-dependent or state-dependent price setting? – micro-evidence from German metal-working industries –	Harald Stahl
26	2005	Money demand and macroeconomic uncertainty	Claus Greiber Wolfgang Lemke
27	2005	In search of distress risk	J. Y. Campbell, J. Hilscher, J. Szilagyi
28	2005	Recursive robust estimation and control without commitment	Lars Peter Hansen Thomas J. Sargent
29	2005	Asset pricing implications of Pareto optimality with private information	N. R. Kocherlakota Luigi Pistaferri
30	2005	Ultra high frequency volatility estimation with dependent microstructure noise	Y. Aït-Sahalia, P. A. Mykland, L. Zhang
31	2005	Umstellung der deutschen VGR auf Vorjahrespreisbasis – Konzept und Konsequenzen für die aktuelle Wirtschaftsanalyse sowie die ökonometrische Modellierung	Karl-Heinz Tödter

32	2005	Determinants of current account developments in the central and east European EU member states – consequences for the enlargement of the euro erea	Sabine Herrmann Axel Jochem
33	2005	An estimated DSGE model for the German economy within the euro area	Ernest Pytlarczyk
34	2005	Rational inattention: a research agenda	Christopher A. Sims
35	2005	Monetary policy with model uncertainty: distribution forecast targeting	Lars E.O. Svensson Noah Williams
36	2005	Comparing the value revelance of R&D reporting in Germany: standard and selection effects	Fred Ramb Markus Reitzig
37	2005	European inflation expectations dynamics	J. Döpke, J. Dovern U. Fritsche, J. Slacalek
38	2005	Dynamic factor models	Sandra Eickmeier Jörg Breitung
39	2005	Short-run and long-run comovement of GDP and some expenditure aggregates in Germany, France and Italy	Thomas A. Knetsch
40	2005	A“wreckers theory” of financial distress	Ulf von Kalckreuth
41	2005	Trade balances of the central and east European EU member states and the role of foreign direct investment	Sabine Herrmann Axel Jochem
42	2005	Unit roots and cointegration in panels	Jörg Breitung M. Hashem Pesaran
43	2005	Price setting in German manufacturing: new evidence from new survey data	Harald Stahl

1	2006	The dynamic relationship between the Euro overnight rate, the ECB's policy rate and the term spread	Dieter Nautz Christian J. Offermanns
2	2006	Sticky prices in the euro area: a summary of new micro evidence	Álvarez, Dhyne, Hoeberichts Kwapil, Le Bihan, Lünnemann Martins, Sabbatini, Stahl Vermeulen, Vilmunen
3	2006	Going multinational: What are the effects on home market performance?	Robert Jäckle
4	2006	Exports versus FDI in German manufacturing: firm performance and participation in international markets	Jens Matthias Arnold Katrin Hussinger
5	2006	A disaggregated framework for the analysis of structural developments in public finances	Kremer, Braz, Brosens Langenus, Momigliano Spolander
6	2006	Bond pricing when the short term interest rate follows a threshold process	Wolfgang Lemke Theofanis Archontakis
7	2006	Has the impact of key determinants of German exports changed? Results from estimations of Germany's intra euro-area and extra euro-area exports	Kerstin Stahn
8	2006	The coordination channel of foreign exchange intervention: a nonlinear microstructural analysis	Stefan Reitz Mark P. Taylor
9	2006	Capital, labour and productivity: What role do they play in the potential GDP weakness of France, Germany and Italy?	Antonio Bassanetti Jörg Döpke, Roberto Torrini Roberta Zizza
10	2006	Real-time macroeconomic data and ex ante predictability of stock returns	J. Döpke, D. Hartmann C. Pierdzioch

11	2006	The role of real wage rigidity and labor market frictions for unemployment and inflation dynamics	Kai Christoffel Tobias Linzert
12	2006	Forecasting the price of crude oil via convenience yield predictions	Thomas A. Knetsch
13	2006	Foreign direct investment in the enlarged EU: do taxes matter and to what extent?	Guntram B. Wolff
14	2006	Inflation and relative price variability in the euro area: evidence from a panel threshold model	Dieter Nautz Juliane Scharff
15	2006	Internalization and internationalization under competing real options	Jan Hendrik Fisch
16	2006	Consumer price adjustment under the microscope: Germany in a period of low inflation	Johannes Hoffmann Jeong-Ryeol Kurz-Kim
17	2006	Identifying the role of labor markets for monetary policy in an estimated DSGE model	Kai Christoffel Keith Küster Tobias Linzert
18	2006	Do monetary indicators (still) predict euro area inflation?	Boris Hofmann
19	2006	Fool the markets? Creative accounting, fiscal transparency and sovereign risk premia	Kerstin Bernoth Guntram B. Wolff
20	2006	How would formula apportionment in the EU affect the distribution and the size of the corporate tax base? An analysis based on German multinationals	Clemens Fuest Thomas Hemmelgarn Fred Ramb

21	2006	Monetary and fiscal policy interactions in a New Keynesian model with capital accumulation and non-Ricardian consumers	Campbell Leith Leopold von Thadden
22	2006	Real-time forecasting and political stock market anomalies: evidence for the U.S.	Martin Bohl, Jörg Döpke Christian Pierdzioch
23	2006	A reappraisal of the evidence on PPP: a systematic investigation into MA roots in panel unit root tests and their implications	Christoph Fischer Daniel Porath
24	2006	Margins of multinational labor substitution	Sascha O. Becker Marc-Andreas Mündler
25	2006	Forecasting with panel data	Badi H. Baltagi
26	2006	Do actions speak louder than words? Household expectations of inflation based on micro consumption data	Atsushi Inoue Lutz Kilian Fatma Burcu Kiraz
27	2006	Learning, structural instability and present value calculations	H. Pesaran, D. Pettenuzzo A. Timmermann
28	2006	Empirical Bayesian density forecasting in Iowa and shrinkage for the Monte Carlo era	Kurt F. Lewis Charles H. Whiteman

Series 2: Banking and Financial Studies

1	2005	Measurement matters – Input price proxies and bank efficiency in Germany	Michael Koetter
2	2005	The supervisor's portfolio: the market price risk of German banks from 2001 to 2003 – Analysis and models for risk aggregation	Christoph Memmel Carsten Wehn
3	2005	Do banks diversify loan portfolios? A tentative answer based on individual bank loan portfolios	Andreas Kamp Andreas Pfingsten Daniel Porath
4	2005	Banks, markets, and efficiency	F. Fecht, A. Martin
5	2005	The forecast ability of risk-neutral densities of foreign exchange	Ben Craig Joachim Keller
6	2005	Cyclical implications of minimum capital requirements	Frank Heid
7	2005	Banks' regulatory capital buffer and the business cycle: evidence for German savings and cooperative banks	Stéphanie Stolz Michael Wedow
8	2005	German bank lending to industrial and non-industrial countries: driven by fundamentals or different treatment?	Thorsten Nestmann
9	2005	Accounting for distress in bank mergers	M. Koetter, J. Bos, F. Heid C. Kool, J. Kolari, D. Porath
10	2005	The eurosystem money market auctions: a banking perspective	Nikolaus Bartzsch Ben Craig, Falko Fecht
11	2005	Financial integration and systemic risk	Falko Fecht Hans Peter Grüner

12	2005	Evaluating the German bank merger wave	Michael Koetter
13	2005	Incorporating prediction and estimation risk in point-in-time credit portfolio models	A. Hamerle, M. Knapp, T. Liebig, N. Wildenauer
14	2005	Time series properties of a rating system based on financial ratios	U. Krüger, M. Stötzel, S. Trück
15	2005	Inefficient or just different? Effects of heterogeneity on bank efficiency scores	J. Bos, F. Heid, M. Koetter, J. Kolatri, C. Kool
01	2006	Forecasting stock market volatility with macroeconomic variables in real time	J. Döpke, D. Hartmann C. Pierdzioch
02	2006	Finance and growth in a bank-based economy: is it quantity or quality that matters?	Michael Koetter Michael Wedow
03	2006	Measuring business sector concentration by an infection model	Klaus Düllmann
04	2006	Heterogeneity in lending and sectoral growth: evidence from German bank-level data	Claudia M. Buch Andrea Schertler Natalja von Westernhagen
05	2006	Does diversification improve the performance of German banks? Evidence from individual bank loan portfolios	Evelyn Hayden Daniel Porath Natalja von Westernhagen

Visiting researcher at the Deutsche Bundesbank

The Deutsche Bundesbank in Frankfurt is looking for a visiting researcher. Among others under certain conditions visiting researchers have access to a wide range of data in the Bundesbank. They include micro data on firms and banks not available in the public. Visitors should prepare a research project during their stay at the Bundesbank. Candidates must hold a Ph D and be engaged in the field of either macroeconomics and monetary economics, financial markets or international economics. Proposed research projects should be from these fields. The visiting term will be from 3 to 6 months. Salary is commensurate with experience.

Applicants are requested to send a CV, copies of recent papers, letters of reference and a proposal for a research project to:

Deutsche Bundesbank
Personalabteilung
Wilhelm-Epstein-Str. 14

D - 60431 Frankfurt
GERMANY