DEUTSCHE
BUNDESBANK
EUROSYSTEM

Comparing the DSGE model with the factor model: an out-of-sample forecasting experiment

Mu-Chun Wang
(Johann Wolfgang Goethe University of Frankfurt)

Discussion Paper
Series 1: Economic Studies
No 04/2008

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

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

Internet http://www.bundesbank.de

ISBN 978-3–86558–379–6 (Printversion)
ISBN 978-3–86558–380–2 (Internetversion)

Abstract

In this paper, we put DSGE forecasts in competition with factor forecasts. We focus on these two models since they represent nicely the two opposing forecasting philosophies. The DSGE model on the one hand has a strong theoretical economic background; the factor model on the other hand is mainly data-driven. We show that by incorporating large information set using factor analysis can indeed improve the short horizon predictive ability, as claimed by many researchers. The micro-founded DSGE model can provide reasonable forecasts for inflation, especially with growing forecast horizons. To a certain extent, our results are consistent with the prevailing view that simple time series models should be used in short-horizon forecasting and structural models should be used in long-horizon forecasting. Our paper compareds both state-of-the-art data-driven and theory-based modelling in a rigorous manner.

Keywords: DSGE models; factor models; forecasting; forecast evaluation

JEL-Classification: C2 C3, C53, E37

Non-technical summary

Recently, a growing literature has dealt with forecasting using large factor models. Typically, the modern factor models extract only a few components (called common factors) representing the overall economic activities of a large dataset. These factors can then be used to forecast some key economic variables. Various applications of such models in forecasting macroeconomic and financial variables claim that forecasts generated from factor models are superior to simple time series benchmark models. A totally different approach to forecast economic variables is to specify and estimate a dynamic stochastic general equilibrium (DSGE) model. DSGE models have recently become popular in terms of policy analysis because they are very useful storytelling devices due to their rigorous microeconomic foundations. In this paper, we put DSGE forecasts in competition with factor forecasts. We focus on these two models since they represent nicely the two opposing forecasting philosophies. The DSGE model, on the one hand, has a strong theoretical economic background. The factor model, on the other hand, is mainly data-driven. In addition, unrestricted vector autoregressive (VAR) and simple univariate autoregressive (AR) models are considered as benchmark models. We implement out-of-sample forecasting experiments, using the aforementioned models, on output growth and inflation in the US economy on a quarterly basis.

The underlying factor forecasting model is the diffusion indexes model (DI) proposed by Stock and Watson (2002b). The DI model produces direct multistep forecasts without having to specify the dynamics of the factors explicitly. We collect, in total, 83 macroeconomic and financial series and use the static principle component method to extract the factors. The underlying DSGE model is a prototypical monetary New-Keynesian model proposed by Del Negro and Schorfheide (2004). The estimation of the DSGE model is carried out using the Bayesian framework.

Regarding the informativeness of forecasts, model-based output growth forecasts are generally not informative for forecast horizons exceeding one year. However, inflation forecasts computed from the models are vastly informative at every forecast horizon. In terms of the forecasting performance measured in mean squared forecast errors, the factor model is, in general, superior to any other model in the short run. The DSGE model is able to outperform the unrestricted VAR. However, the VAR turns out to be an unreliable benchmark since it is, in most cases, the worst forecasting model.

To a certain extent our results are consistent with the prevailing view that time series models should be used in short-horizon forecasting and structural models should be used in long-horizon forecasting. Our paper closes the existing gap in the literature which up to now has investigated the predictive ability of data-driven factor models and theoretical DSGE models separately from each other.

Nicht technische Zusammenfassung

In den letzten Jahren ist die Zahl der Publikationen, welche sich mit großen Faktormodellen beschäftigen, stark angestiegen. Moderne Faktormodelle ziehen typischerweise aus einem großen Datensatz eine kleine Anzahl von Komponenten (auch gemeinsame Faktoren genannt) heraus, die die allgemeine wirtschaftliche Aktivitäten darstellen. Diese Faktoren können dann zum Vorhersagen von wichtigen ökonomischen Variablen herangezogen werden. Verschiedene Anwendungen von Faktormodellen zur Vorhersage von makroökonomischen und Finanzzeitreihen kommen zu dem Schluss, dass Faktormodelle den einfachen Zeitreihenmodellen überlegen sind.

Ein völlig anderer Ansatz um ökonomischen Variablen zu prognostizieren ist die Spezifikation und Schätzung eines dynamisch-stochastischen Allgemeingleichgewichtsmodells (DSGE). Die DSGE-Modelle sind in aktuellen politischen Analysen populär geworden. Wegen ihrer strengen mikroökonomischen Fundierung bieten die DSGE-Modelle sehr nützliche Interpretations-möglichkeiten. Das vorliegende Papier vergleicht Vorhersagen von einem DSGE-Modell in Konkurrenz zu einem Faktormodell. Wir konzentrieren uns auf diese beiden Modelle, weil sie zwei entgegen gesetzte Vorhersagephilosophien repräsentieren. Das DSGE-Modell auf der einen Seite hat eine solide ökonomisch-theoretische Fundierung; das Faktormodell auf der anderen Seite basiert in erster Linie auf verfügbare Daten. Zusätzlich werden dazu unrestringierte vektorautoregressive Modelle (VAR) und einfache univariate autoregressive Modelle (AR) als Referenzen herangezogen. Wir implementieren Vorhersagesimulationen außerhalb des Schätzzeitraums von den genannten Modellen für US-amerikanischen Wachstumsraten der gesamtwirtschaftlichen Produktion und Inflation auf vierteljährlicher Basis.

Das zugrunde gelegte Faktormodell ist das Diffusionsindex-Modell (DI) von Stock und Watson (2002b). Das DI-Model produziert direkte Mehrschrittvorhersagen ohne die Dynamik der Faktoren explizit spezifizieren zu müssen. Wir sammeln insgesamt 83 Zeitreihen aus den Gebieten der Makroökonomie und der Finanzen und benutzen die statische Hauptkomponentenanalyse um die Faktoren zu extrahieren. Das zugrunde gelegte DSGE-Modell ist ein prototypisches monetäres Neu-Keynesianisches Modell von Del Negro und Schorfheide (2004). Die Schätzung des DSGE-Modells erfolgt Bayesianisch.

Inflationsvorhersagen, die auf dem DSGE-Modell basieren, sind überaus informativ für jeden Prognosehorizont. Dagegen sind die auf diesem Modell basierenden Wachstumsvorhersagen im Allgemeinen nicht informativ für Prognosehorizonte, welche ein Jahr überschreiten. Betrachtet man die Vorhersagegüte, gemessen mit mittleren quadratischen Vorhersagefehlern, so ist das Faktormodell allen anderen Modellen in der Kurzfristprognose überlegen. Weiterhin ist das DSGE-Modell ist in der Lage, das unrestringierte VAR-Modell zu schlagen. Allerdings hat sich das VAR-Modell im Allgemeinen als ein unzuverlässiges Referenzmodell herausgestellt, da es in meisten Fällen das schlechteste Vorhersagemodell ist.

In gewissem Maße bestätigen unsere Resultate die vorherrschende Sicht, dass Zeitreihenmodelle für Kurzfristprognosen und strukturelle Modelle für Langfristprognosen herangezogen werden sollen. Unser Beitrag dient der Schließung der Lücke in der Literatur, welche die Vorhersagegüte von daten-basierten Faktormodellen und theoretischen DSGE-Modellen bislang getrennt voneinander untersuchte.

Contents

List of Tables

List of Figures

Comparing the DSGE Model with the Factor Model: An Out-of-Sample Forecasting Experiment*†

1 Introduction

Various time series approaches have been proposed to forecast economic variables. For a long time, Box-Jenkins univariate methodology has been the workhorse tool for the forecasters and it still frequently works as the benchmark forecasting model. On the multivariate side, the VAR and Bayesian VAR model popularized by Litterman (1986) are also able to provide useful forecasts. However, the multivariate models typically contain only a few variables due to their large number of parameters to be estimated and the resulting high estimation uncertainties.

Recently, a growing number of papers have dealt with forecasting using factor models. Stock and Watson (2002a), Bai and Ng (2002) and Bai (2003) have laid theoretical foundations of estimating large-scale static factor models. Forni, Hallin, Lippi, and Reichlin (2000) and Forni, Hallin, Lippi, and Reichlin (2004) developed the framework of estimating dynamic factor models. The factor models try to incorporate as much economic information as possible and, at the same time, are parsimonious by extracting only a few factors out of a large data set. The various applications of such models in forecasting macroeconomic and financial variables can be found in Stock and Watson (2002b), Marcellino, Stock, and Watson (2003) and Forni, Hallin, Lippi, and Reichlin (2003). They all claim that forecasts generated from factor models are superior to the benchmark models.

Usually, the superiority of factor models, documented in the literature, is justified only relative to simple autoregressive models or small multivariate time series models. In general, it is difficult to establish rigorous economic foundations with forecasting models involving factors. The estimated factors themselves cannot be identified or even interpreted. In a sense, the factor models are among the most data-driven general-to-specific approaches in modelling economic time series.

A totally different modelling philosophy is to use theory-based models to forecast

*I am thankful to Sandra Eickmeier, Uwe Hassler, Michel Juillard, Christian Matthes, Christian Schumacher and seminar participants at the University of Frankfurt, the DIW in Berlin, the IWH in Halle, the Deutsche Bundesbank in Frankfurt and the 13th CEF conference in Montréal for providing helpful comments and discussions. I also wish to thank the Deutsche Bundesbank, where part of this paper was written, for its hospitality and support. The opinions expressed in this paper are those of the author and do not necessarily reflect the views of the Deutsche Bundesbank.

†Correspondence to: Mu-Chun Wang, Statistics and Econometric Methods, Goethe-University Frankfurt, Graefstr. 78 / PF 76, D-60054 Frankfurt, Germany. E-mail: muwang@wiwi.uni-frankfurt.de.

economic variables. However, the structural form equation models during the 70's provided notoriously inaccurate forecasts. Since Kydland and Prescott (1982), economists have begun to specify and estimate micro-founded dynamic stochastic general equilibrium (DSGE) models. Initially not considered as a forecasting tool, the DSGE models have become popular among macroeconomists. A growing number of researchers apply Bayesian methods to estimate the structural parameters of the DSGE models. One of the reasons is the straightforward model comparison within the Bayesian framework.[1] In addition, some researchers argue that the Bayesian approach is superior to the frequentist approach like GMM and maximum likelihood in short samples. Recently, there have been serious attempts to move DSGE models into the field of economic forecasting. Smets and Wouters (2004) estimated medium-scale New Keynesian DSGE models with Bayesian approach and claimed that their DSGE models can generate superior forecasts compared to unrestricted VAR, whereas Del Negro and Schorfheide (2004) and Del Negro, Schorfheide, Smets, and Wouters (2007) use DSGE priors in the Bayesian VAR setup and provide supportive results for the structural models. However, a rigorous comparison of DSGE models with sophiscated time series models has not been yet carried out.

This paper is motivated by the fact that although there are some discussions about both factor and DSGE models in the literature, there has been, to our knowledge, no attempt to carry out a forecasting horserace between these two model classes. The DSGE model, on the one hand, has strong theoretical economic background and is widely considered among theorists and central bankers as a useful model for policy analysis. The factor model, on the other hand, is mainly data-driven and has been the subject of various investigations on its predictive ability. Our out-of-sample forecasting experiments focus on the output growth and price inflation of the US economy. To a certain extent, this paper is closely related to the aforementioned forecasting performance surveys by Stock and Watson (2002b), Forni, Hallin, Lippi, and Reichlin (2003) and Smets and Wouters (2004). However, those papers investigate the predictive ability of the factor and DSGE models isolated from each other. The main contribution of this paper is putting the two state-of-the art macroeconometric models in a rigorous out-of-sample forecasting comparison. Thus, the results might be interesting for both professional forecasters and general macroeconomists.

The rest of the paper is organized as follow: In Section 2 we introduce the forecasting models considered in this paper. In section 3 we present the data, the estimation techniques and the design of the out-of-sample forecasting experiment. In Section 4 we

[1] See Schorfheide (2000)

provide estimation results and discuss the outcomes of the forecasting experiments in detail. Section 5 concludes.

2 Forecasting Models

In this section, the forecasting models for the out-of-sample forecasting experiments are presented. First, the factor model and its forecasting approach are discussed briefly. Then, we provide short descriptions of the DSGE model considered in this paper.

2.1 Factor Model

Since Stock and Watson (2002a) and Stock and Watson (2002b), various forecasting models involving factors have been proposed in the literature. Despite different treatment of the estimation techniques and the dynamics of the underlying factors, it remains inconclusive which factor specification can be regarded as the most accurate in terms of forecasting. Forni, Hallin, Lippi, and Reichlin (2003), Schumacher (2007) found supportive results for the dynamic over the static factor models, whereas Boivin and Ng (2005) report little difference among the competitive factor specifications. In a recent paper, Eickmeier and Ziegler (2007) use a meta-analytic approach to assess the predictive ability of factor models. They conclude that the static factor specification from Stock and Watson (2002b) yields better inflation forecasts than the dynamic approach from Forni, Hallin, Lippi, and Reichlin (2004) and vice versa for output forecasts. Tentatively, it seems that the correct choice of factor specification lies on the data at hand. Our investigation relies on the original factor model specification by Stock and Watson (2002a) and Stock and Watson (2002b) which can be regarded as the benchmark factor model due to its widespread application.

Let X_t denote the observed $N \times 1$ vector of stationary time series and $y_t = \Delta Y_t$ be the first difference representation of the I(1) scalar series Y_t.[2] The underlying factor model is

$$\begin{aligned} X_t &= \Lambda F_t + e_t, \\ y_{t+1} &= \beta(L)F_t + \gamma(L)y_t + \epsilon_{t+1}, \end{aligned} \tag{1}$$

where Λ is the $N \times k$ factor loading matrix and F_t is the $k \times 1$ vector of common factors,

[2]Thoughout this paper, Y_t is either logarithmic output or price series so that the differenced Y_t can be interpreted as the associated growth rates.

e_t is the $N \times 1$ vector of idiosyncratic shocks, $\beta(L)$ and $\gamma(L)$ are lag polynomials of finite order and ϵ_t is the stochastic error term. Stock and Watson (2002a) show in theorem 1, that under mild moment conditions, the factors can be estimated consistently using principle component method.

Equation (1) leads to the so-called direct multistep regression model described in Stock and Watson (2002b):

$$y^h_{t+h} = \alpha_h + \beta_h(L)F_t + \gamma_h(L)y_t + \epsilon^h_{t+h}, \tag{2}$$

where y^h_{t+h} is the horizon-specific growth rate of Y_t. The associated predictors are growth forecasts between t and $t+h$, rather than h-step ahead forecasts of one-period growth rates. The multistep forecasts are generated from the horizon-specific regression by minimizing the h step ahead forecast errors directly, i.e. the underlying model is different in every forecast horizon. Hence, we do not have to specify the potential dynamics of the factors explicitly. This procedure leaves only a small set of unrestricted parameters to estimate compared with fully specified dynamic factor models.

2.2 New Keynesian DSGE Model

For our forecasting experiments we adopt a prototypical New Keynesian DSGE model considered in Del Negro and Schorfheide (2004). The model can be regarded as a benchmark DSGE model which has been investigated extensively in the literature. Del Negro and Schorfheide (2004) for instance investigate the predictive ability of the model in the Bayesian DSGE-VAR framework and conclude that the priors derived from the theoretical DSGE model can help improve forecasts over unrestricted VAR models. In our analysis we do not use the approximated VAR representation of the DSGE model. Instead we work with the fully specified model directly. We hope that our approach by taking the structural restrictions of the DSGE model more seriously set contrast with the pure data-driven factor model for our forecasting experiment.

There is a representative household which maximizes an additive separable utility function in consumption, real money balances and hours worked over infinite lifetime. The household gains utility from consumption relative to the level of technology, real balances of money and disutility from hours worked. The household earns interest from holding government bonds and real profits from the firms. It also receives wages from supplying perfect elastic labour to the firms. In addition the household pays lump-sum taxes to the government.

There is a perfectly competitive, representative final goods producer which uses a continuum of intermediate goods as factor inputs and takes the factor prices as given. The intermediate good producers are monopolists which hire labour as the only factor input. The production technology process is the same for all intermediate good producers and fluctuates around the steady-state growth rate. The nominal rigidities are introduced in terms of price adjustment costs for the intermediate goods producers. Each firm maximizes the profits over infinite lifetime by choosing its labor inputs and prices.

The government spends in each period a fraction of total outputs which fluctuates exogenously. It issues bonds and levies lump-sum taxes as part of the budget constraint. The monetary authority follows the standard Taylor rule, which embodies the inflation target and the output gap with an exogenous monetary policy shock. The following log-linearized equations of the model are taken directly from Del Negro and Schorfheide (2004). All variables with ^ are log-deviations from their steady-state values.

$$\begin{aligned}
\hat{x}_t &= E_t(\hat{x}_{t+1}) - \frac{1}{\tau}[\hat{R}_t - E_t(\hat{\pi}_{t+1})] + (1-\rho_g)\hat{g}_t + \rho_z \frac{1}{\tau}\hat{z}_t \\
\hat{\pi}_t &= \frac{\gamma}{r^*} E_t(\hat{\pi}_{t+1}) + \kappa(\hat{x}_t - \hat{g}_t) \\
\hat{R}_t &= \rho_R \hat{R}_{t-1} + (1-\rho_R)(\psi_1 \hat{\pi}_t + \psi_2 \hat{x}_t) + \epsilon_{R,t} \\
\hat{g}_t &= \rho_g \hat{g}_{t-1} + \epsilon_{g,t} \\
\hat{z}_t &= \rho_z \hat{z}_{t-1} + \epsilon_{z,t},
\end{aligned}$$

where x_t denotes the detrended output (divided by the non-stationary technology process), π_t is the gross inflation rate of prices, P_t. R_t is the gross nominal interest paid by holding government bonds. g_t is defined as $1/(1-\xi_t)$, where ξ_t is the fraction of output consumed by the government. The fluctuations around the technology growth rates are denoted as z_t. Both are assumed to follow an AR(1) process with the associated iid normal idiosyncratic shocks $\epsilon_{g,t}$ and $\epsilon_{z,t}$, The monetary policy shock $\epsilon_{R,t}$ is assumed to be iid normal as well. The standard deviations of the shocks are denoted as σ_R, σ_g and σ_z. The set of deep parameters includes τ, which denotes the inverse of the intertemporal elasticity of substitution, γ is the steady-state growth rate of technology, r^* is the steady-state real interest rate, κ is the slope of the Phillips curve, ρ_R measures the degree of central bank's interest rate smoothing, ψ_1 and ψ_2 are the long-run feedback coefficients from the target values of inflation and output respectively.

The set of measurement equations according to the observable variables quarterly GDP growth rates (YGR), quarterly inflation rates ($INFL$) and annualized nominal interest rates (INT) are defined as

$$\begin{aligned} YGR_t &= \ln\gamma + 100(\hat{x}_t - \hat{x}_{t-1} + \hat{z}_t) \\ INFL_t &= \ln\pi^* + 100\hat{\pi}_t \\ INT_t &= 4\ln\pi^* + 4\ln r^* + 400\hat{R}_t, \end{aligned}$$

where π^* is the steady-state inflation. Let

$$\Phi = [\tau, \kappa, \rho_R, \psi_1, \psi_2, \rho_g, \rho_z, \ln\gamma, \ln\pi^*, \ln r^*, \sigma_R, \sigma_g, \sigma_z]'$$

denote the vector of the model deep parameters to be estimated, and S be the collection of the sample of the observables YGR_t, $INFL_t$ and INT_t. The forecasts generated from the DSGE model are h-step ahead forecasts of quarter-to-quarter output growth and inflation. In order to compare with the forecasts generated from the factor model, we have to accumulate the forecasts of the DSGE model to obtain the horizon-specific growth rate forecasts

3 Data and Estimation

This section deals with the underlying dataset and estimation techniques of the various models introduced in the previous section. First, we discuss the data, sample and experimental design of the out-of-sample forecasting experiment. The rest of the section provides detailed descriptions of estimation procedures of factor models, DSGE models and VAR/AR models.

3.1 Data and Experimental Design

The goal of this research paper is to carry out simulated out-of-sample forecasting experiments for the output growth and inflation of the US economy at a quarterly frequency. The dataset for all models spans from 1959:1 to 2006:3. For the factor estimation, we obtain 83 US economic time series measuring overall economic activity on a quarterly basis from Moody's Economy and Datastream Advance. We argue that the dataset is, admittedly, not identical with Stock and Watson (2002b), but that it is rather similar as the data selection and treatment closely follow the data description in Stock and Watson (2002b) and uses the same series where applicable.[3]

[3] Detailed descriptions of the dataset are given in the appendix.

For the estimation of the DSGE model we use quarterly data obtained from BEA to calculate output growth (YGR) as 100 times the logarithmic differences of GDP (GDP-SAAR, Billions Chained (2000) Dollars). Inflation ($INFL$) is represented by 100 times the logarithmic differences of CPI (CPI-U, All Items, SA, 1982-1984=100, source: BLS). The nominal interest rates (INT) are represented by the average effective federal funds rate during the first month of each quarter. (Source: FRED).

For the out-of-sample forecasting experiments, we use rolling regression with sample size fixed at 80. Forecast horizons up to 12 quarters are considered. The first estimation sample starts from 1960:1 and ends in 1979:4 so that the first forecasting date is 1980:1. Earlier observations are used to compute the initial growth rates. After all models have been estimated, the first set of out-of-sample forecasts is computed. Then, sample range shifts one-step forward to 1960:2-1980:1 in order to compute the second set of forecasts. The estimation is performed 96 times to obtain a series of forecasts for each forecast horizon and each model. The last sample is 1983:4-2003:3 and the last forecasting date is 2006:3.

3.2 Data Treatment and Estimation

3.2.1 Factor Model (Diffusion Index)

In terms of factor estimation, we follow closely the pre-estimation data treatment described in Stock and Watson (2002b). First, all series in X_t which are not negative or already expressed as rates or percentage units are taken logarithm. Then, the series are differenced according to preliminary data analysis. All implemented data transformations are documented in the Appendix. Finally, all series are demeaned and standardized to have unit variance. The estimation of the forecasting equation (2) involves estimated factors, which are computed using the principle component method. The forecasts are then generated as

$$\hat{y}^h_{T+h|T} = \hat{a}_h + \sum_{j=1}^{m} \hat{\beta}_{hj} \hat{F}_{T-j+1} + \sum_{i=1}^{p} \hat{\gamma}_{hi} y_{T-i+1}, \tag{3}$$

where $\hat{y}^h_{T+h|T}$ denotes the h-step ahead growth forecasts at time T defined earlier, and $\hat{F}_t$ denotes the vector of k estimated factors. m, p, and k are either fixed or selected with BIC with $1 \leq m \leq 4$, $0 \leq p \leq 4$, and $1 \leq k \leq 6$, which is very similar to the specifications of Stock and Watson (2002b) where they refer (3) to the diffusion index forecasts (DI). Stock and Watson (2002a) show, in theorem 2, that, under mild regularity conditions, (3) can be estimated consistently with OLS. For the out-of-sample forecasting experiments, y_t is represented by YGR and $INFL$ and y_t^h is calculated as 100 times the logarithmic h-th differences of the associated series in level (GDP and CPI).[4]

3.2.2 New Keynesian Model

We estimate the New Keynesian DSGE model using a Bayesian approach. Since the model has been log-linearized, it can be written in the state space representation. The system is solved under rational expectations. The support of the prior distributions has been restricted such that the uniqueness of the solution is guaranteed, then the Kalman filter is applied on the systems to evaluate the likelihood function. The posterior densities estimation of the parameters is carried out with the Markov Chain Monte Carlo algorithm (MCMC).[5] We proceed with a slightly modified version of the Random Walk Metropolis-Hastings algorithm (RWMH) documented in Schorfheide (2000). Let $l(\Phi|S)$ denote the conditional log-likelihood function of the deep parameters and $\ln p(\Phi)$ denote the log of prior density.

1. Find the posterior mode $\tilde{\Phi}$ and Hessian $\tilde{\Sigma}$ of Φ by numerical maximizing the $l(\Phi|S) + \ln p(\Phi)$.

2. For J parallel chains, draw the starting value Φ_j^0 from $N(\tilde{\Phi}, c_j\tilde{\Sigma}^{-1})$, where c_j denotes the scale parameter of the jumping distribution for each chain j.

3. At step i, take a candidate draw Φ_j^* from $N(\Phi_j^{i-1}, c_j\tilde{\Sigma}^{-1})$.

4. Calculate the acceptance probability $\alpha_j^i = \min\Big(1, p(\Phi_j^*|S)/p(\Phi_j^{i-1}|S)\Big)$, where $p(.|S)$ denotes the posterior density.

5. Accept the candidate draw by setting $\Phi_j^i = \Phi_j^*$ with probability α_j^i. Otherwise, set $\Phi_j^i = \Phi_j^{i-1}$.

6. Calculate the average acceptance ratio $\bar{\alpha}_j$. Adjust c_j every q-th iteration according to $c_j = c_j\frac{\bar{\alpha}_j}{\alpha^*}$, where α^* denotes the target average acceptance ratio. Do not adjust after the replications exceed the threshold I.

7. Go to step 3 until maximum iteration $\bar{I}$ is reached.

[4] Notice that, implicitly, we treat prices as I(1) series. This is given by the fact that our DSGE model assumes inflation to be stationary.

[5] The estimation and forecasting of the DSGE models are performed with modified version of Dynare. Part of the algorithm presented in this paper is documented in Dynare.

The algorithm is different from Schorfheide (2000) in two respects. First, we use J independent chains in order to compute the convergence diagnostic plots proposed by Brooks and Gelman (1998). Second, we adjust the scale parameter automatically within the RWMH algorithm since all models have to be re-estimated several times according to the out-of-sample forecasting experiments. Finding the appropriate scale parameter manually for the jumping distribution for each rolling sample is too tedious and a time-consuming task. Notice that our algorithm does not violate any regularity conditions which guarantee the convergence of the Markov chains since we only use the draws with constant jumping distribution, i.e. the draws below threshold I with moving jumping distributions are discarded.

Posterior inference is based on combinations of the remaining draws of all parallel chains. To be precise, we set $J = 2$, $I = 25000$, $\bar{I} = 50000$, $q = 500$ and $\alpha^* = 0.3$. The mean of the posterior forecast distributions is taken as the point forecast of the relevant variable. The MCMC diagnostics proposed by Brooks and Gelman (1998) show that all Markov chains for each estimation sample have converged nicely.

3.2.3 Benchmarks: VAR and AR Model

Since the solution of the DSGE model can be approximated with a restricted VAR model, it is natural to estimate a corresponding unrestricted VAR as a counterpart in the out-of-sample forecasting experiment. According to our DSGE model, the unrestricted VAR contains the same observed variables: YGR, $INFL$ and INT. We report results for both the fixed-lag specification and the information criterion-based lag selection (BIC). The horizon specific growth rate forecasts are constructed in the same way as in the case of the DSGE model.

The benchmark AR model follows the specification of the direct multistep regression model:

$$y_{t+h}^h = \alpha_h + \gamma_h(L)y_t + \epsilon_{t+h}. \tag{4}$$

The model has the same dynamic specification as our factor forecasting model in (2). The only difference is the lack of factors in (4). The lag length selection is carried out with BIC.

4 Forecasting Results

Section 4 presents the findings of this paper. First, summary results of the DSGE models estimation are provided. The remainder of section 4 discusses in detail the outcomes of the main out-of-sample forecasting experiments of output growth and inflation.

4.1 Estimation Results of the DSGE model

Table 1 reports the priors and means of the posterior mean of the deep parameters with the associated 90% highest posterior density interval (HPDI) for the first and last estimation samples. The first sample ranged from 1960:1 to 1979:4 and can be regarded as the pre-Volcker era, or "Great Inflation". The last sample 1983:4-2003:3 encompasses the Volcker-Greenspan era, or the "Great Moderation". The priors are very similar to those from Del Negro and Schorfheide (2004) with the means centred on various calibration exercises and micro-estimates.[6]

The estimated moments of the posteriors are mostly in line with estimates reported in the literature. Interestingly, most of the parameter estimates differ substantially between the two samples and reflect the change of the economy from "Great Inflation" to "Great Moderation". The steady-state growth rate of technology rises from 0.50 for the pre-Volcker period to 0.60 for the Volcker-Greenspan sample. The steady-state inflation rate during the Volcker-Greenspan period (0.86) is only moderately lower than in the pre-Volcker period (0.89). In terms of monetary policy, the estimated responses of the central bank to deviations of output and inflation to their targets have hardly changed. However, the estimate of the interest rate smoothing parameter is substantially higher during the Volcker-Greenspan period (from 0.64 to 0.85). These results indicate that the US central bank may have shifted toward a more gradual monetary policy.

The other major differences between the two sub-periods are the estimated volatilities of the structural shocks. Compared with the pre-Volcker period, the standard deviations of all structural shocks have decreased considerably in the latter period. The standard deviation of monetary shock drops from 0.26 to 0.17, and the standard deviation of the technology shock nearly halved from 0.89 to 0.49. The estimated standard deviation of the government spending shock has also decreased (from 0.44 to 0.40). The slope of the Phillips curve has become flatter in the second period. Usually, it is a hint that the prices have become stickier. Nevertheless, the slope coefficient κ is a nonlinear combination of both the price stickiness and the elasticity of intermediate goods demand, which cannot be identified separately. Therefore, one should be wary of this interpretation.

[6] We have investigated the sensitivity of our results by increasing the prior standard deviations of the DSGE model by 50 percent. The overall results remain very similar.

4.2 Informativeness of Forecasts

In order to access informativeness of forecasts generated from the models, we report in Table 2 the relative mean squared forecast errors defined as

$$rMSFE(m) = 1 - \frac{MSFE_m}{MSFE_{\bar{y}}},$$

where $MSFE_m$ denotes the mean squared forecast errors generated from the particular forecast model, and $MSFE_{\bar{y}}$ denotes the mean squared forecast errors with the in-sample mean of the series as the forecast. The values can be interpreted as gain (or loss if the sign is negative) in $MSFEs$ of the conditional forecasts based on the particular model relative to the unconditional mean. If a model forecast is considered as informative, its $rMSFE$ should be larger than zero.[7]

Figure 1 shows the $rMSFEs$ of our out-of-sample forecasting experiments.[8] The upper panel shows the results of output growth forecasts. The DI model is able to provide substantially informative forecasts up to two-years-ahead horizon. At the one-quarter-ahead horizon, the DI model yields 28 percent lower mean squared forecast errors than unconditional forecasts. At the one-year-ahead horizon the improvement is even higher at 32 percent. Overall the DI model is superior to any other models considered in this paper. Compared to the simple AR model, the improvement of the DI model is 18 percent at the one-quarter-ahead horizon. With regard to the DSGE model, the improvement is even higher at 35 percent. The DSGE and the VAR model are not able to provide informative forecasts for output growth at all. This is not really surprising given that the output growth is a particularly difficult series to predict.[9] Nevertheless, the results indicate that the prevailing approach of treating the unrestricted VAR model as the benchmark is somewhat problematic, since the unrestricted VAR model itself is outperformed by simple unconditional mean in terms of output growth forecasts.

The lower panel of figure 1 shows the results for inflation forecasts. Now, all models except the VAR model can clearly generate informative forecasts for all forecast horizons. For both the DI and the DSGE model, the improvements in mean squared forecast errors are over 60 percent up to one-year-ahead horizon. The DI model also has a better performance than the DSGE model within this range. At the one-quarter-ahead horizon, the DI model yields 29 percent lower mean squared forecast errors than the DSGE model.

[7] Brisson, Campbell, and Galbraith (2003) also refer to the relative mean squared forecast errors as the forecast content function.

[8] Detailed tables of the $rMSFEs$ can be provided upon request.

[9] See Brisson, Campbell, and Galbraith (2003).

On the other hand, the DSGE model has, overall, the best predictive ability beyond the one-year-ahead horizon. It is able to outperform the DI model by 23 percent at the six-quarter-ahead horizon. The AR model performs somewhere in between. It is worse than the DI model and better than the DSGE model up to five-quarters-ahead horizon and the reverse is the case afterwards. In sharp contrast to the output growth forecasts, the VAR model is able to generate considerably informative forecasts up to two-years-ahead horizon. However, it remains the worst forecasting model for all forecast horizons.

4.3 Various Factor and VAR Specifications

In this subsection we provide cross-check with the results of cited literature, namely Stock and Watson (2002b) for the factor model where the AR model is considered as the benchmark and Del Negro and Schorfheide (2004), where the performance of the DSGE-VAR model is documented using the unrestricted VAR model as the benchmark. Hence, we focus on the number of estimated factors used in the DI model and the lag length of the VAR model. Figure 2 shows the gain in $MSFE$ of a different number of estimated factors over the simple AR model, where BIC denotes the automatic factor selection with Bayesian information criterion, and k denotes the fixed number of factors specification of the DI model. In general, the results are very similar to the findings of Stock and Watson (2002b). In terms of forecasting output growth, only a few factors are needed to outperform the benchmark AR model. There are considerable gains in the predictive ability of the DI model if one uses more than one factor. The improvement over the AR model varies mostly around 20% for all forecast horizons.

The results of inflation forecasts are quite different from the output growth forecasts. Now, using one factor is the only case where the DI model is able to beat the AR model for all forecast horizons. A Higher number of factors, including BIC based selection, generate inferior forecasts to the AR model beyond the five-quarters-ahead horizon. Interestingly, there seems to be no advantage of using BIC to automatically determine the number of factors for both output growth and inflation forecasts.

Figure 3 shows the gains of the DSGE model over different lag length specification of the unrestricted VAR model, where BIC denotes automatic lag length selection with Bayesian information criterion and p denotes fixed lag length specification of the VAR. Again, the results are in line with the associated results from Del Negro and Schorfheide (2004), where a VAR(4) is used as the benchmark. Clearly, the DSGE model is able to outperform the VAR(4) decisively in forecasts of both output growth and inflation. There are only a few occasions where the VAR model can generate superior forecasts. The

VAR(1) yields 10 percent and 6 percent lower $MSFE$ for ouput growth at one-quarter-ahead and two-quarter-ahead horizon respectively. The VAR(3) has at its best (at two-years-ahead horizon for output growth) 6 percent lower $MSFE$ than the DSGE model. No VAR model specification can compete with the DSGE model in forecasting inflation for all horizons. Overall, the predictive ability of the VAR model tends to deteriorate quickly relative to the DSGE model as the forecasting horizon grows. However, as we documented in the previous subsection, the VAR model is an unreliable benchmark in terms of informativeness. We argue that the supportive results for the DSGE models in the literature may be due partly to the fact that the underlying benchmark models (VAR) are easy to outperform.

4.4 Testing equal predictive ability

We show that exploiting large datasets with the factor framework can indeed lead to more accurate forecasts than is possible with simple time series models. Also, it seems that the DSGE model is not able to compete with the factor model except for long horizon inflation forecasts. However, there is a general caveat of evaluating forecast accuracy based on error metrics suchb as $MSFE$: it is the lack of statistical significance. It is not possible to make rigorous statistical statements by simply ignoring the sampling uncertainties of such metrics. To trace the systematic differences between competing forecast models, we need statistical hypothesis tests.

Diebold and Mariano (1995) and West (1996) discuss some significance tests of forecast accuracy in the literature and present their own asymptotic test of equal forecast accuracy, which we refer to the Diebold-Mariano-West test (DMW test). Let P be the total number of forecasts, and consider two forecast error series u_{1p} and u_{2p}. The null hypothesis of equal accuracy of two competitive forecasts is

$$H_0 : E\left[g(u_{1p}) - g(u_{2p})\right] = E(d_p) = 0.$$

where $g(.)$ denotes a loss function and d_p denotes the loss differential series. The large sample $N(0,1)$ test statistics under the null of equal accuracy is

$$S_{DMW} = \frac{\bar{d}}{\sqrt{\frac{2\pi \hat{f}_d(0)}{P}}},$$

where $\bar{d}$ denotes the sample mean loss differential and $f_d(0)$ denotes the spectral density of the loss differential at frequency zero

$$f_d(0) = \frac{1}{2\pi} \sum_{\tau=-\infty}^{\infty} \gamma_d(\tau),$$

where $\gamma_d(\tau)$ are the autocovariances of the loss differential at displacement τ. In practice, $f_d(0)$ can be consistently estimated with the HAC estimator proposed by Newey and West (1987)

$$2\pi \hat{f}_d(0) = \hat{\gamma}_d(0) + 2 \sum_{\tau=1}^{B} \left(1 - \frac{\tau}{B}\right) \hat{\gamma}_d(\tau),$$

where B denotes the truncation lag and $\hat{\gamma}_d(\tau)$ denote thes sample estimates of the autocovariances. In our paper, we consider quadratic loss function and $B = h - 1$. [10]

The test statistic requires the limiting variance of the loss differentials to be finite and strictly positive. With nested models, this is unfortunately not the case when both the estimation sample R and the number of forecasts P goes to infinity, because, in this case, the forecast errors are, under the null hypothesis, asymptotically the same and, therefore, perfectly correlated. Clark and McCracken (2001) show that, when comparing two nested forecasting models, the DMW test statistics have a non-standard limiting distribution. However, in Giacomini and White (2006), theorem 4, it is demonstrated that, under the rolling window regression and taking account of the estimation uncertainties, the limiting distribution is still standard normal. In addition, Giacomini and White (2006) argue that by focusing on the estimated rather than the population values of the model parameters, their framework can accommodate Bayesian estimation methods. Hence, we argue that the proposed test statistic is suitable for our forecasting experiments.

Table 2 presents the results of the pairwise unconditional equal predictive ability tests. Plus (minus) signs of the test statistics indicate that the forecasting model in rows have lower (higher) squared forecast errors in mean than the corresponding forecasting model in columns. Strikingly, looking at output growth forecasts, there are only very few occasions showing significant differences between the competing models. The test reveals that only the AR model at one-quarter-ahead horizon is able to outperform the naive mean significantly. The factor model produces significantly better forecasts than the DSGE and the VAR model also at one-quarter-ahead horizon. From two-years-ahead

[10] We also tried different spectral density estimator like quadratic spectral kernel and automatic truncation lag selection proposed by Andrews (1991). Overall, the test outcomes remain very similar.

horizon onwards, no significant differences can be found between the competing models.

In terms of the inflation forecasts, the results are quite different. Now, all models are able to generate significant informative forecasts up to one-year-ahead horizon. Especially the DI model dominates all competing models at one-quarter-ahead horizon. Interestingly, the DSGE model tends to outperform the competing models as the forecasting horizon grows. However, on no occasion is the DSGE model significantly superior to the DI model. Overall, the unconditional predictive ability tests report few differences across the competing forecasting models considered in this paper. The DI model seems to be superior in short-horizon output growth and inflation forecasts. The DSGE model is arguably more accurate in long-horizon inflation forecasts. Nevertheless, it is documented that the DMW test has low power in the finite sample, particularly when nested models are involved.[11] Hence, the results of the test should be regarded as complementary rather than definitive.

5 Conclusion

This paper conducts out-of-sample forecasting horse race between data-driven and theory-based forecasting models, with large factor models and small-scale New Keynesian DSGE models serving as the prominent state-of-the-art candidate models. The results of factor models confirm findings from various surveys claiming that the factor models are able to provide reasonable short-run forecasts for both output growth and inflation. However, the choice of the number of factors clearly matters for the predictive ability of factor models, although how the "correct" number of factors should be selected is unclear, since the automatic information criterion-based procedure does not uniformly outperform the manually fixed procedure.

The estimation results of the small-scale New Keynesian DSGE model are quite standard compared with the related empirical literature. Some key notions of the economic evolution can be captured with the rolling sample regressions of the models. The most surprising result is the strong predictive ability of the DSGE model in forecasting inflation in the long run. We show that, by taking the economic meaningful restrictions imposed on the DSGE models seriously, one can still obtain reasonable forecasts. The unrestricted VAR model, as the common benchmark in assessing the predictive power of DSGE models, turns out to be very weak. It is not able to provide informative forecasts for most cases. Therefore we suggest future out-of-sample forecasting experiments involving DSGE models to include more reasonable benchmarks.

[11] See Clark and McCracken (2001).

Given the wide variety of DSGE models in the literature, this paper should not be understood as an exhaustive research into the predictive ability of DSGE models. Nevertheless, the findings of the forecasting experiments are surely encouraging for further research on this topic, since the DSGE model is a very useful storytelling device for both theorists and policymakers. It is important that such a device exhibits reasonable predictive ability. Future research on topics should try to dissect the impact of different structural restrictions on the predictive ability of the DSGE models in order to develop a deeper understanding of the underlying structural economy.

Appendix

This appendix gives an overview of the dataset used to construct the factors. The data are presented in the following ordering: series number, series mnemonic, series description and transformation code. The transformation codes are 1 = no transformation, 2 = first difference, 3 = first difference of logs, 4 = second difference of logs. All price series are obtained from Moody's Economy and all other series are obtained from Datastream. The series mnemonics and descriptions are taken directly from the associated sources. The interest rate spreads are calculated using the average federal funds rate obtained from Moody's Economy. The abbreviations appearing in the series descriptions are sa/sadj = seasonally adjusted, cura = current prices, seasonally adjusted, vola = volumn index, seasonally adjusted.

	Prices		
1	cpiuaa_us	cpi: urban consumer - apparel, (1982-84=100, sa)	4
2	cpiuac_us	cpi: urban consumer - commodities, (1982-84=100, sa)	4
3	cpiuad_us	cpi: urban consumer - durables, (1982-84=100, sa)	4
4	cpiuam_us	cpi: urban consumer - medical care, (1982-84=100, sa)	4
5	cpiuas_us	cpi: urban consumer - services, (1982-84=100, sa)	4
6	cpiuat_us	cpi: urban consumer - transportation, (1982-84=100, sa)	4
7	cpiul1_us	cpi: urban consumer - all items less food, (1982-84=100, sa)	4
8	cpiul2_us	cpi: urban consumer - all items less shelter, (1982-84=100, sa)	4
9	cpiul5_us	cpi: urban consumer - all items less medical care, (1982-84=100, sa)	4
10	ppisp1000_us	ppi: stage of processing - crude materials, (index 1982=100, sa)	4
11	ppisp2000_us	ppi: stage of processing - intermediate materials, (index 1982=100, sa)	4
12	ppisp3000_us	ppi: stage of processing - finished goods, (index 1982=100, sa)	4
13	ppisp3100_us	ppi: stage of processing - finished consumer goods, (index 1982=100, sa)	4
	Consumption		
14	uscdtan_b	pce - durables, new autos (ar) cura	3
15	uscondurb	personal consumption expenditures - durables (ar) cura	3
16	usconndrb	personal consumption expenditures - nondurables (ar) cura	3
17	usconsrvb	personal consumption expenditures - services (ar) cura	3
18	usperconb	personal consumption expenditures (ar) cura	3

	Employment		
19	usem21__o	employed - mining vola	3
20	usem23__o	employed - construction vola	3
21	usem42__o	employed - wholesale trade vola	3
22	usem81__o	employed - otherservices vola	3
23	usemig__o	employed - government vola	3
24	usemimd_o	employed - durable goods vola	3
25	usemip__o	employed - totalprivate vola	3
26	usemir__o	employed - retail trade vola	3
27	usemit__o	employed - trade, transportation, & utilities vola	3
28	usempallo	employed - nonfarm industries total (payroll survey) vola	3
29	usempg__o	employed - goods-producing vola	3
30	usempmano	employed - manufacturing vola	3
31	usemps__o	employed - service-providing vola	3
32	usemptoto	total civilian employment vola	3
33	ushlpwadq	help wanted-proportion of labor markets w/rising want-ad vola	1
34	usun_totq	unemployment rate sadj	2
35	usundurne	average durationof unemployment (weeks) vola	1
36	usunw14_q	unemployed distribution - 5 to 14 weeks sadj	1
37	usunw15_q	unemployed distribution - 15 weeks & over sadj	1
38	usunw26_q	unemployed distribution - 15 to 26 weeks & over sadj	1
39	usunw5__q	unemployed distribution - less than 5 weeks sadj	1
40	usvactoto	index of help wanted advertising vola	3
	Housing		
41	ushbrm__o	housing started - midwest (ar) vola	3
42	ushbrn__o	housing started - northeast (ar) vola	3
43	ushbrs__o	housing started - south (ar) vola	3
44	ushbrw__o	housing started - west (ar) vola	3
45	ushouse_o	new private housing units started (ar) vola	3
	Hours and earnings		
46	ushkim__o	avg wkly hours -manufacturing vola	3
47	ushxpmano	avg overtime hours - manufacturing vola	3
48	uswr23__b	avg hrly earn -construction cura	4
49	uswrim__b	avg hrly earn -manufacturing cura	4

	<u>Output and income</u>		
50	usipmbuqg	indl prod - business equipment vola	3
51	usipmcogg	indl prod - consumer goods vola	3
52	usipmducg	indl prod - durable consumer goods vola	3
53	usipmfgsg	industrial production - manufacturing (sic) vola	3
54	usipmfing	indl prod - final products, total vola	3
55	usipmmatg	indl prod - materials, total vola	3
56	usipmnocg	indl prod - nondurable consumer goods vola	3
57	usipmprog	indl prod - final products & nonindustrial supplies vola	3
58	usiptot_g	industrial production - total index vola	3
59	usiumfgsq	indl utilization- manufacturing (sic) sadj	1
60	uspdispib	disposable personal income (ar) cura	3
61	uspersinb	personal income (ar) cura	3
	<u>Interest rates</u>		
62	uscrbbaa	corporate bond yield - moody's baa, seasoned issues	2
63	uscrbyld	corporate bond yield - moody's aaa, seasoned issues	2
64	ustrb3av	treasury bill secondary market rate on discount basis-3 month	2
65	ustrcn10	treasury yield adjusted to constant maturity - 10 year	2
66	ustrcn1_	treasury yield adjusted to constant maturity - 1 year	2
67	ustrcn5_	treasury yield adjusted to constant maturity - 5 year	2
68	usytb6sm	treasury bill secondary market rate on discount basis-6 month	2
69	ussfycrbyld	spread uscrbyld - federal funds	1
70	ussfycrbbaa	spread uscrbbaa - federal funds	1
71	ussfytrb3av	spread ustrb3av - federal funds	1
72	ussfyytb6sm	spread usytb6sm - federal funds	1
73	ussfytrcn1_	spread ustrcn1_ - federal funds	1
74	ussfytrcn10	spread ustrcn10 - federal funds	1
75	ussfytrcn5_	spread ustrcn5_ - federal funds	1
	<u>Other</u>		
76	usm0____b	monetary base cura	4
77	usnbrrsab	nonborrowed reserves of depository institutions cura	3
78	uspmchin	chicago purchasingmanager diffusion index-inventories(sa)	1
79	uspmchlt	chicago purchasingmanager diffusion index-deliveries(sa)	1
80	uspmchp_	chicago purchasingmanager diffusion index-prodn. (sa) sadj	1
81	ustotrsab	total reserves of depository institutions cura	3
82	usexpgdsb	exports f.a.s. cura	3
83	uscnfbusq	ism purchasing managers index (mfg survey) sadj	1

References

Andrews, D. (1991): "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

Bai, J. (2003): "Inferential Theory For Factor Models of Large Dimension," *Econometrica*, 71, 135–171.

Bai, J., and S. Ng (2002): "Determine the Number of Factors in Approximate Factor Models," *Econometrica*, 70, 191–221.

Boivin, J., and S. Ng (2005): "Understanding and Comparing Factor Based Macroeconomic Forecast," *International Journal of Central Banking*, 1, 117–152.

Brisson, M., B. Campbell, and J. W. Galbraith (2003): "Forecasting Some Low-Predictability Time Series Using Diffusion Indices," *Journal of Forecasting*, 22, 515–531.

Brooks, S., and A. Gelman (1998): "General Methods for Monitoring Convergence of Iterative Simulations," *Journal of Computational and Graphical Statistics*, 7, 434–455.

Clark, T., and M. McCracken (2001): "Tests of Equal Forecast Accuracy and Encompassing for Nested Models," *Journal of Econometrics*, 105, 85–110.

Del Negro, M., and F. Schorfheide (2004): "Priors from General Equilibrium Models for VARs," *International Economic Review*, 45, 643–673.

Del Negro, M., F. Schorfheide, F. Smets, and R. Wouters (2007): "On the Fit and Forecasting Performance of New Keynesian Models," *Journal of Business and Economic Statistics*, 25, 123–143.

Diebold, F. X., and R. S. Mariano (1995): "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, 253–263.

Eickmeier, S., and C. Ziegler (2007): "How Successful are Dynamic Factor Models at Forecasting Output and Inflation? A Meta-Analytic Approach," forthcoming in Journal of Forecasting.

Forni, M., M. Hallin, M. Lippi, and L. Reichlin (2000): "The Generalized Factor Model: Identification and Estimation," *The Review of Economics and Statistics*, 82, 540–554.

——— (2003): "Do Financial Variables Help Forecasting Inflation and Real Activity in the EURO Area?," *Journal of Monetary Economics*, 50, 1243–1255.

——— (2004): "The Generalized Dynamic Factor Model: Consistency and Rates," *Journal of Econometrics*, 119, 231–255.

Giacomini, R., and H. White (2006): "Tests of Conditional Predictive Ability," *Econometrica*, 74, 1545–1578.

Kydland, F., and E. Prescott (1982): "Time to Build and Aggregate Fluctuations," *Econometrica*, 50, 1345–1370.

Litterman, R. (1986): "Forecasting with Bayesian Vector Autoregressions - 5 Years of Experience," *Journal of Business and Economic Statistics*, 4, 25–38.

Marcellino, M., J. Stock, and M. Watson (2003): "Macroeconomic Forecasting in the Euro Area: Country-Specific versus Area-Wide Information," *European Economic Review*, 47, 1–18.

Newey, W., and K. West (1987): "A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

Schorfheide, F. (2000): "Loss Function-Based Evaluation of DSGE Models," *Journal of Applied Econometrics*, 15, 645–670.

Schumacher, C. (2007): "Forecasting German GDP Using Alternative Factor Models Based on Large Datasets," *Journal of Forecasting*, 26, 271–302.

Smets, F., and R. Wouters (2004): "Forecasting with a Bayesian DSGE Model: An Application to the Euro Area," *Journal of Common Market Studies*, 42, 841–867.

Stock, J., and M. Watson (2002a): "Forecasting Using Principal Components from a Large Number of Predictors," *Journal of the American Statistical Association*, 97, 1167–1179.

——— (2002b): "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business and Economic Statistics*, 20, 147–162.

West, K. (1996): "Asymptotic Inference About Predictive Ability," *Econometrica*, 64, 1067–1084.

Table 1: Priors and posterior distributions of the DSGE model

	Priors		Pre-Volcker		Volcker-Greenspan	
Parameters	Distribution	*Mean* (*Std*)	Posterior Mean	HPDI	Posterior Mean	HPDI
$\ln\gamma$	Normal	0.50 (0.25)	0.50	[0.34; 0.65]	0.60	[0.47; 0.71]
$\ln\pi^*$	Gamma	1.00 (0.50)	0.89	[0.32; 1.39]	0.86	[0.29; 1.39]
$\ln r^*$	Gamma	0.50 (0.25)	0.35	[0.21; 0.49]	0.62	[0.50; 0.74]
τ	Gamma	2.00 (0.50)	2.73	[1.93; 3.54]	2.92	[2.12; 3.66]
κ	Gamma	0.30 (0.15)	0.23	[0.07; 0.39]	0.18	[0.07; 0.28]
ψ_1	Gamma	1.50 (0.25)	1.34	[1.05; 1.64]	1.31	[1.04; 1.54]
ψ_2	Gamma	0.125 (0.10)	0.28	[0.12; 0.44]	0.31	[0.13; 0.51]
ρ_R	Beta	0.50 (0.20)	0.64	[0.54; 0.74]	0.85	[0.80; 0.89]
ρ_g	Beta	0.80 (0.10)	0.93	[0.89; 0.97]	0.98	[0.96; 0.99]
ρ_z	Beta	0.30 (0.10)	0.34	[0.18; 0.50]	0.39	[0.22; 0.54]
$100\sigma_R$	Inv. Gamma	0.25 (0.14)	0.26	[0.22; 0.29]	0.17	[0.14; 0.19]
$100\sigma_g$	Inv. Gamma	0.63 (0.32)	0.44	[0.29; 0.59]	0.40	[0.26; 0.54]
$100\sigma_z$	Inv. Gamma	0.88 (0.43)	0.89	[0.72; 1.07]	0.49	[0.40; 0.58]

Note: Std denotes the prior standard deviation.

Table 2: Test of equal predictive ability

Output growth					Inflation				
	h=1					h=1			
	Mean	AR	DI	VAR		Mean	AR	DI	VAR
AR	1.9255*				AR	3.9263**			
DI	1.5973	0.9307			DI	4.4500**	1.7828*		
VAR	0.0717	-0.5121	-1.7564*		VAR	3.0615**	-1.6074	-2.1113**	
DSGE	-0.9024	-2.6596**	-1.7920*	-0.4876	DSGE	3.7594**	-0.6776	-1.7600*	1.4286
	h=4					h=4			
AR	1.3480				AR	4.4020**			
DI	1.3946	1.0625			DI	4.2876**	0.7651		
VAR	-0.1501	-0.5066	-1.2801		VAR	2.8762**	-2.0583**	-1.7374*	
DSGE	-0.3808	-1.7051*	-1.5758	0.0272	DSGE	3.5416**	-0.5707	-0.9553	1.5682
	h=8					h=8			
AR	0.8712				AR	2.7239**			
DI	-0.0348	-0.1469			DI	2.0294**	-0.5322		
VAR	-0.1056	-0.2922	-0.0382		VAR	0.4962	-1.4578	-0.6770	
DSGE	-0.0696	-0.1934	-0.0413	0.0041	DSGE	2.4887**	1.0221	0.7389	1.6901*
	h=12					h=12			
AR	0.2276				AR	1.6331			
DI	0.0733	-0.0001			DI	0.9783	-0.1480		
VAR	-1.2491	-1.1778	-1.1886		VAR	-0.2035	-1.0909	-0.6896	
DSGE	-0.7834	-0.7779	-0.6828	0.5709	DSGE	1.9470*	1.8237*	0.7566	1.6608*

Note: The entries are test statistics of pairwise tests of equal predictive ability. An asterisk denotes rejection of the null of equal predictive ability at 10%. Double asterisks denote rejection at 5%. For example the first entry means, at one-quarter-ahead horizon, the AR model outperforms the unconditional mean forecasts for output growth. The associated test statistic is 1.9255 and hence significant at 10%

Figure 1: Gain (Loss) in MSFEs: models versus unconditional mean

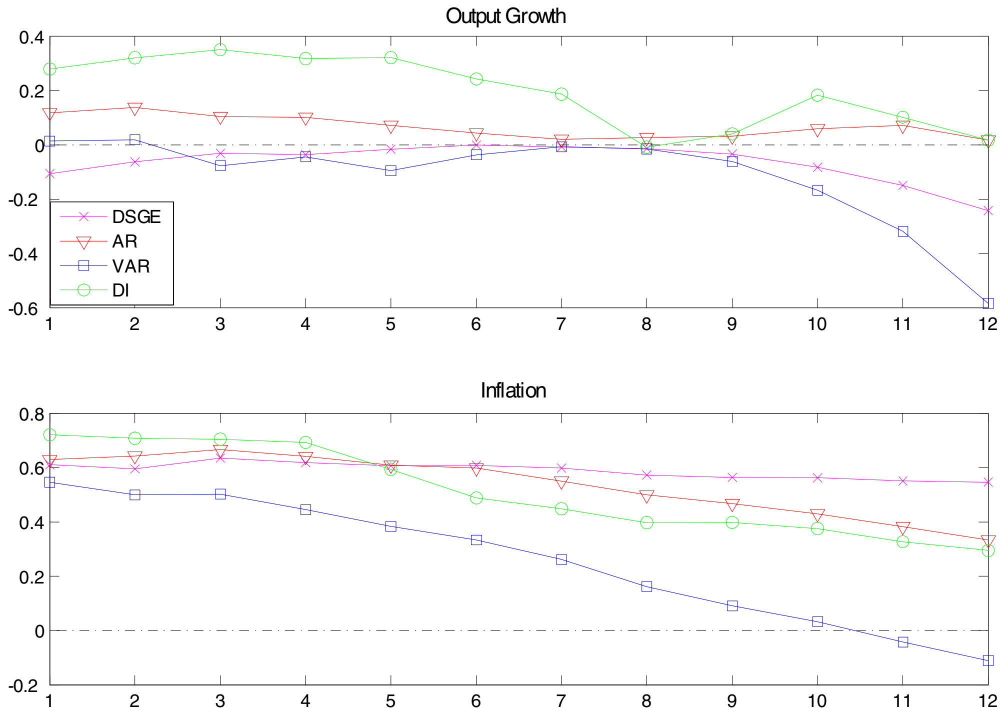


Figure 2: Gain (Loss) in MSFEs: DI versus AR

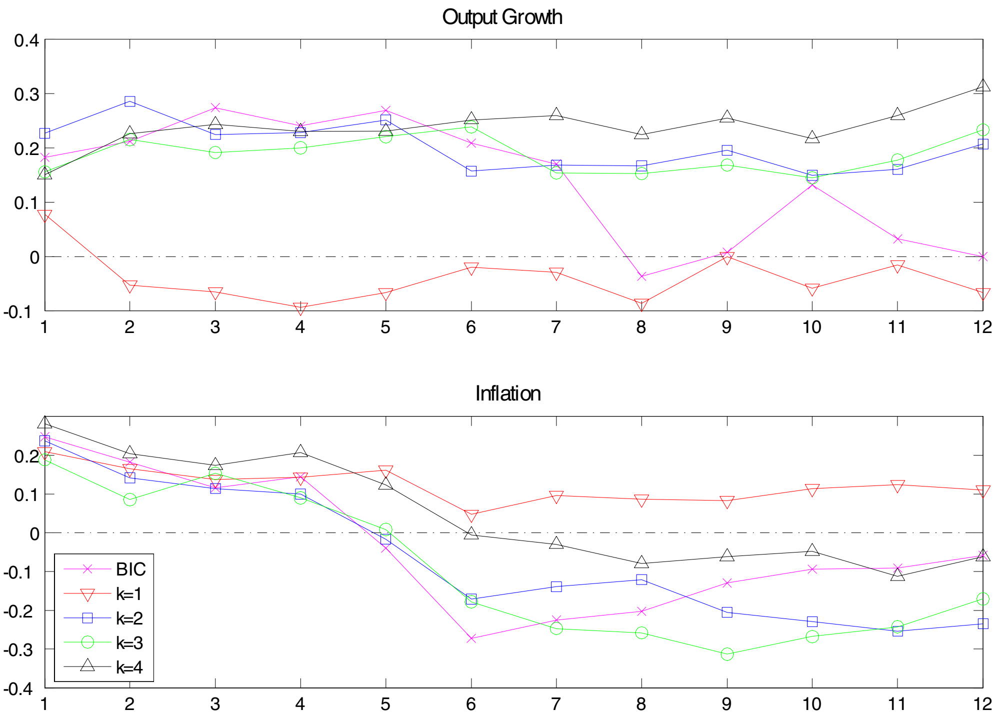
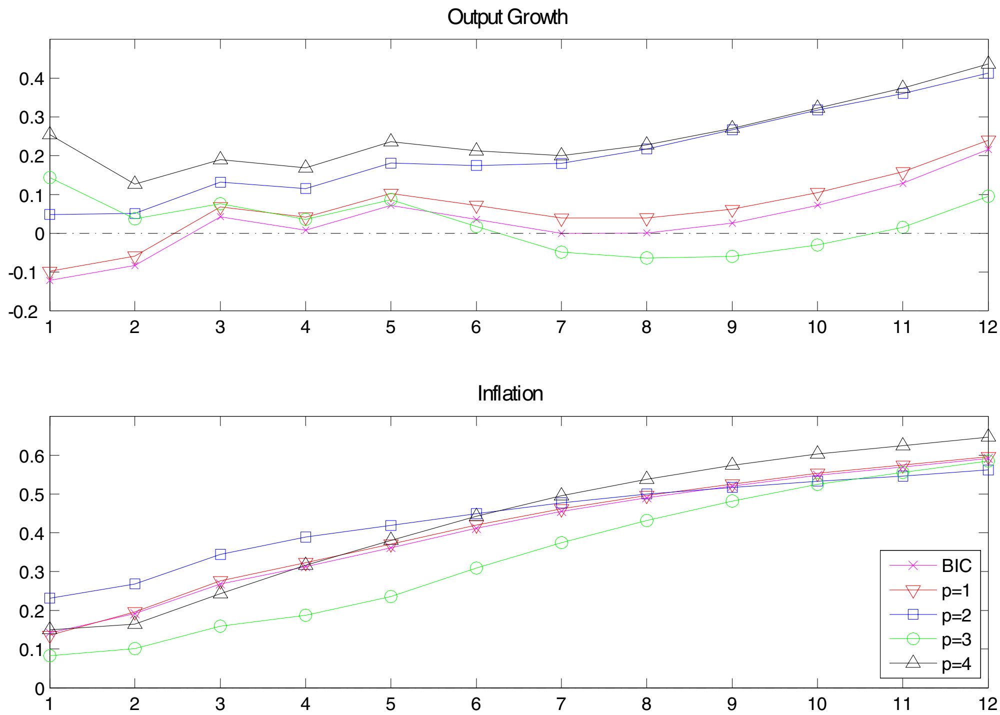


Figure 3: Gain (Loss) in MSFEs: DSGE versus VAR



The following Discussion Papers have been published since 2007:

Series 1: Economic Studies

01	2007	The effect of FDI on job separation	Sascha O. Becker Marc-Andreas Mündler
02	2007	Threshold dynamics of short-term interest rates: empirical evidence and implications for the term structure	Theofanis Archontakis Wolfgang Lemke
03	2007	Price setting in the euro area: some stylised facts from individual producer price data	Dias, Dossche, Gautier Hernando, Sabbatini Stahl, Vermeulen
04	2007	Unemployment and employment protection in a unionized economy with search frictions	Nikolai Stähler
05	2007	End-user order flow and exchange rate dynamics	S. Reitz, M. A. Schmidt M. P. Taylor
06	2007	Money-based interest rate rules: lessons from German data	C. Gerberding F. Seitz, A. Worms
07	2007	Moral hazard and bail-out in fiscal federations: evidence for the German Länder	Kirsten H. Heppke-Falk Guntram B. Wolff
08	2007	An assessment of the trends in international price competitiveness among EMU countries	Christoph Fischer
09	2007	Reconsidering the role of monetary indicators for euro area inflation from a Bayesian perspective using group inclusion probabilities	Michael Scharnagl Christian Schumacher
10	2007	A note on the coefficient of determination in regression models with infinite-variance variables	Jeong-Ryeol Kurz-Kim Mico Loretan

11	2007	Exchange rate dynamics in a target zone - a heterogeneous expectations approach	Christian Bauer Paul De Grauwe, Stefan Reitz
12	2007	Money and housing - evidence for the euro area and the US	Claus Greiber Ralph Setzer
13	2007	An affine macro-finance term structure model for the euro area	Wolfgang Lemke
14	2007	Does anticipation of government spending matter? Evidence from an expectation augmented VAR	Jörn Tenhofen Guntram B. Wolff
15	2007	On-the-job search and the cyclical dynamics of the labor market	Michael Krause Thomas Lubik
16	2007	Heterogeneous expectations, learning and European inflation dynamics	Anke Weber
17	2007	Does intra-firm bargaining matter for business cycle dynamics?	Michael Krause Thomas Lubik
18	2007	Uncertainty about perceived inflation target and monetary policy	Kosuke Aoki Takeshi Kimura
19	2007	The rationality and reliability of expectations reported by British households: micro evidence from the British household panel survey	James Mitchell Martin Weale
20	2007	Money in monetary policy design under uncertainty: the Two-Pillar Phillips Curve versus ECB-style cross-checking	Günter W. Beck Volker Wieland
21	2007	Corporate marginal tax rate, tax loss carryforwards and investment functions – empirical analysis using a large German panel data set	Fred Ramb

22	2007	Volatile multinationals? Evidence from the labor demand of German firms	Claudia M. Buch Alexander Lipponer
23	2007	International investment positions and exchange rate dynamics: a dynamic panel analysis	Michael Binder Christian J. Offermanns
24	2007	Testing for contemporary fiscal policy discretion with real time data	Ulf von Kalckreuth Guntram B. Wolff
25	2007	Quantifying risk and uncertainty in macroeconomic forecasts	Malte Knüppel Karl-Heinz Tödter
26	2007	Taxing deficits to restrain government spending and foster capital accumulation	Nikolai Stähler
27	2007	Spill-over effects of monetary policy – a progress report on interest rate convergence in Europe	Michael Flad
28	2007	The timing and magnitude of exchange rate overshooting	Hoffmann Sondergaard, Westelius
29	2007	The timeless perspective vs. discretion: theory and monetary policy implications for an open economy	Alfred V. Guender
30	2007	International cooperation on innovation: empirical evidence for German and Portuguese firms	Pedro Faria Tobias Schmidt
31	2007	Simple interest rate rules with a role for money	M. Scharnagl C. Gerberding, F. Seitz
32	2007	Does Benford's law hold in economic research and forecasting?	Stefan Günnel Karl-Heinz Tödter
33	2007	The welfare effects of inflation: a cost-benefit perspective	Karl-Heinz Tödter Bernhard Manzke

34	2007	Factor-MIDAS for now- and forecasting with ragged-edge data: a model comparison for German GDP	Massimiliano Marcellino Christian Schumacher
35	2007	Monetary policy and core inflation	Michele Lenza
01	2008	Can capacity constraints explain asymmetries of the business cycle?	Malte Knüppel
02	2008	Communication, decision-making and the optimal degree of transparency of monetary policy committees	Anke Weber
03	2008	The impact of thin-capitalization rules on multinationals' financing and investment decisions	Buettner, Overesch Schreiber, Wamser
04	2008	Comparing the DSGE model with the factor model: an out-of-sample forecasting experiment	Mu-Chun Wang

Series 2: Banking and Financial Studies

01	2007	Granularity adjustment for Basel II	Michael B. Gordy Eva Lütkebohmert
02	2007	Efficient, profitable and safe banking: an oxymoron? Evidence from a panel VAR approach	Michael Koetter Daniel Porath
03	2007	Slippery slopes of stress: ordered failure events in German banking	Thomas Kick Michael Koetter
04	2007	Open-end real estate funds in Germany – genesis and crisis	C. E. Bannier F. Fecht, M. Tyrell
05	2007	Diversification and the banks' risk-return-characteristics – evidence from loan portfolios of German banks	A. Behr, A. Kamp C. Memmel, A. Pfingsten
06	2007	How do banks adjust their capital ratios? Evidence from Germany	Christoph Memmel Peter Raupach
07	2007	Modelling dynamic portfolio risk using risk drivers of elliptical processes	Rafael Schmidt Christian Schmieder
08	2007	Time-varying contributions by the corporate bond and CDS markets to credit risk price discovery	Niko Dötz
09	2007	Banking consolidation and small business finance – empirical evidence for Germany	K. Marsch, C. Schmieder K. Forster-van Aerssen
10	2007	The quality of banking and regional growth	Hasan, Koetter, Wedow
11	2007	Welfare effects of financial integration	Fecht, Grüner, Hartmann
12	2007	The marketability of bank assets and managerial rents: implications for financial stability	Falko Fecht Wolf Wagner

13	2007	Asset correlations and credit portfolio risk – an empirical analysis	K. Düllmann, M. Scheicher C. Schmieder
14	2007	Relationship lending – empirical evidence for Germany	C. Memmel C. Schmieder, I. Stein
15	2007	Creditor concentration: an empirical investigation	S. Ongena, G.Tümer-Alkan N. von Westernhagen
16	2007	Endogenous credit derivatives and bank behaviour	Thilo Pausch
17	2007	Profitability of Western European banking systems: panel evidence on structural and cyclical determinants	Rainer Beckmann
18	2007	Estimating probabilities of default with support vector machines	W. K. Härdle R. A. Moro, D. Schäfer

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

60431 Frankfurt
GERMANY