

## WORKING PAPERS

14 | 2011

### RATIONAL VS. PROFESSIONAL FORECASTS

João Valle e Azevedo

João Tovar Jalles

May 2011

---

The analyses, opinions and findings of these papers represent the views of the authors, they are not necessarily those of the Banco de Portugal or the Eurosystem

---

Please address correspondence to  
João Valle e Azevedo  
Banco de Portugal, Economics and Research Department  
Av. Almirante Reis 71, 1150-012 Lisboa, Portugal;  
Tel.: 351 21 313 0163, email: [jvazevedo@bportugal.pt](mailto:jvazevedo@bportugal.pt)



*Banco de Portugal*  
EUROSYSTEM

**BANCO DE PORTUGAL**

Av. Almirante Reis, 71

1150-012 Lisboa

*www.bportugal.pt*

***Edition***

Economics and Research Department

***Pre-press and Distribution***

Administrative Services Department

Documentation, Editing and Museum Division

Editing and Publishing Unit

***Printing***

Administrative Services Department

Logistics Division

Lisbon, May 2011

***Number of copies***

150

ISBN 978-989-678-083-8

ISSN 0870-0117 (print)

ISSN 2182-0422 (online)

Legal Deposit no. 3664/83

# Rational vs. Professional Forecasts<sup>\*</sup>

João Valle e Azevedo<sup>†</sup>

João Tovar Jalles<sup>‡</sup>

Banco de Portugal

University of Cambridge

Universidade NOVA de Lisboa

May 2, 2011

## Abstract

We compare theoretical and empirical forecasts computed by rational agents living in a model economy to those produced by professional forecasters. We focus on the variance of the prediction errors as a function of the forecast horizon and analyze the speed with which it converges to a constant (which can be seen as a measure of the speed of convergence of the economy to the steady state). We look at a standard sticky-prices-wages model, concluding that it delivers a strong theoretical forecastability of the variables under scrutiny, at odds with the data (professional forecasts). The flexible prices-wages version delivers a forecastability closer to the data and performs relatively better empirically (with actual data), but mainly because forecasts deviate little from the unconditional mean. These results can be interpreted in at least two ways: first, actual deviations from the steady-state

---

<sup>\*</sup>Acknowledgements: The authors are grateful to Valerio Ercolani, António Antunes, Nikolay Iskrev, Pedro Teles, Isabel Correia, Bernardino Adão, Andrew C. Harvey, Petra Geraats, Mikhail Pranovich, Eva Ortega, Gabriel Perez-Quiros, James Costain, Samuel Hurtado, Juan Francisco Jimeno, Javier Ortega and Michael Bengad for insightful comments and suggestions as well as seminar participants at the University of Cambridge, Banco de Portugal, Banco de España, University of Birmingham, University of Liverpool and City University London. The usual disclaimer applies. Most research was done while João Tovar Jalles was a Visiting Researcher at the Banco de Portugal, whose hospitality is greatly appreciated. The views expressed are those of the authors and do not necessarily represent those of the Banco de Portugal or the Eurosystem.

<sup>†</sup>Corresponding author: Banco de Portugal, Research Department, Av. Almirante Reis, 71-6th floor, 1150-012 Lisboa, Portugal. E-mail: azevedojv@gmail.com

<sup>‡</sup>University of Cambridge, Faculty of Economics, Sidgwick Avenue, CB3 9DD, Cambridge UK. E-mail: jodstj2@cam.ac.uk

are not persistent, in which case the implications of the specific formulation of nominal rigidities for short-run dynamics are unrealistic; second, and still looking through the lens of a model, exogenous (or unmodelled) steady-state shifts attributable to, e.g., changes in monetary-policy, taxation, regulation or in the growth of the technological frontier, occur in such a way as to strongly limit the performance of professional forecasters.

*JEL Classification:* C32, C53, C68, E12, E13,E17

*Keywords:* Survey of Professional Forecasters, Green Book forecasts, DSGE models, Nominal rigidities

# 1 Introduction

Despite tremendous efforts over the past decades, macroeconomic forecasting seems as difficult as before. For most variables forecast accuracy is low, simple models prove hard to beat and sophisticated statistical methods provide marginal (if any) improvements at long horizons. This degree of uncertainty should perhaps be considered a feature of the economy, as the same difficulty characterizes professional forecasts (say, from the Philadelphia Survey of Professional Forecasters, henceforth Phil-SPF, the Federal Reserve Green Book, Fed Green-Book, or the European Central Bank Survey of Professional Forecasters, ECB-SPF). Still, there is clear evidence they rank very well in comparison with various statistical methods, being less prone to the structural instabilities of macroeconomic time series (e.g., Faust and Wright, 2007 and Bernanke and Boivin, 2003 for evidence on Green-Book forecasts or Ang, Bekaert and Wei, 2007 for Phil-SPF inflation forecasts). Moreover, professional forecasts can be seen as a fortunate aggregation of various individual forecasts that probably adapt fast to changes in the economy, each using different data, different methods and even some judgment. The question we address is whether the behavior of these forecasts shares characteristics of theoretical and empirical forecasts produced by typical dynamic stochastic general equilibrium (DSGE) models. We will call the first set of forecasts “Professional” and the latter “Rational”.

We view Professional forecasts as the best publicly available proxy for the forecasts produced by well-informed agents in the economy, providing a natural benchmark against which to confront the forecasts produced by rational agents living in a model economy that is taken seriously. Departing from the common practice of comparing moments and impulse response functions, we assess the fit of models to data by analyzing the differences in the relative performance of Rational vs. Professional forecasts, paying special attention to the variance of the forecast errors as a function of the forecast horizon as well as to the speed with which it converges to a constant (the speed with which the forecasts converge to the unconditional mean of the variables under scrutiny). Under the strong assumption that unmodelled features of the economy affecting the steady-state are unimportant for the analysis of short-run dynamics, we

can view this exercise as useful to understand in which dimensions (or for which variables) theoretical models provide a reasonable description of the process (or speed) of convergence of the economy to the steady-state. We will interpret small differences in forecast accuracy (as a function of the forecast horizon) in the two worlds (Rational and Professional) as sign that the model is able to replicate an important dimension of actual data.

Our benchmark model economy will be the one discussed in Smets and Wouters (2007) in its New-Keynesian (NK) and Real Business Cycle (RBC) versions (i.e., with and without nominal rigidities). The comparison of the two paradigms is instructive. In fact, nominal rigidities (along with real features adding persistence such as habit formation and investment adjustment costs) are often incorporated with the justification that they enable the models to replicate the persistence of the response to various shocks identified with VAR's, i.e., impulse response functions that are still "alive" after two or even three years horizons. This translates into theoretical forecastability of the corresponding variables at very long horizons. However, we conclude that for the growth of real variables such as consumption, investment or output, Professional forecasters don't do better than the mean at horizons greater than 3 or 4 quarters. The notable exception occurs with the unemployment rate (which we take as a proxy for hours worked in the theoretical model). For nominal interest rates both forecasters and agents that live in a sticky-prices-wages economy can form forecasts that are superior to the mean at very long horizons (in practice, simple time series models also can do that due to the very high persistence of nominal rates). For inflation, forecasters still add to the mean after 5 quarters, but little, whereas the standard sticky-price-wage model (under standard parametrizations) is still far from the mean of the process at long horizons, clearly at odds with the data. The RBC version is silent with respect to nominal interest rates and inflation but delivers forecasts of real variables that are closer (in terms of relative performance with respect to the unconditional mean and speed of convergence) to Professional forecasts. The notable exception relates to hours. Once we use these models to forecast actual data, the performance is extremely weak but less so with the RBC version (mainly because forecasts converge to the unconditional mean more rapidly). Again, the exception to this pattern is found with hours/unemployment.

The remainder of the paper is organized as follows: Section 2 shows that, for a host of variables, the predictive power of Professional forecasts vanishes fast as the forecast horizon increases, i.e., the gains (if any) that one obtains by using these forecasts instead of a (real-time) estimate of the unconditional mean of the variables are small. Section 3 confronts these facts with the theoretical and empirical performance of a standard DSGE model, whereas section 4 pays special attention to what occurs during recessions. Section 5 concludes.

## 2 Professional Forecasts: how much they deliver

Here we assess the predictive power of U.S. Professional forecasts, measuring simply their performance relative to an estimate of the unconditional mean of the variables analyzed.<sup>1</sup> In this way we investigate how informative these forecasts are and until when (in terms of forecast horizon) they provide relevant signal relative to what can be viewed as a steady-state forecast.

### 2.1 Data

We analyze 15 macroeconomic indicators from the Phil-SPF,<sup>2</sup> namely: Nominal GNP/GDP (NOUTPUT), Real GNP/GDP (ROUTPUT), Industrial Production Index - Total (IPT), Real Personal Consumption Expenditures - Total (RCON), Net Exports (NETEXP), GDP deflator (GDPDEF), Consumer Price Index (CPI), Real Gross Private Domestic Investment – Residential (RINVRESID), Real Gross Private Domestic Investment – Nonresidential (RINVBF), Real Government Consumption and Gross Investment – State and Local (RGLS), Real Government Consumption and Gross Investment – Federal (RGF), Housing Starts (HSTARTS), unemployment (UNRATE), T-bill (TB3MS), T-bond (GS10). We look only at point forecasts and define these as the median forecast in every release of the survey (results with the mean

---

<sup>1</sup>Analysis of the forecast performance of Phil-SPF is routinely conducted at the Federal Reserve Bank of Philadelphia, see Stark (2010) for a recent discussion.

<sup>2</sup>For complete information on the survey's background see <http://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/spf-documentation.pdf> as well as Zarnowitz (1969), Zarnowitz and Braun (1993) and Croushore (1993).

forecast are very similar). The individual respondent's point forecasts are generally close to the central tendencies of their subjective distributions (e.g., Engelberg et al. 2009) while there is clear evidence that this aggregation produces forecasts that are in general superior to individual forecasts. Obviously, a not so straightforward aggregation can result in forecast improvements, and this can be achieved even when there is (as in Phil-SPF) entry and exit of forecasters, see Capistrán and Timmerman (2009).

Regarding the Fed Green-Book, we analyze predictions for real GDP (ROUTPUT) and GDP deflator (GDPDEF) only, and only those constructed with approximately the same information set as Phil-SPF's (i.e., presented at the FOMC meeting closest to the middle of the second month of a given quarter). The sample periods are as follows: Phil-SPF, 1981q3-2009q2; Fed Green-Book, 1981q3-2003q4.<sup>3</sup> All data is firstly aggregated quarterly when necessary (to be consistent with the variables forecasted in the Phil-SPF) and except for unemployment and interest rates, all data is in growth rates. For a more detailed description of data treatment and transformations please refer to Appendix A.

## 2.2 Methodology

We begin our analysis by taking the real-time vintage data from 1968q4 through 1981q3  $-h$  quarters, for  $h = 1, \dots, 5$ .<sup>4</sup> We then estimate the unconditional mean of the variables under scrutiny by simply computing the average of each variable for this vintage, which is our benchmark forecast for 1981q3. As a reference, we also compute forecasts from an estimated Direct Autoregression (AR) using data through 1981q3  $-h$ . We repeat this procedure using the vintage from 1968q4 through 1981q4  $-h$ ,  $h = 1, \dots, 5$  and so forth until 2009q2. It should be noted that most variables are available with a delay of one quarter. Hence, to properly compare these benchmark forecasts with Phil-SPF and the Fed-Green Book forecasts we re-label the forecast horizon so that the information sets with each method approximately coincide (so, the  $h$  step

---

<sup>3</sup>The Fed Green-Book only releases forecasts every five years and 2003 is the latest available year with data.

<sup>4</sup>These series were retrieved from the Philadelphia Fed website. <http://www.phil.frb.org/research-and-data/real-time-center/real-time-data/data-files/>. See, e.g., Croushore and Stark (2001, 2003) for a discussion of real-time data.

ahead, with  $0 \leq h \leq 4$ , Phil-SPF and Green-Book forecasts will be considered as  $h + 1$  step ahead forecasts since the latest observation of the variable to be forecasted does not in general refer to the forecast moment,<sup>5</sup> which is approximately the middle of the quarter since at least 1990q3).<sup>6</sup>

We then compare the forecast accuracy of the Professional predictions, AR and also real-time average by computing the ratio of Root Mean Square Forecast Error (RMSFE) of both the AR and Professional forecasts relative to that of the benchmark forecast (real-time average). It should be noted that the forecast error is defined as the difference between the forecast and the corresponding observation of the **latest** vintage of data available (considering the  $h$  quarters ahead vintage alters little the results, at least in relative terms, see also Stark 2010 for a thorough analysis of this issue). Following Fair and Shiller (1989) we also run the following forecast encompassing regression:

$$y_{t+h} = \alpha + \beta_0 f_{t+h}^{\text{real}} + \beta_1 f_{t+h}^x + \varepsilon_{t+h} \quad (1)$$

where  $y_{t+h}$  is the observation of the variable forecasted,  $f^{\text{real}}$  is the forecast from the real-time average,  $f^x$  is the forecast from the candidate model  $x$ , in our case either the AR process or the Professional forecasts and  $\varepsilon_{t+h}$  is a (most likely serially correlated) regression error. Obviously, if  $\beta_1 \neq 0$ , then forecasts from candidate model  $x$  add information relative to the real-time average. Both the RMSFE ratios and the  $\beta_1$  coefficients are computed and presented for the full sample as well as for an aggregation of recession quarters as defined by the NBER (results for each decade, that is, the 1980s, 1990s and the period 2000q1-2009q2 are available upon request).

Additionally, we will take a closer look at Phil-SPF and Fed-Green Book forecasts for real GDP (ROUTPUT) and the GDP deflator (GDPDEF), to understand if Phil-SPF forecasts represent best practice, or close to best practice, within professional forecasts. In fact, using

---

<sup>5</sup>This does not apply, e.g., to interest rates, whose quarterly average is to be forecasted but are obviously available in the middle of the quarter, when the forecast is made.

<sup>6</sup>The timing of the previous American Statistical Association/NBER survey (that was taken by the Philadelphia FED) is not known exactly but it is believed that it followed closely Phil-SPF's.

data until 1991, Romer and Romer (2000) show that the Fed-Green Book forecasts of inflation and real GDP are statistically unbiased and dominate private sector forecasts (i.e., suggesting that the Federal Reserve has considerably more information beyond what is known to the private sector). The period of “Great Moderation”<sup>7</sup> between 1982-2007 has affected the time-series properties of many variables as well as the forecasting performance of different models. In particular, D’Agostino and Whelan (2008) show that the superior forecasting performance of the Fed-Green Book forecasts deteriorated considerably after 1991. Similarly, Gamber and Smith (2008) find that the Fed’s relative forecasting superiority has declined relative to Phil-SPF forecasts for both inflation and real GDP growth after 1994. Thus, and in line with Chong and Hendry (1986) and Granger and Newbold (1986), we will conduct efficiency-type encompassing tests based on the estimation of the following equations:

$$ef_{t+h}^x = \gamma_0 + \gamma_1(ef_{t+h}^B - ef_{t+h}^x) + \varepsilon_{t+h} \quad (2)$$

$$ef_{t+h}^B = \delta_0 + \delta_1(ef_{t+h}^x - ef_{t+h}^B) + \varepsilon_{t+h} \quad (3)$$

where  $ef_{t+h}^x (= y_{t+h} - f_{t+h}^x)$  is a general representation for the alternative  $x$  forecast error (Fed-Green Book or Phil-SPF) whereas  $ef_{t+h}^B (= y_{t+h} - f_{t+h}^B)$  is the benchmark forecast error (respectively Phil-SPF or Fed-Green Book). Under the null hypothesis that  $\gamma_1 = 0 (\delta_1 = 0)$ ,  $f_{t+h}^x(f_{t+h}^B)$  encompasses  $f_{t+h}^B(f_{t+h}^x)$ .

It is important to note that a forecast performing as well as an estimate of the unconditional mean in terms of RMSFE (or encompassed by it) may nonetheless be useful if it can more often accurately predict the direction of change in the actual series (Joutz and Stekler, 2000). With this in mind we will examine the sign forecast accuracy of the forecasts by constructing the following two-by-two contingency table in which the actual and forecast data for each quarter are classified (i) by whether the actual change in a given variable is positive (+) or negative/zero (-,0), and (ii) whether the forecast correctly or incorrectly predicted the sign:

---

<sup>7</sup>See, e.g., McConnell and Perez-Quiros (2000), Stock and Watson (2003) and Giannone, Lenza and Reichlin (2008)

Contingency Table	
$n_{11} : \Delta y_{t+h}(+), \Delta f_{t+h}^x(+)$	$n_{12} : \Delta y_{t+h}(-, 0), \Delta f_{t+h}^x(+)$
$n_{21} : \Delta y_{t+h}(+), \Delta f_{t+h}^x(-, 0)$	$n_{22} : \Delta y_{t+h}(-, 0), \Delta f_{t+h}^x(-, 0)$

where the actual change is  $\Delta y_{t+h} = y_{t+h} - y_t$  and the predicted change is  $\Delta f_{t+h}^x = f_{t+h}^x - y_t$ . Note that  $y_t$  is the most recent (quarterly) value known at the time of the forecast. The diagonal cells include numbers of correct sign forecasts and the off-diagonal cells include the numbers of incorrect sign forecasts. In testing the null hypothesis of no association between the frequency of actual and predicted changes we use Fisher's exact test.

## 2.3 Results and Discussion

### 2.3.1 Forecast Accuracy

Our main empirical results regarding forecast accuracy are presented in Tables 1.1 and 1.2 for the Phil-SPF and the Fed-Green Book, respectively.<sup>8</sup> Table 1.1 presents the forecast performance with respect to the 15 macroeconomic variables defined before. It contains the ratio of the Root Mean Square Forecast Error (RMSFE) of both the AR and Professional forecasts relative to the benchmark forecast (real-time average) as well as the estimate of  $\beta_1$  resulting from OLS estimation of Eq. (1) at different forecast horizons (robust standard errors are reported below the estimated coefficient).<sup>9</sup> The main conclusions from Table 1.1 follow:

- considering the full sample, Phil-SPF forecasts add signal relative to the benchmark (real-time average) only up to  $h = 2$  when looking at the significance of the  $\beta_1$  coefficients for most variables under scrutiny. The exceptions are Phil-SPF's CPI inflation, GDP deflator, unemployment and interest rates (both T-bond and T-bill rates) predictions throughout

<sup>8</sup>For reasons of parsimony, these are condensed tables that focus solely on the full sample and recession periods that result from simple aggregation of all periods corresponding to the 4 recessions identified by the NBER. Versions of these tables with sub-samples are available from the authors upon request.

<sup>9</sup>It is advisable to use Newey-West standard errors since the error terms are generally serially correlated and heteroskedastic. For further details see Harding and Pagan (2006). Other estimators could have been used instead, e.g., the method outlined in Kiefer and Vogelsang (2002) and Phillips et al. (2003).

the different forecast horizons and Phil-SPF's RGLS (State and Local Government Consumption and Gross investment growth) up to  $h = 4$ .

- considering the full sample, the relative (to the real-time average) RMSFE for Phil-SPF's forecasts is clearly less than one for all horizons only in the case of unemployment, interest rates and, in a lesser extent, inflation (CPI and GDP deflator). In the case of 10 year bond interest rates the AR outperforms Phil-SPF whereas for the 3-month T-bill rate the opposite is true. For output (nominal and real) and specially industrial production, housing starts and net exports this ratio indicates mostly useless Phil-SPF forecasts at horizons greater than or equal to  $h = 3$ . For consumption, investment (residential and non-residential) and Government expenditures (federal and local) there is still some superiority on average (relative to the real-time average) at horizons  $h = 3, 4, 5$ . In these cases however, it would in general suffice to use a simple autoregression as the rel. RMSFE compares favorably with Phil-SPF's.<sup>10</sup>
- for all variables except (again) interest rates, inflation (CPI and GDP deflator) and unemployment, Phil-SPF (and AR) forecasts that correspond to recession periods have a quite poor performance relative to the real-time average except at  $h = 1$  (for many variables). Afterwards the rel. RMSFEs are higher than those obtained with the full sample and more frequently well above 1. This evidence is in line with e.g. Zarnowitz (1992) , Zarnowitz and Braun (1992), McNees (1992) and McNees and Ries (1983) who reported a number of systematic errors made by forecasters regarding recession periods. For  $h = 2$ ,  $\beta_1$  is nonetheless still significant for Phil-SPF in the case of consumption and for AR forecasts in the cases of state and local government expenditures, non-residential private investment and industrial production, despite the fact that rel. RMSFE is above 1 in two of the three cases.

---

<sup>10</sup>One should bear in mind that regressions including only recession periods contain solely 28 observations and this fact may be the underlying cause of some marginally significant coefficient estimates for longer forecast horizons in some variables (e.g. nominal and real GDP or residential investment).

We now focus on Table 1.2 which is organized similarly to Table 1.1 but refers to Fed-Green Book’s predictions instead, computed using the same information set as Phil-SPF forecasts (i.e., we focus on the four quarterly forecasts produced before the FOMC meetings occurring in the middle of a given quarter). In this case the sample size is shorter, going up to 2003q4. We look only at Fed-Green Book’s forecasts of real output and GDP deflator inflation, the ones available from the Federal Reserve Bank of Philadelphia. For real output, the overall performance is worse than the Phil-SPF counterpart in terms of rel. RMSFE,<sup>11</sup> in accordance with Gavin and Mandal (2003) whose evaluation of real output growth forecasts finds no evidence of the Fed’s superior information about the economy. As for GDP deflator-based inflation, we find relative RMSFEs smaller than unity together with highly significant  $\beta_1$  coefficients at all forecast horizons for the full sample. We will later compare in more detail Green book and Phil-SPF forecasts for these variables, to see if Phil-SPF forecasts can be considered best practice among well-informed agents in the economy.

Putting it simply, this exercise shows that for most variables a real-time estimate of the conditional mean seems a hard to beat forecast (even) at short horizons. Regarding unemployment, nominal interest rates and inflation, Professional forecasts do contain relevant information beyond that of our crude benchmark forecast. In these cases, however, it is more clear that the distance between these forecasts and the real-time average forecast is surely overstated, in the sense that the latter is supposed to measure a steady-state value that may be varying over time (e.g., due to changes in monetary policy or labor market institutions). This is not damaging for our purposes as it allows us to refer to this distance as an upper bound on what a theoretical model (without regime shifts in monetary policy or labor market institutions) ought to deliver in terms of forecast accuracy relative to the steady state forecast.

### 2.3.2 Sign forecast accuracy

Table 2.1 reports the cell counts for the contingency table described in section 2.2 and p-values for the null of no association between actual and predicted changes for the Phil-SPF, Real-

---

<sup>11</sup>Constraining Phil-SPF’s sample to the one available for the Fed-Green Book does not alter this conclusion.

time average and Fed-Green Book forecasts. First, it is clear that most p-values for Phil-SPF and Green-book forecasts are less than 0.1, or the null hypothesis of no association between actual and predicted changes is rejected, indicating that, in general, these forecasts accurately predict the direction of change in the actual series. What is more interesting for our purposes is to compare the behavior of Professional's forecasts to that of the benchmark (real-time average) forecast. In fact, it is perfectly possible that this rough benchmark forecast accurately predicts the direction of change at long horizons. First, we observe that for  $h = 1, 2$  Phil-SPF and Green Book forecasts are, in general, clearly more informative than the real-time average (lower associated p-values), according to this criterion. Second, for real output Green-book (not Phil-SPF) forecasts clearly look more useful than the benchmark even at  $h = 3, 4, 5$ .<sup>12</sup> But the main result emerging from Table 2.1 is that at horizons greater than  $h = 2$  and for all other variables except interest rates, CPI inflation, unemployment and to a lesser extent State and local Government spending, the null of no association (no valuable prediction of the direction of change) is either rejected for both Phil-SPF, Green-Book and real-time average forecasts or, when the null is not rejected for Phil-SPF and Green-Book forecasts, it is often rejected in the case of real-time average forecasts. All in all, the main message is that (with the exceptions mentioned) Professional forecasts certainly lose marginal informational content when compared to the benchmark after 2/3 quarters, in line with the previous subsection.

### 2.3.3 Phil-SPF vs. Fed Green-Book

We now compare more carefully Phil-SPF and Green book forecasts to assess whether the former can be recognized as best practice, or close to best practice among professional forecasts. Table 3 reports the p-values for testing  $\gamma_1 = 0$  and  $\delta_1 = 0$  in Eqs. (4) and (5) with Phil-SPF as the benchmark, that is, to assess whether Phil-SPF forecasts encompass Fed-Green Book forecasts or vice-versa (with a common sample up to 2003q4). For every forecast horizon in the case of real output, we reject  $\gamma_1 = 0$  but we cannot reject  $\delta_1 = 0$ , indicating that Phil-SPF forecasts

---

<sup>12</sup>Recall however that the Fed-Green Book sample is different from the Phil-SPF one. We are still postponing a proper comparison between these forecasts.

for this variable embody useful predictive information above that contained in Fed-Green Book forecasts. In the case of GDP deflator inflation the evidence is not clear cut but favors the conclusion that Green book forecasts encompass Phil-SPF forecasts at  $h = 4, 5$ .

We proceed to an alternative way of comparing the forecast performance of these Professional forecasts by constructing a table with the ratio of MSFEs (Fed-Green Book's to Phil-SPF's) for different sample periods with and without recessions. Table 4 reports MSFE ratios for the full sample, an aggregation of recession periods, the sample up to 2001q4 and the sample up to 1991q4 for forecast horizons ranging from  $h = 1$  through  $h = 5$  both for real output growth and GDP deflator inflation.

Focusing on real GDP growth first, we observe that the MSFE ratios are above 1 for short forecast horizons in the various sub-samples that include recessions, implying that Phil-SPF's forecasts perform superiorly on average. For  $h = 4, 5$  the reverse is true, as a result of relatively large Phil-SPF forecast errors in the very beginning of the evaluation period, coinciding with the 1981/82 recession, see below. More generally, if recession periods are excluded from the full sample and both sub-samples, the MSE ratio increases (in general). As for GDP deflator inflation, we observe that Fed-Green Book's forecasts only perform superiorly when  $h$  is above 1 and for the samples (that include recessions) ending in 1991q4 and (less so) 2001q4 (the exception occurs with  $h = 5$ , due to Fed-Green Book's substantially larger errors from 1983 through 1989, see below). The general superiority observed in these earlier sub-samples clearly vanishes with the inclusion of more recent data. To test the significance of superior predictive ability, each MSFE ratio entry is matched with the p-value from a modified Diebold-Mariano (1995) test. The null hypothesis is that Fed-Green Book's and Phil-SPF's forecasts have equal predictive accuracy. Now, for both variables most p-values are above 10% when  $h = 3, 4, 5$ , suggesting that even though the MSFE is more often lower for Phil-SPF, the differences in forecast accuracy are not significant. More frequent exceptions to this pattern are found at very short horizons ( $h = 1, 2$ ) and favoring Phil-SPF's predictions.

We now move to a graphical comparison of the performance of Fed-Green Book's and Phil-SPF's predictions of real output growth and GDP deflator inflation. Figures 1.a-i show in the

upper panel 1-year (4 quarters) moving averages of squared forecast errors (Rolling-SFEs, with the last observation of the moving average in the  $x$  axis) and also (bottom panel) the cumulative squared forecast errors (from 1981Q1 through the date in the  $x$ -axis) for the two variables. The grey-shaded regions identify NBER recession periods. First, one can conclude that Rolling-SFE's rise during business cycle turning points and recessions. Second, for both variables and most forecast horizons we conclude that Fed-Green Book's forecast errors are generally lower than Phil-SPF's (albeit similar) up until the early 1990s (the exceptions occurs with output growth when  $h = 1, 2$  and GDP deflator inflation when  $h = 1$ ); in general, after this point we observe that the series of accumulated squared forecast errors get close and eventually cross (whenever the Phil-SPF series is above that of the Fed-Green book initially; when this did not occur, as with GDP growth with  $h = 1, 2$  and GDP deflator inflation when  $h = 1$ , the series drift farther apart). All this clearly indicates an overall relative deterioration of Fed-Green book's forecasts.

All in all, it seems fair to conclude that at least for these two variables Phil-SPF's forecasts are not inferior to Fed-Green Book's. In the remainder of the paper we will focus almost exclusively on Phil-SPF forecasts, regarding them as a good proxy of the forecasts produced by well-informed agents in the economy.

### 3 How does a Standard DSGE model forecast?

#### 3.1 The model

We move now towards the core of the paper, comparing previous results with the theoretical and empirical forecast performance of the medium-scale model analyzed and estimated in Smets and Wouters (2007) (henceforth SW07), based on Smets and Wouters (2003) and Christiano et al. (2005). The model has many of the features now popular in the growing so-called DSGE literature<sup>13</sup> including monopolistic competition in the goods and labour markets, ingredients

---

<sup>13</sup>see, e.g., Adolfson et. al. (2007, 2008) and Christiano et al. (2009) for further (and growing) models.

aimed at improving the fit of the model to observables such as habit formation in consumption, investment adjustment costs, variable capacity utilization (all of these implying amplification of the effects of shocks) and crucially, nominal frictions such as Calvo sticky prices and wages along with partial backward-looking indexation. Monetary policy follows a Taylor rule and has real effects when nominal frictions are important. Seven shocks are included (total factor productivity, investment productivity, monetary policy, government spending and risk premium following an AR(1) process along with price and wage markup shocks following an ARMA(1, 1) process) as well as seven observables: output, investment, consumption, wages (all in log differences) as well as inflation, nominal interest rates and (log of) hours. Further details can be found in Appendix B.

We start by solving a first order log-linear approximation of the model using the algorithm of Swanson et al. (2004). We then build the state space representation of the solution, including measurement equations linking deviations from the steady-state (or first differences of deviations from the steady-state whenever observed data is non-stationary) to the (quarterly) observables: (log differences of) output, consumption, investment, wages, (log of) hours and (levels of) nominal interest rate and inflation. We use exactly the same data treatment as in SW07, implying that the match between the model's variables and Phil-SPF's counterparts is not perfect. Specifically, SW07 observables for output, consumption, investment and wages are expressed in per capita (working age population) terms and nominal interest rates are measured with the Federal funds rate (quite close to the 3-month T-bill rate from Phil-SPF nonetheless). The inflation measure in the model is GDP deflator inflation (i.e., perfect match with Phil-SPF) whereas (minus) Phil-SPF's unemployment, while following closely hours, surely drifts somehow from the concept in the model.

We analyze the forecast performance of two versions of SW07: the original one featuring nominal rigidities, or New-Keynesian (NK) version, and another where we shut down these rigidities (i.e., flexible prices or RBC version, where we further reduce the observables by eliminating inflation and nominal interest rates while keeping the wage markup shock). We use Smets and Wouters's estimated parameters (mode of the posterior distribution, obtained from

combining the likelihood function with a set of independent priors for 36 of the 41 structural parameters included in the model. The other 5 parameters are fixed) using data from 1984q1 through 2004q2 (see details in Appendix B<sup>14</sup>). We choose this sample to avoid quibbles regarding the onset of the “Great Moderation” and likely changes in monetary policy within the period starting in 1966q1 (SW07’s beginning of the sample). We arguably go against the RBC version by not re-estimating the model, i.e., we keep fixed the structural parameters not related to nominal rigidities. Forecasts of the observables are just conditional expectations given the state-space model and are obtained with the Kalman filter, which is also used to derive the theoretical covariance matrix of the forecast errors for various horizons.<sup>15</sup>

We start with a theoretical analysis of the forecastability of the various variables implied by the model, i.e., we assume the model is the economy and derive analytically the standard deviation of the forecast errors at various horizons.<sup>16</sup> The (artificial) sample size is set at  $T = 160$  (thinking in 40 years of post war quarterly data), although it could be lower given the fact that the Kalman filter converges fast. Figure 2 presents the theoretical relative (to the standard deviation of the variables) root mean squared forecast error (RMSFE) for output, consumption, investment, inflation, hours, nominal interest rate and wages of the original SW07 (NK version). As easily seen, for nominal interest rates, inflation but also hours, there is a very strong predictability at short horizons, the relative RMSFE converges slowly and after 20 quarters this ratio is still around 0.4 for inflation and nominal interest rates and 0.7 for hours. For consumption, output and investment the initial level lies around 0.45-0.55 but convergence is fast except for investment. Wages is the least predictable variable, with a relative RMSFE starting around 0.8. All this means that a rational agent understanding this economy should

---

<sup>14</sup>Results for the empirical and theoretical forecast performance across all the variations considered in SW07 (i.e., shutting off, individually, the various real and nominal frictions) are available upon request. The main conclusions in this paper are robust to these different parametrizations.

<sup>15</sup>For the theoretical analysis this only implies that agents would be using a minimum mean square criterion if they were to pick this as a point forecast, i.e., they know the parameters of the model and produce conditional expectations given the state space model. Regarding the empirical analysis in the paper, it is fair to say that Bayesian estimation of the models would make natural using as point forecasts the mean of the predictive density of future observations, see e.g., Adolfson, Lindé, and Villani (2007).

<sup>16</sup>Here and in the rest of the paper, we could have looked as well at the correlation of the forecast errors for different variables, but avoided so for reasons of parsimony.

be able to forecast in such a way as to clearly outperform the unconditional mean in the case of hours, inflation, nominal interest rates even at very long horizons. For consumption, output and specially investment, he would clearly beat the mean even at 6 quarters ahead.

In the case of the RBC version (Figure 3) the conclusions are naturally quite different. The model becomes silent with respect to inflation and nominal interest rates but for the remaining variables the convergence of the RMSFE towards the standard deviation of the variables is much faster. For output, the relative RMSFE is around 0.8 for 1-step ahead forecasts and above 0.9 afterwards. For consumption and investment the speed of convergence is lower but clearly higher than that of the sticky prices/wages version. For wages, there is only significant predictability at 1-step ahead whereas for hours convergence of the RMSFE towards the standard deviation is slow but at a level clearly above that of the NK version. Now, it is important to note that this feature of the specific NK model analyzed here is certainly common to any model featuring price and wage setting frictions along with an important indexation mechanism (to target or current inflation or a combination of the two) aimed at rationalizing the observed persistence of inflation, see e.g., the models in Christiano et al. (2005), Adolfson et al. (2007), Ireland (2007) or Schorfheide (2005). This occurs because indexation generates high persistence in inflation and in other variables (and thus strong forecastability). In other way, any deviation of inflation from target in this kind of world represents a persistent (forecastable) deviation of the economy from its steady-state.

### 3.2 Model vs. Data

Here we confront the results in section 2, regarding Phil-SPF's forecasts, with the theoretical and empirical forecast accuracy of the NK and RBC versions of SW07 analyzed above. To be clear, we view the relative (to the standard deviation) RMSFE of well-informed agents in the economy (Professional forecasters), as a statistic that should be matched by a realistic DSGE model, just as it should deliver steady-state ratios, volatilities and correlations that are close to what is observed in the data. E.g., if this relative RMSFE for output growth at 1 quarter

horizon is 0.3 in the model and 0.8 in the case of Professional forecasters (data), we view this as an indication that the model delivers a forecastability that is at odds with the data. And similarly if after 10 quarters the Rational agent living in the model economy is still able to clearly outperform the mean whereas Professional forecasters don't. Comparison of Professional and Rational (given the model) forecasts can thus inform theory or at least show the limitations of the theoretical models, even though the mapping from Rational to Professional forecasts may be considered loose.

If nothing else, we believe Professional forecasts allow us to measure how fast (from the perspective of the forecasters) the economy is moving towards the steady-state. Specifically, we can measure this convergence to the steady-state through the speed with which the RMSFE converges to the standard deviation of the variables. In fact, if after some time (horizon) the forecast is (on average) very close to the unconditional mean of the variable under scrutiny, this means the forecaster believes the economy (or at least that variable) takes that much time to reach the steady-state (in the absence of unpredictable shocks). With rational expectations this must be a characteristic of the process generating the data.<sup>17</sup>

Now, results in the previous section suggest that for most real variables (and in particular output, investment and consumption) Professional forecasts lose grip after 2 quarters, meaning that using as forecast an estimate of the unconditional mean of the variables does not imply losing valuable information. Professional forecasts of unemployment and nominal interest rates are still clearly superior to the mean after one year whereas for inflation (CPI and GDP) there is forecastability but in a lesser extent. Notice further that we are using as benchmark a real-time estimate of the mean. If this mean is time-varying or shifts occasionally, e.g., if the steady-state changes due to permanent changes in taxation or in monetary policy (that changes for instance target inflation), the real-time average will not be efficient whereas professional forecasters are probably aware of these shifts. This is useful for our purposes as it allows us to interpret the relative RMSFE of Professional forecasts (which is thus deflated) as a **lower bound** on what

---

<sup>17</sup>We are certainly aware of the difficulty of characterizing as rational a consensus (mean or median) forecast, see e.g., Bonham and Cohen (2001). Rationality should arguably be analysed at the individual level but exit and entry of forecasters in the surveys makes this a difficult task.

a realistic theoretical model (without steady-state shifts) ought to deliver in terms of forecast accuracy relative to the steady-state forecast. Similarly then, in the mapping from Professional forecasts to the theoretical performance of SW07, we must see the model as corrected for regime changes, hence we cannot be as demanding when using the models in a pseudo out-of sample forecasting exercise with actual data.

Table 5 compares the results, for the theoretical and empirical (with actual data) relative (to the standard deviation) RMSFE of the NK and RBC versions of SW07, vis-a-vis that obtained with Phil-SPF forecasts. In the analysis of the empirical performance we focus on the sample 1981q3- 2009q2 (coinciding with the previous SPF's evaluation sample). If we take first the theoretical rel. RMSFE for output and investment, it is clear that the distance between the rel. RMSFE of Phil-SPF and that of the theoretical model is in general smaller for the RBC model, clearly so for all  $h$  in the case of investment and for  $h = 3, 4, 5$  in the case of output. At  $h = 1, 2$  in the case of output, the RBC has a clearly lower forecastability. This result for output and investment contrasts to what obtains with the NK version, where the strong predictability at  $h = 1$  and even at long horizons is at odds with Phil-SPF. In the case of consumption the RBC is more successful at matching the data when  $h = 1, 2$ , whereas for  $h = 3, 4, 5$  the evidence favors the NK version (notice however that it may well be the case that the rel. RMSFE obtained with Phil-SPF may not be statistically different from 1). With respect to hours/unemployment (we recall that Phil-SPF forecasts unemployment, which explains nonetheless around 80% of the variation in hours), the RBC is closer to Phil-SPF at all horizons, although the rel. RMSFE is consistently above that of Phil-SPF for  $h \geq 2$ . This is in clear contrast with the strong predictability implied by the NK model. The RBC version is silent with respect to the nominal interest rate and inflation but for the NK model it is clear that while the behavior of the rel. RMSFE function is very close to that of Phil-SPF in the case of nominal interest rates, for inflation the very high forecastability of the NK model does not match data from Phil-SPF forecasts. We notice also that even if the rational agent uses the forecasts produced by the univariate representation of inflation given the model (NK univariate, i.e, using only past inflation to produce the forecast), the strong forecastability of

inflation is almost unchanged. This is a consequence of the degree of backward-looking behavior (indexation) of inflation in the NK model. Once the rational agent observes current inflation and its history, information on other shocks is almost irrelevant to form a close to efficient conditional expectation of inflation at some point in the future. If the model is realistic, this implies that a forecaster would only need to nail the univariate representation of inflation in order to obtain a close to efficient forecast.

Now, demanding from the models forecasts of actual (observed) data changes radically, in absolute terms, the picture above, with a clear deterioration of the empirical counterpart of the statistics above.<sup>18</sup> Nonetheless, Table 5 (bottom panel) shows that for real output and investment the RBC is close to Phil-SPF (and dramatically superior to the NK version). For wages (no data for Phil-SPF) the performance of both models is very similar whereas for consumption both the RBC and NK versions have a very weak performance (although the latter performs relatively better at horizons greater than 5 quarters, despite the fact that forecasts are close to the mean). For nominal interest rates the NK model is close to Phil-SPF at  $h = 1, 2$  but it drifts quite fast afterwards, becoming useless after 6 quarters (in clear contrast with the theoretical result). For inflation, the empirical performance of the NK model is beyond terrible, a qualification also deserved for the behavior of the RBC version with respect to hours (in this case the NK version is clearly more informative but not much compared to Phil-SPF at  $h > 2$ ).

As far as we are aware, only Rubaszek and Skrzypczynski (2008) compared forecasts from a (3 equations prototypical) DSGE model to SPF forecasts while using real-time data for estimation and forecasting (instead of the latest vintage of data and a fixed set of parameters, useful for our purposes). Their sample size is also larger than usual, spanning 1994:q1–2006:q2. The main conclusions are that while for a few horizons in the case of GDP growth the DSGE model seems to outperform SPF (not statistically significant difference in accuracy), in the case of inflation and short-term nominal interest rate SPF clearly outperforms the DSGE model.<sup>19</sup>

---

<sup>18</sup>Again, it is fair to recognize that the literature acknowledges the likely misspecification of DSGE models. E.g., Del Negro et al. (2007) approximate a DSGE model by a vector autoregression (VAR) and then relax the implied cross-equation restrictions in order to improve fit. It is possible to optimally relax these restrictions and it is found that forecast accuracy improvements obtain.

<sup>19</sup>Edge et al. (2010) do compare the forecast performance of an alternative DSGE model to Green-Book

All in all, the results above suggest that the nominal rigidities apparatus of the NK model, which greatly amplifies the effects of shocks, tended to produce an excessively large theoretical forecastability, extending over long horizons. This seems clearly at odds with the data. The stripped down flexible prices version (RBC) delivers a forecastability resembling more that of the Phil-SPF while performing relatively better empirically (the important exception relates to hours/unemployment). This is due to the fact that deviations from the steady-state tend not to be persistent, hence forecasts (conditional expectations) are closer to the mean of the variables. Thus, not taking risks (or not assuming a detailed knowledge of short-run dynamics) compensated in this context. The RBC model seemed more immune to misspecification (notice also that the RBC version was not even re-estimated, it keeps all the parameters from the estimated NK model).

Next we repeat the analysis for recession periods.

## **4 How do Rational and Professional Forecasts behave during Recessions?**

There is clear evidence that macroeconomic forecasts fail to predict business-cycle's turning points and, moreover, forecasting the beginning of a recession one or two quarters in advance never occurred. In this aspect data (professional forecasts) are in line with standard models, where recessions must be seen as the result of large exogenous shocks (or at least unpredictable shocks in size and moment). Hence, one should not demand (or expect) accurate forecasts referred to the first period (quarter) of a recession. Afterwards, the theoretical mechanisms embodied in the models should be helpful in determining the path of observed variables.

Here we show that the conclusions above seem to carry over to recession<sup>20</sup> periods, and are certainly magnified. That is, the performance of the NK version of SW07 is quite poor

---

forecasts from 1996 through 2004, arguing for a positive contribution of the model in some instances (specially for output growth).

<sup>20</sup>As identified by the NBER dates. For the purposes of this section we include a quarter before and a quarter after the recessions to capture turning points.

compared to that of the RBC version. First, we recall that Professional forecasts (from the Fed-Green Book or SPF) have a poorer (relative to an estimate of the unconditional mean) performance during recession periods (Tables 1.1 and 1.2), specially at horizons greater or equal to 3 quarters. The exceptions occur with inflation and nominal interest rates as well as with housing market variables for short horizons (housing starts and residential investment). Despite this, they are clearly more accurate than model forecasts. To analyze this we simply plot the various forecasts (Fed-Green Book, Phil-SPF, NK and RBC) for  $h = 1, \dots, 5$  for real GDP growth, inflation and interest rates<sup>21</sup> (see Figures 4.a-i). Analysis of other real variables conveys a very similar message. As easily seen, Professional forecasts of real GDP have no clue about the beginning and dynamics of recessions with an anticipation of 2 or more quarters ( $h \geq 3$ ) whereas 1 quarter earlier they have some signal and for the current quarter they are accurate ( $h = 1$ , we recall that one step ahead forecasts in the case of Professional forecasts is really a nowcast). Now, although the RBC model performs poorly relative to professional forecasts, the characterization is very similar. The RBC obviously does not anticipate the recessions but provides signal about subsequent developments when  $h = 1, 2$ . The performance of the NK model is clearly disastrous, specially during the last recession, where observed deflation and very low nominal interest rates contribute to forecasts that never consider consecutive negative growth (but instead a quick way out of the recession). This is clearly not the case in the 1991 and 2001 recessions. Again, the “defensive” (or not implying persistent deviations from the steady -state) dynamics implicit in the RBC version seems to at least produce forecasts that have some signal (although definitely close to the steady-state, or unconditional mean). For inflation and nominal interest rates we observe that professional forecasts are very accurate at short horizons and convey some signal at longer horizons. For nominal interest rates, the NK model does not produce out of bounds forecasts, but they are weak compared to those of Professionals. For inflation, NK forecasts are very poor and do seem out of bounds, except during the last recession.

---

<sup>21</sup>Graphs for other real variables under scrutiny in this paper are available from the authors upon request.

## 5 Concluding remarks

It seems unwise to expect too much from macroeconomic forecasts. For what really matters (real variables, but except for unemployment) best practice has little to say at horizons greater than 2, 3 quarters. If statistics derived from these facts inform general equilibrium modelling, in the sense that a rational agent living in the economy should deduce similar statistics, they probably say the economy has not been deviating persistently from the steady-state. In the theoretical models, this should translate into low forecastability (relative to a naive, or steady-state forecast and, again, except for unemployment - hours) of most variables. This occurs with the RBC version of the model analyzed here but clearly not with the NK version. Furthermore, even recognizing limitations in a model without nominal frictions and (or) limited (not persistent) departures from the steady-state, the fact is that empirical forecasts seem to indicate that the model minimizing type II errors (or less prone to misspecification) is the RBC version. Forecasts are closer to naive (or to steady-state values) but provide some signal. The alternative (relying on a particular description of nominal rigidities) is not reliable. In our view, and given the effects of the inclusion of nominal frictions on forecast performance (theoretical and empirical), care should be taken at least on the way trend inflation (or varying central bank target) is modeled. In the model analyzed here and many others, the central bank target (steady-state inflation) is fixed, which implies that any deviation of inflation from target is necessarily interpreted by the model as a deviation from the steady-state (inflation gap). We are persuaded by Cogley and Sbordone's (2008) analysis that once movements in trend inflation are taken into account, the (backward-looking) indexation component of a general new-Keynesian Phillips curve is not needed to fit the data well. If indexation is incorrectly assumed, it implies a supposedly high theoretical forecastability of inflation (even if a rational agent only looks past inflation) as we have shown. This is clearly at odds with the data (Professional forecasts) and does not survive a forecast evaluation with actual data. Another interpretation of the results rests on the observation that theoretical models used to fit several decades of data are likely missing relevant changes in monetary policy, product and labor market regulation, taxation or in the

trend growth of technology. If these changes are reasonably unpredictable, there is potential compatibility between professional forecasters having a hard time and the NK model becoming seriously misspecified only along those dimensions, i.e., nominal rigidities can play an important role which is hidden due to lack of control for what can be seen as steady-state shifts.

## References

- [1] Adolfson, M., J. Lindé and M. Villani (2007). “Forecasting Performance of an Open Economy DSGE Model”, *Econometric Reviews*, 26: 289–328.
- [2] Adolfson, M., S. Laséen, J. Lindé and M. Villani (2007). “Bayesian Estimation of an Open Economy DSGE Model with Incomplete Pass-Through”, *Journal of International Economics*, 72:481–511.
- [3] Adolfson, M., S. Laséen, J. Lindé and M. Villani (2008). “Evaluating an Estimated New Keynesian Small Open Economy Model”, *Journal of Economic Dynamics and Control*, 32: 2690–2721.
- [4] Ang, A., Bekaert, G. and M. Wei. (2007). “Do Macro Variables, Asset Markets, or Surveys Forecast Inflation Better?”, *Journal of Monetary Economics*, 54(4): 1163-1212.
- [5] Bernanke, B.S. and J. Boivin (2003). “Monetary Policy in a Data-Rich Environment”, *Journal of Monetary Economics*, 50: 525-546.
- [6] Bonham, C. and R. H. Cohen (2001). “To Aggregate, Pool, or Neither: Testing the Rational Expectations Hypothesis Using Survey Forecasts,” *Journal of Business and Economic Statistics*, 19(3): 278-91.
- [7] Capistrán, C. and A. Timmermann (2009). “Forecast Combination with Entry and Exit of Experts.” *Journal of Business and Economic Statistics*, 27: 428-440.
- [8] Chong, Y. Y. and D. F. Hendry (1986). “Econometric evaluation of linear macroeconomic models”. *Review of Economic Studies*, 53, 671–690.
- [9] Christiano, L., Eichenbaum, M. and C. Evans (2005). “Nominal rigidities and the dynamic effects of a shock to monetary policy.” *Journal of Political Economy*, 113 (1): 1-45.
- [10] Christiano, L. J., Trabandt, M. and K. Walentin (2009). “Introducing Financial Frictions and Unemployment into a Small Open Economy Model”, *Sveriges Riksbank Working Paper Series No. 214*.
- [11] Cogley, T. and Sbordone, A. M. (2008). “Trend Inflation, Indexation, and Inflation Persistence in the New Keynesian Phillips Curve”, *American Economic Review*, 98(5): 2101-26.
- [12] Croushore, D. and T. Stark (2001). “A Real-Time Data Set for Macroeconomists”, *Journal of Econometrics*, 105:111-130.
- [13] Croushore, D. and T. Stark (2003). “A Real-Time Data Set for Macroeconomists: Does the Data Vintage Matter?” *Review of Economics and Statistics* 85:605–617.
- [14] Croushore, D. (1993). “The Survey of Professional Forecasters”, *Business Review*, Federal Reserve Bank of Philadelphia, November/December, 3-15.

- [15] Croushore, D. (2006). “An evaluation of inflation forecasts from surveys using real-time data”, Working Paper No. 06-19, Federal Reserve Bank of Philadelphia.
- [16] D’Agostino, A. and K. Whelan (2008). “Federal Reserve information during the Great Moderation”, *Journal of the European Economic Association*, 6(2-3): 609-620.
- [17] Del Negro, M., Schorfheide, F., Smets, F. and R. Wouters (2007). “On the Fit of New Keynesian Models”, *Journal of Business and Economic Statistics*, 25: 123-143
- [18] Diebold, F. X. and R. S. Mariano (1995). “Comparing predictive accuracy”, *Journal of Business and Economic Statistics*, 13:252-263.
- [19] Edge, R. M., Kiley, M. T. and J. P. Laforge (2010). “A comparison of forecast performance between Federal Reserve staff forecasts, simple reduced-form models, and a DSGE model”, *Journal of Applied Econometrics*, 25(4): 720-754
- [20] Engelberg, J., Manski, C. F. and J. Williams (2009). “Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters”, *Journal of Business and Economic Statistics*, 27: 30-41.
- [21] Fair, R. C. and R. J. Shiller (1989). “The Informational Context of Ex Ante Forecasts”, *The Review of Economics and Statistics*, 71(2): 325-31
- [22] Faust, J. and J. H. Wright (2007). “Comparing Greenbook and Reduced Form Forecasts using a Large Realtime Dataset”, NBER Working Papers 13397
- [23] Faust, J. and J. H. Wright (2008). “Efficient forecast tests for conditional policy forecasts”, *Journal of Econometrics*, 146(2), 293-303.
- [24] Gamber, E. and J. Smith (2009). “Are the Fed’s inflation forecasts still superior to the the Private Sector’s?”, *Journal of Macroeconomics*, 31(2): 240-251.
- [25] Garcia, J. (2003). “An introduction to the ECB’s Survey of Professional Forecasters”, ECB Occasional Paper, No.8, September.
- [26] Gavin, W. T. and R. J. Mandal, (2003). “Evaluating FOMC forecasts”, *International Journal of Forecasting*, 19: 655–667.
- [27] Giannone, D., Lenza, M. and L. Reichlin (2008). “Explaining The Great Moderation: It Is Not The Shocks”, *Journal of the European Economic Association*, 6(2-3), pp. 621-633
- [28] Harding, D. and A. Pagan (2006). “Synchronization of cycles”, *Journal of Econometrics*, 132(1), 59-79.
- [29] Ireland, Peter N. (2007). “Changes in the Federal Reserve’s Inflation Target: Causes and Consequences” *Journal of Money, Credit and Banking*, 39(8): 1851-2110.
- [30] Granger, C. W. J. and P. Newbold (1986). *Forecasting economic time series* (2nd ed.). London: Academic Press.

- [31] Joutz, F. and H. O. Stekler, (2000). “An evaluation of the predictions of the Federal Reserve”, *International Journal of Forecasting*, 16, 17–38.
- [32] Kiefer, N. and T. Vogelsang (2002). “Heteroskedasticity autocorrelation robust standard errors using the Bartlett kernel without truncation”, *Econometrica* 70: 2093–2095.
- [33] McConnell, M. and G. Perez-Quiros (2000). “Output Fluctuations in the United States: What Has Changed Since the Early 1980’s?”, *American Economic Review*, 90 (5): 1464-1476
- [34] McNees, S.K., (1992), “How large are economic forecast errors?” *New England Economic Review*, Jul/Aug 25–33.
- [35] McNees, S.K. and J. Ries (1983). “The track record of macroeconomic forecasts”. *New England Economic Review*, Nov/Dec 5–18.
- [36] Phillips, P.C.B., Sun, Y. and S. Jin (2003). “Consistent HAC Estimation and Robust Regression Testing when using Sharp Original Kernels with no Truncation”, *Cowles Foundation for Economic Research*, Yale University.
- [37] Romer, C. and D. Romer (2000), “Federal Reserve information and the behaviour of interest rates”, *American Economic Review*, 90: 429-457.
- [38] Rubaszek, M., and P. Skrzypczynski (2008). “On the Forecasting Performance of a Small-Scale DSGE Model,” *International Journal of Forecasting*, 24: 498-512.
- [39] Schorfheide, F. (2005). “Learning and monetary policy shifts.” *Review of Economic Dynamics*, 8(2): 392-419.
- [40] Smets, F. and R. Wouters (2003). “An Estimated Dynamic Stochastic General Equilibrium Model of the Euro Area”, *Journal of the European Economic Association*, 1(5): 1123-1175.
- [41] Smets, F. and R. Wouters. (2007). “Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach”, *American Economic Review*, 97(3): 586-606.
- [42] Stark, T. (2010). “Realistic evaluation of real-time forecasts in the Survey of Professional Forecasters”, *Federal Reserve Bank of Philadelphia, Research Rap Special Report*, May 2010
- [43] Stock, J.H. and M.W. Watson (2003). “Has the Business Cycle Changed and Why?,” *NBER Chapters*, in: *NBER Macroeconomics Annual 2002*, Vol17, pp 159-230, National Bureau of Economic Research
- [44] Stock, J.H., and M.W. Watson (2007). “Why Has U.S. Inflation Become Harder to Forecast?”, *Journal of Money, Credit, and Banking*, 39:3-34
- [45] Swanson, E., Anderson, G. and A. Levin (2005). “Higher-Order Perturbation Solutions to Dynamic, Discrete-Time Rational Expectations Models”, *mimeo*, Federal Reserve Bank of San Francisco.
- [46] Zarnowitz, V. (1969). “The new ASA-NBER Survey of Forecasts by Economic Statisticians”, *American Statistician*, 23: 12-16.

- [47] Zarnowitz, V., (1992). Business cycles: Theory, history, indicators and forecasting. In: National Bureau of Economic Research Studies in Business Cycles, vol. 27. University of Chicago Press, Chicago.
- [48] Zarnowitz, V., Braun, P., (1992). Twenty two years of the NBER-ASA Quarterly Outlook Surveys: Aspects and comparisons of forecasting performance. NBER Working Paper 3965, New York.
- [49] Zarnowitz, V. and Braun, P. (1993). Twenty-Two years of the NBER-ASA Quarterly Economic Outlook Surveys: Aspects and comparisons of forecasting performance, in J.H. Stock and M.W. Watson (eds.), Business Cycles, Indicators and Forecasting (NBER Studies in Business Cycles, Volume 28), 11-84, Chicago: University of Chicago Press.

## Appendix A - Variables' definitions and sources

Our data for both the Phil-SPF and Fed-Green-Book predictions comes directly from the Federal Reserve Bank of Philadelphia website and covers the period 1968:q4-2009:q2<sup>22</sup>; The latest vintage of all series (against which forecasts are compared) was downloaded from the FRED database (Federal Reserve Bank of St. Louis). Data is converted to quarterly whenever the series are available at a higher frequency (by averaging the observations within each quarter). Except for unemployment and interest rates, all data is in growth rates. Except for interest rates, all published data is seasonally adjusted, in accordance with the targets of Professional's forecasts. Prior to 1992, nominal and real output forecasts refer to nominal GNP. GDP deflator forecasts refer to GNP deflator prior to 1992, to GDP deflator from 1992 through 1995 and to chain-weighted price index for GDP since 1996.

The following table shows the definition of all series, Phil-SPF's and FRED's id code, and observation range.

Definition	FRED code	SPF code
Gross Domestic Product (Nominal)	GDP	NOUTPUT
Gross Domestic Product (Real)	GDPC1	ROUTPUT
Real Personal Consumption Expenditures	PCECC96	RCONS
Real Private Nonresidential Fixed Investment	PNFIC96	RINVBFI
Real Private Residential Fixed Investment	PRFIC96	RINVRESID
Housing Starts Total: New Privately Owned Housing Units	HOUST	HSTARTS
Industrial Production Index	INDPRO	IPT
Real Federal Cons. Exp. & Gross Investment	FGCEC1	RGF
Real State & Local Cons. Exp. & Gross Investment	SLCEC1	RGLS
Net Exports of Goods & Services	NETEXP	NETEXP
Civilian Unemployment Rate	UNRATE	UNRATE
Consumer Price Index: All Items	CPIAUCSL	CPI
Gross Domestic Product Deflator	GDPDEF	GDPDEF
10-Year Treasury Constant Maturity Rate	GS10	GS10
3-Month Treasury Bill: Secondary Market Rate	TB3MS	TB3MS

As for data used in SW07, we note that the latest available vintage of the following series is used: FRED's GDPC1 (real output), PCECC96 (consumption) and PNFIC96 (investment),

<sup>22</sup><http://www.phil.frb.org/econ/spf/spfpage.html>. For a recent discussion about the Phil-SPF see Croushore (2006). For a recent discussion on differences between surveys, do refer to Engelberg et al. (2007) and Clements (2008).

which are divided by the Bureau of Labor Statistics (BLS) series LNU00000000Q (Civilian noninstitutional population, 16 years and over). Short term nominal interest rate is the Federal funds rate (FRED's FEDFUNDS), inflation is measured with FRED's GDPDEF (perfect match with SPF), hours worked is obtained as average weekly hours (nonfarm business, BLS's PRS8500602) multiplied by civilian employment (16 and over, FRED's CE16OV) and then divided by BLS's series LNU00000000Q (Civilian noninstitutional population, 16 years and over)

## Appendix B - Smets and Wouters (2007) model

Table 1  
Log-linearized equations of the SW07 model (sticky-price-wage economy)

---



---

(1)	$y_t = c_y c_t + i_y i_t + r^{kss} k_y z_t + \varepsilon_t^g$
(2)	$c_t = \frac{\lambda/\gamma}{1 + \lambda/\gamma} c_{t-1} + \frac{1}{1 + \lambda/\gamma} E_t c_{t+1} + \frac{w^{ss} l^{ss} (\sigma_c - 1)}{c^{ss} \sigma_c (1 + \lambda/\gamma)} (l_t - E_t l_{t+1})$ $-\frac{1-\lambda/\gamma}{(1+\lambda/\gamma)\sigma_c} (r_t - E_t \pi_{t+1}) - \frac{1-\lambda/\gamma}{(1+\lambda/\gamma)\sigma_c} \varepsilon_t^b$
(3)	$i_t = \frac{1}{1+\beta\gamma(1-\sigma_c)} i_{t-1} + \frac{\beta\gamma(1-\sigma_c)}{1+\beta\gamma(1-\sigma_c)} E_t i_{t+1} + \frac{1}{\varphi\gamma^2(1+\beta\gamma(1-\sigma_c))} q_t + \varepsilon_t^i$
(4)	$q_t = \beta(1-\delta)\gamma^{-\sigma_c} E_t q_{t+1} - r_t + E_t \pi_{t+1} + (1-\beta(1-\delta)\gamma^{-\sigma_c}) E_t r_{t+1}^k - \varepsilon_t^b$
(5)	$y_t = \phi_p(\alpha k_t^s + (1-\alpha)l_t + \varepsilon_t^a)$
(6)	$k_t^s = k_{t-1} + z_t$
(7)	$z_t = \frac{1-\psi}{\psi} r_t^k$
(8)	$k_t = (1-\delta)/\gamma k_{t-1} + (1-(1-\delta)/\gamma) i_t + (1-(1-\delta)/\gamma) \varphi\gamma^2(1+\beta\gamma(1-\sigma_c)) \varepsilon_t^i$
(9)	$\mu_t^p = \alpha(k_t^s - l_t) - w_t + \varepsilon_t^a$
(10)	$\pi_t = \frac{\beta\gamma(1-\sigma_c)}{1+\iota_p\beta\gamma(1-\sigma_c)} E_t \pi_{t+1} + \frac{\iota_p}{1+\beta\gamma(1-\sigma_c)\iota_p} \pi_{t-1} - \frac{(1-\beta\gamma(1-\sigma_c)\xi_p)(1-\xi_p)}{(1+\iota_p\beta\gamma(1-\sigma_c))(1+(\phi_p-1)\varepsilon_p)\xi_p} \mu_t^p + \varepsilon_t^p$
(11)	$r_t^k = l_t + w_t - k_t$
(12)	$\mu_t^w = w_t - \sigma_l l_t - \frac{1}{1-\lambda/\gamma} (c_t - \lambda/\gamma c_{t-1})$
(13)	$w_t = \frac{\beta\gamma(1-\sigma_c)}{1+\beta\gamma(1-\sigma_c)} (E_t w_{t+1} + E_t \pi_{t+1}) + \frac{1}{1+\beta\gamma(1-\sigma_c)} (w_{t-1} + \iota_w \pi_{t-1}) - \frac{1+\beta\gamma(1-\sigma_c)\iota_w}{1+\beta\gamma(1-\sigma_c)} \pi_t$ $-\frac{(1-\beta\gamma(1-\sigma_c)\xi_w)(1-\xi_w)}{(1+\beta\gamma(1-\sigma_c))(1+(\phi_w-1)\varepsilon_w)\xi_w} \mu_t^w + \varepsilon_t^w$
(14)	$r_t = \rho r_{t-1} + (1-\rho)(r_\pi \pi_t + r_y(y_t - y_t^*)) + r_{\Delta y}((y_t - y_t^*) - (y_{t-1} - y_{t-1}^*)) + \varepsilon_t^r$
(15)	$\varepsilon_t^a = \rho_a \varepsilon_{t-1}^a + \eta_t^a$
(16)	$\varepsilon_t^b = \rho_b \varepsilon_{t-1}^b + \eta_t^b$
(17)	$\varepsilon_t^g = \rho_g \varepsilon_{t-1}^g + \rho_{ga} \eta_t^a + \eta_t^g$
(18)	$\varepsilon_t^i = \rho_I \varepsilon_{t-1}^i + \eta_t^i$
(19)	$\varepsilon_t^r = \rho_r \varepsilon_{t-1}^r + \eta_t^r$
(20)	$\varepsilon_t^p = \rho_p \varepsilon_{t-1}^p + \eta_t^p - \mu_p \eta_{t-1}^p$
(21)	$\varepsilon_t^w = \rho_w \varepsilon_{t-1}^w + \eta_t^w - \mu_w \eta_{t-1}^w$

---

Note: The model variables are: output ( $y_t$ ), consumption ( $c_t$ ), investment ( $i_t$ ), utilized and installed capital ( $k_t^s$ ,  $k_t$ ), capacity utilization ( $z_t$ ), rental rate of capital ( $r_t^k$ ), Tobin's  $q$  ( $q_t$ ), price and wage markup ( $\mu_t^p$ ,  $\mu_t^w$ ), inflation rate ( $\pi_t$ ), real wage ( $w_t$ ), total hours worked ( $l_t$ ), and nominal interest rate ( $r_t$ ). The shocks are: total factor productivity ( $\varepsilon_t^a$ ), investment-specific technology ( $\varepsilon_t^i$ ), government purchases ( $\varepsilon_t^g$ ), risk premium ( $\varepsilon_t^b$ ), monetary policy ( $\varepsilon_t^r$ ), wage markup ( $\varepsilon_t^w$ ) and price markup ( $\varepsilon_t^p$ ).

Table 2  
Log-linearized equations of the SW07 model (flexible-price-wage economy)

---



---

(1)	$y_t^* = c_y c_t^* + i_y i_t^* + r^{kss} k_y z_t^* + \varepsilon_t^g$
(2)	$c_t^* = \frac{\lambda/\gamma}{1 + \lambda/\gamma} c_{t-1}^* + \frac{1}{1 + \lambda/\gamma} E_t c_{t+1}^* + \frac{w^{ss} l^{ss} (\sigma_c - 1)}{c^{ss} \sigma_c (1 + \lambda/\gamma)} (l_t^* - E_t l_{t+1}^*)$ $- \frac{1-\lambda/\gamma}{(1+\lambda/\gamma)\sigma_c} r_t^* - \frac{1-\lambda/\gamma}{(1+\lambda\lambda/\gamma)\sigma_c} \varepsilon_t^b$
(3)	$i_t^* = \frac{1}{1+\beta\gamma^{(1-\sigma_c)}} i_{t-1}^* + \frac{\beta\beta\gamma^{(1-\sigma_c)}}{1+\beta\gamma^{(1-\sigma_c)}} E_t i_{t+1}^* + \frac{1}{\varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)})} q_t^* + \varepsilon_t^i$
(4)	$q_t^* = \beta(1-\delta)\gamma^{-\sigma_c} E_t q_{t+1}^* - r_t^* + (1-\beta(1-\delta)\gamma^{-\sigma_c}) E_t r_{t+1}^{k*} - \varepsilon_t^b$
(5)	$y_t^* = \phi_p(\alpha k_t^{s*} + (1-\alpha)l_t^* + \varepsilon_t^a)$
(6)	$k_t^{s*} = k_{t-1}^* + z_t^*$
(7)	$z_t^* = \frac{1-\psi}{\psi} r_t^{k*}$
(8)	$k_t^* = (1-\delta)/\gamma k_{t-1}^* + (1-(1-\delta)/\gamma) i_t^* + (1-(1-\delta)/\gamma) \varphi\gamma^2(1+\beta\gamma^{(1-\sigma_c)}) \varepsilon_t^i$
(9)	$\mu_t^{p*} = \alpha(k_t^{s*} - l_t^*) - w_t^* + \varepsilon_t^a$
(10)	$\mu_t^{p*} = 1$
(11)	$r_t^{k*} = l_t^* + w_t^* - k_t^*$
(12)	$\mu_t^{w*} = -\sigma_l l_t^* - \frac{1}{1-\lambda/\gamma} (c_t^* + \lambda/\gamma c_{t-1}^*)$
(13)	$w_t^* = \mu_t^{w*}$

---

Note: The model variables are: output ( $y_t^*$ ), consumption ( $c_t^*$ ), investment ( $i_t^*$ ), utilized and installed capital ( $k_t^{s*}$ ,  $k_t^*$ ), capacity utilization ( $z_t^*$ ), rental rate of capital ( $r_t^{k*}$ ), Tobin's  $q$  ( $q_t^*$ ), price and wage markup ( $\mu_t^{p*}$ ,  $\mu_t^{w*}$ ), real wage ( $w_t^*$ ), and total hours worked ( $l_t^*$ ).

Table 3  
Prior and posterior distribution

par.	interpretation	prior			posterior
		density	mean	std.	mode
$\varphi$	invest. adj. cost	$\mathcal{N}$	4.000	1.500	6.23
$\sigma_c$	inv. elast. intert. subst.	$\mathcal{N}$	1.500	0.375	1.47
$\lambda$	habit	$\mathcal{B}$	0.700	0.100	0.68
$\xi_w$	wage rigidity	$\mathcal{B}$	0.500	0.100	0.74
$\sigma_l$	inv. elast. hours	$\mathcal{N}$	2.000	0.750	2.30
$\xi_p$	price rigidity	$\mathcal{B}$	0.500	0.100	0.73
$\iota_w$	wage indexation	$\mathcal{B}$	0.500	0.150	0.46
$\iota_p$	price indexation	$\mathcal{B}$	0.500	0.150	0.21
$\psi$	cap. utilization cost	$\mathcal{B}$	0.500	0.150	0.69
$\Phi$	fixed cost	$\mathcal{N}$	1.250	0.125	1.54
$r_\pi$	response to inflation	$\mathcal{N}$	1.500	0.250	1.77
$\rho$	int rate smoothing	$\mathcal{B}$	0.750	0.100	0.84
$r_y$	response to output	$\mathcal{N}$	0.125	0.050	0.08
$r_{\Delta y}$	response to output growth	$\mathcal{N}$	0.125	0.050	0.16
$\bar{\pi}$	st.state infl.	$\mathcal{G}$	0.625	0.100	0.67
$\beta'$	discount factor*	$\mathcal{G}$	0.250	0.100	0.12
$\bar{l}$	st. state hours	$\mathcal{N}$	0.000	2.000	-0.55
$\gamma$	trend growth rate	$\mathcal{N}$	0.400	0.100	0.44
$\alpha$	capital share	$\mathcal{N}$	0.300	0.050	0.21
$\delta$	depreciation rate <sup>†</sup>	-	-	-	0.025
$\lambda_w$	wage markup <sup>†</sup>	-	-	-	1.50
$g_y$	government/output <sup>†</sup>	-	-	-	0.18
$\rho_a$	AR prod. shock	$\mathcal{B}$	0.500	0.200	0.94
$\rho_b$	AR risk premium	$\mathcal{B}$	0.500	0.200	0.14
$\rho_g$	AR government	$\mathcal{B}$	0.500	0.200	0.96
$\rho_I$	AR investment	$\mathcal{B}$	0.500	0.200	0.64
$\rho_r$	AR mon. policy	$\mathcal{B}$	0.500	0.200	0.29
$\rho_p$	AR price markup	$\mathcal{B}$	0.500	0.200	0.74
$\rho_w$	AR wage markup	$\mathcal{B}$	0.500	0.200	0.82
$\mu_p$	MA price markup	$\mathcal{B}$	0.500	0.200	0.59
$\mu_w$	MA wage markup	$\mathcal{B}$	0.500	0.200	0.62
$\rho_{ga}$	prod. shock in G	$\mathcal{B}$	0.500	0.200	0.39
$\sigma_a$	st.dev. prod. shock	$\mathcal{IG}$	0.100	2.000	0.35
$\sigma_b$	st.dev. risk premium	$\mathcal{IG}$	0.100	2.000	0.18
$\sigma_g$	st.dev. government	$\mathcal{IG}$	0.100	2.000	0.41
$\sigma_I$	st.dev. investment	$\mathcal{IG}$	0.100	2.000	0.39
$\sigma_r$	st.dev. mon. policy	$\mathcal{IG}$	0.100	2.000	0.12
$\sigma_p$	st.dev. price markup	$\mathcal{IG}$	0.100	2.000	0.11
$\sigma_w$	st.dev. wage markup	$\mathcal{IG}$	0.100	2.000	0.21

Note:  $\mathcal{N}$  is Normal distribution,  $\mathcal{B}$  is Beta-distribution,  $\mathcal{G}$  is Gamma distribution,  $\mathcal{IG}$  is Inverse Gamma distribution.

\*  $\beta' 100(\beta^{-1} - 1)$  where  $\beta$  is the discount factor.

† these parameters are assumed known in SW07.

**TABLE 1.1: Forecast Performance of the Philadelphia Survey of Professional Forecasters (PARTIAL TABLE – full table available upon request)**

	Real Time Average Forecast (RMSFE)	Rel. RMSE AR	$\beta_1$ AR - OLS	Rel. RMSE SPF	$\beta_1$ SPF - OLS	Real Time Average Forecast (RMSFE)	Rel. RMSE AR	$\beta_1$ AR - OLS	Rel. RMSE SPF	$\beta_1$ SPF - OLS	Real Time Average Forecast (RMSFE)	Rel. RMSE AR	$\beta_1$ AR - OLS	Rel. RMSE SPF	$\beta_1$ SPF - OLS	Real Time Average Forecast (RMSFE)	Rel. RMSE AR	$\beta_1$ AR - OLS	Rel. RMSE SPF	$\beta_1$ SPF - OLS	Real Time Average Forecast (RMSFE)	Rel. RMSE AR	$\beta_1$ AR - OLS	Rel. RMSE SPF	$\beta_1$ SPF - OLS
Spec.	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)
Time/H	1					2					3					4					5				
NOMINAL GNP/GDP (NOUTPUT)																									
Full	2.93	1.06	0.04 0.05	0.68	0.11** 0.05	2.93	1.13	0.01 0.06	0.81	0.11* 0.06	2.93	1.18	0.03 0.06	0.93	0.05 0.06	2.93	1.19	0.02 0.07	0.99	0.03 0.06	2.93	1.21	-0.004 0.07	1.07	0.02 0.05
Rec.:	3.29	1.48	0.10 0.26	0.79	0.12 0.09	3.29	1.55	0.22 0.24	1.05	0.09 0.07	3.29	1.64	0.09 0.16	1.30	0.02 0.08	3.28	1.62	0.32 0.25	1.42	0.01 0.08	3.26	1.66	0.41* 0.24	1.56	0.02 0.07
REAL GNP/GDP (ROUTPUT)																									
Full	3.02	0.86	0.03 0.14	0.66	0.26*** 0.08	3.03	0.89	0.07 0.20	0.78	0.21* 0.12	3.02	0.94	0.15 0.38	0.91	0.01 0.17	3.03	0.96	0.38 0.39	0.96	-0.09 0.19	3.04	1.01	0.13 0.20	0.98	-0.03 0.17
Rec.:	3.05	1.30	-0.01 0.11	0.72	0.11 0.12	3.06	1.36	0.11 0.18	1.05	-0.15 0.16	3.05	1.45	0.09 0.28	1.33	-0.31 0.17	3.08	1.48	0.88* 0.49	1.46	-0.34* 0.17	3.08	1.52	0.58 0.50	1.49	0.31* 0.16
INDUSTRIAL PRODUCTION INDEX – TOTAL (IPT)																									
Full	4.92	0.76	0.75*** 0.09	0.65	0.49*** 0.08	4.93	0.94	1.85*** 0.24	0.91	0.23 0.22	4.93	0.98	1.71*** 0.27	1.07	-0.12 0.23	4.93	1.16	0.20 0.16	1.13	-0.23 0.23	4.93	1.18	-0.18 0.15	1.18	-0.27 0.18
Rec.:	5.84	0.92	0.73*** 0.26	0.76	0.52*** 0.13	5.82	1.24	0.56*** 0.14	1.19	0.13 0.29	5.82	1.30	0.52*** 0.11	1.48	-0.20 0.29	5.82	1.59	0.36 0.24	1.59	-0.25 0.33	5.82	1.59	-0.24 0.24	1.66	-0.23 0.31
REAL PERSONAL CONSUMPTION EXPENDITURES – TOTAL (RCON)																									
Full	2.87	0.75	0.35* 0.22	0.68	0.42*** 0.15	2.87	0.76	0.33 0.25	0.74	0.57** 0.23	2.90	0.77	0.35 0.28	0.78	0.35 0.38	2.90	0.80	0.69 0.57	0.84	-0.24 0.42	2.90	0.86	-0.15 0.33	0.83	0.34 0.26
Rec.:	2.02	1.25	0.37 0.24	0.97	0.29* 0.18	2.03	1.28	0.38 0.35	1.10	0.45** 0.19	2.05	1.33	0.50 0.33	1.18	-0.19 0.33	2.04	1.49	0.91** 0.35	1.42	-0.45 0.34	2.03	1.58	-0.001 0.36	1.42	0.23 0.32
CONSUMER PRICE INDEX (CPI)																									
Full	2.13	1.41	-0.28 0.44	0.56	1.40*** 0.24	2.13	1.44	-0.38 0.33	0.78	0.78*** 0.26	2.14	1.46	-0.35 0.29	0.85	0.56*** 0.21	2.14	1.50	-0.24 0.18	0.88	0.42*** 0.16	2.15	1.56	-0.15 0.12	0.89	0.39** 0.15
Rec.:	3.17	1.24	-0.06 0.40	0.51	0.89*** 0.21	3.18	1.28	-0.22 0.32	0.74	0.52** 0.17	3.18	1.30	-0.11 0.22	0.81	0.46** 0.15	3.18	1.34	-0.07 0.15	0.85	0.43** 0.15	3.18	1.35	-0.07 0.12	0.86	0.41** 0.16
REAL GROSS PRIVATE DOMESTIC INVESTMENT – RESIDENTIAL (RINVRESID)																									
Full	19.76	0.60	0.19* 0.10	0.61	0.21*** 0.08	19.76	0.73	0.41*** 0.21	0.76	-0.02 0.07	19.74	0.81	0.59*** 0.21	0.88	-0.05 0.06	19.75	0.86	0.11 0.14	0.95	-0.12 0.08	19.76	0.87	0.09 0.11	0.97	-0.17* 0.10
Rec.:	20.27	0.55	0.09 0.23	0.42	0.04 0.11	20.28	0.79	0.29 0.41	0.76	-0.12 0.08	20.31	0.91	0.52 0.55	1.02	-0.03 0.10	20.28	0.95	0.06 0.11	1.17	-0.08 0.07	20.30	0.95	0.15 0.12	1.19	-0.07 0.06
REAL GROSS PRIVATE DOMESTIC INVESTMENT – NONRESIDENTIAL (RINVBFI)																									
Full	10.20	0.83	0.20*** 0.07	0.65	0.19* 0.10	10.23	0.84	0.23** 0.09	0.73	0.18 0.14	10.21	0.91	0.22 0.16	0.79	0.20 0.15	10.22	0.99	0.08 0.23	0.90	0.15 0.11	10.23	1.05	-0.18 0.21	0.94	0.16 0.17
Rec.:	11.61	1.02	0.24* 0.14	0.71	-0.02 0.24	11.66	1.07	0.32 0.19	0.78	-0.26 0.31	11.64	1.16	0.54 0.41	0.88	-0.26 0.36	11.68	1.22	0.65 0.70	1.00	-0.24 0.24	11.70	1.30	0.11 0.20	1.08	0.17 0.66
REAL GOVERNMENT CONSUMPTION AND GROSS INVESTMENT – STATE AND LOCAL (RGLS)																									
Full	3.02	0.83	0.31* 0.18	0.82	0.19*** 0.07	3.04	0.84	0.16 0.21	0.82	0.35*** 0.08	3.04	0.82	0.11 0.12	0.83	0.24* 0.12	3.03	0.82	0.36** 0.14	0.85	0.48*** 0.11	3.05	0.83	0.49** 0.20	1.01	0.06 0.07
Rec.:	2.61	1.08	0.73*** 0.23	1.08	0.001 0.08	2.63	1.05	0.41** 0.18	1.04	-0.06 0.12	2.63	1.04	0.14 0.14	1.03	-0.06 0.13	2.67	1.03	0.43** 0.16	0.98	0.19 0.22	2.68	1.05	0.56** 0.20	0.99	-0.01 0.05
REAL GOVERNMENT CONSUMPTION AND GROSS INVESTMENT – FEDERAL (RGF)																									
Full	7.22	1.03	-0.03 0.05	0.95	-0.11** 0.05	7.23	1.02	0.08 0.05	0.96	-0.01 0.07	7.23	1.08	0.06 0.06	0.98	0.08 0.09	7.24	1.06	0.01 0.06	1.01	0.11 0.08	7.24	1.01	0.06 0.09	0.96	0.06 0.10
Rec.:	7.52	1.00	-0.002 0.09	0.89	0.17*** 0.06	7.53	0.86	0.09 0.08	0.91	-0.13 0.09	7.55	1.02	-0.01 0.15	0.93	-0.04 0.16	7.53	1.04	0.02 0.13	0.98	-0.08 0.14	7.55	0.92	0.16 0.18	0.97	-0.07 0.23

Table 1.1 (continued)

HOUSING STARTS (HSTARTS)																									
Full	29.32	1.02	0.39	0.84	1.02***	29.33	1.04	0.04	1.03	0.72**	29.33	1.10	-0.92	1.11	0.42	29.33	1.07	0.03	1.18	-0.14	29.33	1.05	0.83	1.16	-0.29
			0.44		0.26			0.81		0.35			2.64		0.31			0.88		0.33			0.61		0.41
Rec.:	34.00	0.88	0.39	0.59	1.55***	34.03	0.96	-0.79	1.21	0.61	34.03	1.05	-7.44**	1.37	0.56	34.03	0.97	-0.19	1.46	0.40	34.03	0.94	0.77	1.37	0.75
			0.62		0.36			0.96		0.49			2.61		0.52			1.12		0.57			0.53		0.86
GROSS DOMESTIC PRODUCT: IMPLICIT PRICE DEFLATOR (GDPDEF)																									
Full	1.92	1.43	0.22	0.67	0.49***	1.94	1.44	0.14	0.70	0.58***	1.94	1.44	0.08	0.72	0.63***	1.94	1.53	0.07	0.75	0.65***	1.94	1.59	0.07	0.77	0.53***
			0.23		0.15			0.22		0.17			0.18		0.15			0.14		0.15			0.12		0.16
Rec.:	2.15	1.56	0.04	0.72	0.44*	2.17	1.58	-0.26	0.73	0.52**	2.17	1.59	-0.14	0.75	0.67**	2.17	1.68	0.03	0.77	0.63**	2.17	1.67	-0.04	0.76	0.55**
			0.52		0.24			0.54		0.24			0.48		0.23			0.33		0.25			0.27		0.24
CIVILIAN UNEMPLOYMENT RATE (UNRATE)																									
Full	1.71	0.21	-0.03	0.20	-0.06*	1.73	0.22	0.02	0.22	-0.09**	1.73	0.25	0.03	0.24	0.09	1.76	0.33	0.04	0.30	0.12*	1.75	0.41	0.04*	0.36	0.11**
			0.03		0.03			0.02		0.04			0.02		0.08			0.03		0.07			0.02		0.06
Rec.:	2.23	0.26	-0.01	0.24	-0.14*	2.27	0.28	0.11	0.25	-0.15**	2.25	0.31	0.14*	0.28	0.26*	2.27	0.37	0.13**	0.30	0.25***	2.29	0.39	0.12*	0.37	0.22***
			0.05		0.08			0.09		0.06			0.08		0.14			0.06		0.07			0.07		0.04
NET EXPORTS OF GOODS AND SERVICES (NETEXP)																									
Full	55.59	0.89	-0.04	2.32	0.02***	55.97	0.46	-0.22	2.00	-0.02	56.15	1.21	-0.07	1.98	-0.01	56.07	0.99	-0.06	1.77	-0.02	56.24	1.19	-0.05	1.68	-0.02
			0.04		0.01			0.25		0.01			0.08		0.01			0.12		0.02			0.08		0.02
Rec.:	111.91	0.12	-0.02	0.29	0.42	111.69	0.07	-0.24	0.27	0.56	111.49	0.09	-0.32	0.28	0.47	111.28	0.10	-0.41	0.28	0.15	111.08	0.14	-0.49	0.29	0.01
			0.04		0.54			0.29		0.63			0.38		0.66			0.46		0.69			0.55		0.61
10-YEAR TREASURY CONSTANT MATURITY RATE (GS10)																									
Full	1.98	0.14	0.89***	0.38	0.69***	2.02	0.16	0.81	0.33	1.04***	2.03	0.19	0.71***	0.32	1.08***	2.03	0.22	0.69***	0.32	1.08***	2.03	0.26	0.64***	0.33	1.06***
			0.09		0.09			0.08		0.08			0.08		0.05			0.10		0.21			0.09		0.05
Rec.:	2.58	0.07	0.73***	0.25	-0.08	2.61	0.11	0.72***	0.23	0.93***	2.65	0.16	0.56***	0.21	1.03***	2.67	0.19	0.49***	0.21	1.04***	2.67	0.24	0.44***	0.21	1.08***
			0.07		0.18			0.10		0.21			0.08		0.20			0.12		0.15			0.09		0.20
3-MONTH TREASURY BILL (TB3MS)																									
Full	2.01	0.32	0.96***	0.09	0.99***	2.00	0.33	1.07***	0.22	0.99***	2.00	0.37	1.07***	0.25	1.02***	2.00	0.46	1.05***	0.28	1.08***	2.00	0.53	1.24***	0.33	1.15***
			0.06		0.05			0.03		0.01			0.04		0.02			0.04		0.01			0.06		0.06
Rec.:	2.83	0.36	0.87***	0.09	1.05***	2.84	0.33	1.00	0.28	1.02***	2.83	0.32	0.99***	0.28	1.06***	2.86	0.35	0.98***	0.26	1.14***	2.84	0.40	1.15***	0.27	1.24***
			0.06		0.04			0.03		0.03			0.04		0.05			0.03		0.04			0.06		0.08

Notes: The table comprises the forecasting performance results of 15 macroeconomic variables defined in Appendix 1 in which the first column indicates whether one is taking the full-sample period (1981:3-2009:2) (“Full”) or the aggregation of recession periods (“Rec”) – simple combination of the 4 main recessions as identified by the NBER. Variables names are in accordance to the Philadelphia Fed’s designation. The remaining of the table is organized as follows: row 2 identifies the specification; row 3 indicates the forecast horizon ranging from h=1 to h=5. Columns 1,6,11,16 and 21 report the Root Mean Squared Forecast Error (RMSFE) associated to using the real-time average as forecast for each vintage available at the forecast date (available at <http://www.phil.frb.org/research-and-data/real-time-center/real-time-data/data-files/>) at different time horizons. Columns 2, 7, 12, 17 and 22 report the relative RMSFE obtained from an Autoregressive (AR) equivalent process to the real-time RMSFE – a value larger than one indicates that the AR’s RMSE is larger than the real-time RMSFE and this is emphasized by a grey shade throughout the table’s entries. Columns 4, 9, 14, 19 and 24 report the relative RMSFE obtained from using the Survey of Professional Forecasters’ (SPF) predictions to the Real Time RMSFE – a value larger than one indicates that the SPF’s RMSFE is larger than the real-time RMSFE and this is emphasized by a grey shade throughout the table’s entries. Columns 3,8,13,18 and 23 (5,10,15,20 and 25) report the coefficient  $\beta_1$  resulting from OLS estimation of Equation (3) for forecast accuracy at different time horizons for the comparison of real-time and the AR process (SPF); Newey-West robust standard errors are presented below each coefficient entry and \*, \*\*, \*\*\* denote significance at 10,5 and 1% levels, respectively (i.e., rejection of the null hypothesis that the coefficient  $\beta_1$  is zero).

**TABLE 1.2: Forecast Performance of the Fed-Green Book (PARTIAL TABLE – full table available upon request)**

REAL GNP/GDP (ROUTPUT)																									
	Real Time Average Forecast (RMSFE)	Rel. RMSE AR	$\beta_1$ AR - OLS	Rel. RMSE Fed	$\beta_1$ SPF - OLS	Real Time Average Forecast (RMSFE)	Rel. RMSE AR	$\beta_1$ AR - OLS	Rel. RMSE Fed	$\beta_1$ SPF - OLS	Real Time Average Forecast (RMSFE)	Rel. RMSE AR	$\beta_1$ AR - OLS	Rel. RMSE Fed	$\beta_1$ SPF - OLS	Real Time Average Forecast (RMSFE)	Rel. RMSE AR	$\beta_1$ AR - OLS	Rel. RMSE Fed	$\beta_1$ SPF - OLS	Real Time Average Forecast (RMSFE)	Rel. RMSE AR	$\beta_1$ AR - OLS	Rel. RMSE Fed	$\beta_1$ SPF - OLS
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)	(21)	(22)	(23)	(24)	(25)
Time/H	1					2					3					4					5				
Full	2.55	1.02	0.04 0.12	0.82	0.08* 0.05	2.57	1.03	0.59 0.62	0.95	0.42 0.29	2.57	1.07	0.35 0.30	1.08	0.29 0.34	2.57	1.09	0.13 0.16	1.13	0.40 0.33	2.56	1.13	-0.34 0.36	1.15	0.21 0.40
Rec.:	3.49	1.09	0.02 0.17	0.63	0.01 0.10	3.55	1.10	-1.30 1.72	0.91	0.42 0.53	3.62	1.12	0.70 0.54	1.10	-0.19 0.75	3.68	1.11	0.01 0.40	1.16	-0.07 0.66	3.75	1.08	0.69 1.01	1.13	-0.35 0.62
PRICE INDEX FOR GNP/GDP (GDPDEF)																									
Full	1.70	1.66	0.91*** 0.25	0.83	0.74*** 0.13	1.70	1.71	0.08 0.20	0.81	0.66*** 0.17	1.70	1.72	0.23 0.21	0.83	0.66*** 0.13	1.70	1.81	-0.08 0.16	0.83	0.56*** 0.12	1.70	1.88	0.01 0.11	0.88	0.56*** 0.18
Rec.:	1.90	1.92	0.70 0.44	0.96	0.64*** 0.21	1.90	1.97	-0.17 0.48	0.91	0.54** 0.24	1.91	1.98	0.41 0.63	0.93	0.75*** 0.24	1.91	2.09	0.16 0.31	0.95	0.72** 0.32	1.91	2.07	-0.04 0.23	1.21	1.14** 0.41

Notes: Vide Table 1.1 for the full set of explanations. As for the Fed-Green Book, the sample runs only up to 2003:4, henceforth the last recession (2008:1) is not considered when doing the aggregation for “Rec” (hence, we are left with even fewer observations). Moreover, due to few observations in the decade starting in 2000, the  $\beta_1$  coefficient is not computed. The averaged real-time vintage data used here is the same as before. For further information on the data/variables refer to Appendix 1.

TABLE 2. Sign forecast accuracy tests

## PHIL-SPF and Real-Time average

H	Variable	Correct				Incorrect				p-value	
		$\Delta Y_{t+h}(+), \Delta \widehat{Y}_{t+h}(+), n_{11}$		$\Delta Y_{t+h}(-,0), \Delta \widehat{Y}_{t+h}(-,0), n_{22}$		$\Delta Y_{t+h}(-,0), \Delta \widehat{Y}_{t+h}(+), n_{12}$		$\Delta Y_{t+h}(+), \Delta \widehat{Y}_{t+h}(-,0), n_{21}$			
		Phil-SPF	Real-Time	Phil-SPF	Real-Time	Phil-SPF	Real-Time	Phil-SPF	Real-Time	Phil-SPF	Real-Time
1	NOUTPUT	24	20	15	25	29	33	41	31	0.00	0.03
2		21	20	14	27	29	30	44	31	0.00	0.06
3		44	35	11	21	15	24	37	27	0.17	0.15
4		36	27	17	30	13	22	40	27	0.16	0.11
5		37	30	10	25	16	23	42	27	0.08	0.14
1	IPT	24	21	32	40	24	27	29	21	0.15	0.09
2		22	16	26	41	29	31	31	20	0.08	0.16
3		37	18	23	37	16	27	31	25	0.07	0.16
4		41	14	14	39	13	30	38	23	0.17	0.14
5		38	15	15	40	13	27	39	23	0.17	0.16
1	HSTARTS	24	26	14	21	34	32	37	30	0.00	0.05
2		18	19	14	21	33	32	43	36	0.00	0.00
3		35	34	12	18	21	22	39	33	0.05	0.15
4		30	28	15	23	12	14	49	41	0.15	0.16
5		31	32	13	19	22	21	39	33	0.03	0.15
1	RCONS	31	13	9	37	29	44	40	15	0.00	0.13
2		24	23	13	37	24	33	47	15	0.00	0.07
3		26	21	10	34	26	34	45	18	0.00	0.15
4		29	16	11	45	17	37	49	8	0.02	0.03
5		24	21	12	39	22	31	47	14	0.00	0.05
1	RINVBFB	20	25	15	23	30	25	44	36	0.00	0.08
2		20	25	13	22	31	26	44	35	0.00	0.07
3		32	26	11	21	22	28	42	32	0.01	0.07
4		34	25	8	21	19	28	45	32	0.01	0.06
5		33	23	6	23	18	28	48	31	0.00	0.07
1	RINVRESID	33	37	13	14	20	16	43	42	0.04	0.14
2		32	39	11	15	21	14	44	40	0.01	0.17
3		23	14	10	12	28	37	45	43	0.00	0.00
4		19	11	10	16	29	37	48	42	0.00	0.00
5		15	7	12	21	26	34	52	43	0.00	0.00
1	RGF	21	14	10	39	37	44	40	12	0.00	0.18
2		20	11	12	39	35	44	40	14	0.00	0.13
3		43	43	11	37	12	12	41	15	0.19	0.00
4		35	42	15	37	23	16	33	11	0.11	0.00
5		40	36	11	43	9	13	45	13	0.19	0.00
1	RGLS	25	21	14	32	30	41	40	15	0.00	0.16
2		24	19	11	32	32	42	41	15	0.00	0.17
3		33	20	14	36	17	41	43	10	0.09	0.08
4		33	21	12	32	19	39	42	14	0.05	0.15
5		33	21	14	30	20	39	38	15	0.08	0.16
1	UNRATE	26	17	46	30	10	21	13	30	0.00	0.15
2		26	19	51	36	10	20	17	32	0.00	0.16
3		30	18	43	37	6	18	27	33	0.00	0.16
4		29	17	34	36	5	16	37	35	0.00	0.16
5		26	19	29	34	7	14	42	37	0.03	0.15
1	NETEXP	28	26	21	40	21	23	39	20	0.11	0.02
2		28	23	20	42	19	24	41	19	0.11	0.03
3		28	20	21	45	16	24	42	18	0.15	0.03
4		27	24	22	41	18	21	39	20	0.15	0.02
5		27	21	24	44	15	21	39	19	0.16	0.02
1	CPI	11	18	18	23	42	35	38	33	0.00	0.01
2		12	18	17	21	42	36	37	33	0.00	0.00
3		38	30	23	26	18	26	28	25	0.06	0.14
4		41	29	22	28	10	22	33	27	0.01	0.11
5		39	34	23	22	13	18	30	31	0.02	0.12
1	TB3MS	23	58	48	3	23	46	12	0	0.00	0.09
2		25	58	47	3	23	48	13	0	0.00	0.09
3		38	59	42	3	8	46	19	0	0.00	0.09
4		35	63	42	2	7	42	21	0	0.00	0.17
5		36	63	47	2	5	41	17	0	0.00	0.16
1	GS10	45	61	62	2	1	47	1	0	0.00	0.19
2		40	65	66	2	1	42	1	0	0.00	0.16
3		41	65	66	2	0	41	0	0	0.00	0.16
4		39	66	66	2	0	39	1	0	0.00	0.15
5		42	62	62	2	0	42	1	0	0.00	0.17

H	Variable	Correct						Incorrect						p-value		
		$\Delta Y_{t+h}(+), \Delta \widehat{Y}_{t+h}(+), n_{11}$			$\Delta Y_{t+h}(-,0), \Delta \widehat{Y}_{t+h}(-,0), n_{22}$			$\Delta Y_{t+h}(-,0), \Delta \widehat{Y}_{t+h}(+), n_{12}$			$\Delta Y_{t+h}(+), \Delta \widehat{Y}_{t+h}(-,0), n_{21}$					
		SPF	Real	Green	SPF	Real	Green	SPF	Real	Green	SPF	Real	Green	SPF	Real	Green
1	GDPDEF	33	32	15	9	25	10	19	46	27	48	6	35	0.01	0.02	0.00
2		31	30	15	10	24	9	23	47	28	44	7	34	0.00	0.05	0.00

3		32	34	21	20	22	12	19	44	21	36	7	31	0.16	0.03	0.02
4		36	35	22	20	21	11	14	43	18	36	7	33	0.12	0.03	0.03
5		33	37	21	19	20	10	15	41	13	38	7	32	0.16	0.03	0.08
1	ROUTPUT	23	19	19	16	38	10	28	37	22	42	15	36	0.00	0.14	0.00
2		22	17	20	13	39	10	30	38	22	43	14	34	0.00	0.15	0.00
3		37	15	29	12	39	3	16	38	17	42	15	36	0.11	0.17	0.00
4		32	18	24	17	40	7	17	33	17	40	15	36	0.14	0.11	0.01
5		36	16	24	12	37	4	17	34	17	40	18	31	0.10	0.17	0.00

Notes: The first column indicates the number of quarters associated to the forecast horizon.  $\Delta Y_{t+h} = Y_{t+h} - Y_{t-1}$  is the actual change, and

$\Delta \hat{Y}_{t+h} = \hat{Y}_{t+h} - Y_{t-1}$  is the predicted change in each variable.  $Y_{t-1}$  is the most recent (quarterly) rate known at the time of the forecast. (+) indicates that  $\Delta Y_{t+h} (\Delta \hat{Y}_{t+h})$  is positive; (-,0) indicates that  $\Delta Y_{t+h} (\Delta \hat{Y}_{t+h})$  is negative or zero. P-value (or Fisher's exact test) is for testing the null hypothesis of no association between the direction of change in the actual and forecast series. For both the ROUTPUT and GDP deflator variables (last two considered above) we subdivide the each column into an additional one as to incorporate the Fed-Green Book. Recall that the Fed-Green Book's and Phil-SPF's samples are different in size.

**TABLE 3. Comparing Predictive information content**

H	Variable	p-value for testing the null hypothesis that	
		Green Book encompasses SPF	SPF encompasses Green Book
1	ROUTPUT	0.00	0.77
2		0.00	0.09
3		0.00	0.11
4		0.00	0.21
5		0.00	0.21
1	GDP Deflator	0.00	0.00
2		0.01	0.97
3		0.11	0.24
4		0.31	0.08
5		0.39	0.08

Notes: The first column indicates the number of quarters associated to the forecast horizon. A grey shade indicates a P-value smaller than 10%.

**TABLE 4. MSE-ratio Comparison during Recessions: Fed-Green Book vs. Phil--SPF**

h	Full Sample		Up to 2001:4		Up to 1991:4		Full Sample		Up to 2001:4		Up to 1991:4		Full Recessions	
	MSE- Ratio	DB	MSE- Ratio	DB	MSE- Ratio	DB	MSE- Ratio_w	DB	MSE- Ratio_w	DB	MSE- Ratio_w	DB	MSE- Ratio_r	DB
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
<b>ROUTPUT</b>														
1	1.51	0.06	1.46	0.04	1.61	0.13	1.38	0.05	1.22	0.08	1.27	0.18	1.74	0.06
2	1.43	0.00	1.35	0.00	1.40	0.01	1.49	0.00	1.31	0.00	1.23	0.00	1.38	0.15
3	1.12	0.19	1.08	0.26	1.05	0.69	1.32	0.04	1.13	0.08	1.02	0.61	0.98	0.56
4	0.99	0.91	0.97	0.81	0.90	0.55	1.32	0.04	1.12	0.02	1.05	0.10	0.79	0.50
5	0.94	1.00	0.95	0.88	0.85	0.22	1.22	0.22	1.03	0.34	0.87	0.33	0.77	0.43
<b>GDP Deflator</b>														
1	1.25	0.01	1.17	0.02	1.19	0.00	1.13	0.14	1.22	0.29	1.35	0.04	1.35	0.00
2	1.07	0.62	0.99	0.84	0.96	0.77	1.12	0.15	1.22	0.81	1.25	0.88	1.03	0.91
3	1.03	0.83	0.92	0.60	0.83	0.13	1.09	0.18	1.08	0.73	0.99	0.32	0.98	0.57
4	0.98	0.57	0.88	0.44	0.80	0.15	1.04	0.23	0.99	0.41	0.87	0.14	0.93	0.65
5	1.04	0.87	0.92	0.72	0.93	0.12	1.07	0.71	0.98	0.30	1.02	0.00	1.02	0.00

Note: MSE-Ratio is the ratio of the MSE of Fed-Greenbook forecasts to Phil-SPF forecasts. MSE-Ratio\_w refers to the MSE-Ratio without recession periods. We present the full sample as well as two sub-periods up until 2001:4 and 2001:4, respectively, with and without recessions. Columns 14-15 present the same information for the aggregation of recessions only. DB refers to the modified Diebold-Mariano p-values and it tests whether Fed-Greenbook forecasts are significantly better than the Phil-SPF forecasts. P-values in excess of 10% mean the non-rejection of the null that both forecasts are not significantly different from each other.

**TABLE 5. RMSFE comparison between the SPF's predictions vis-a-vis the theoretical and empirical NK and RBC models' predictions****PANEL A - THEORETICAL**

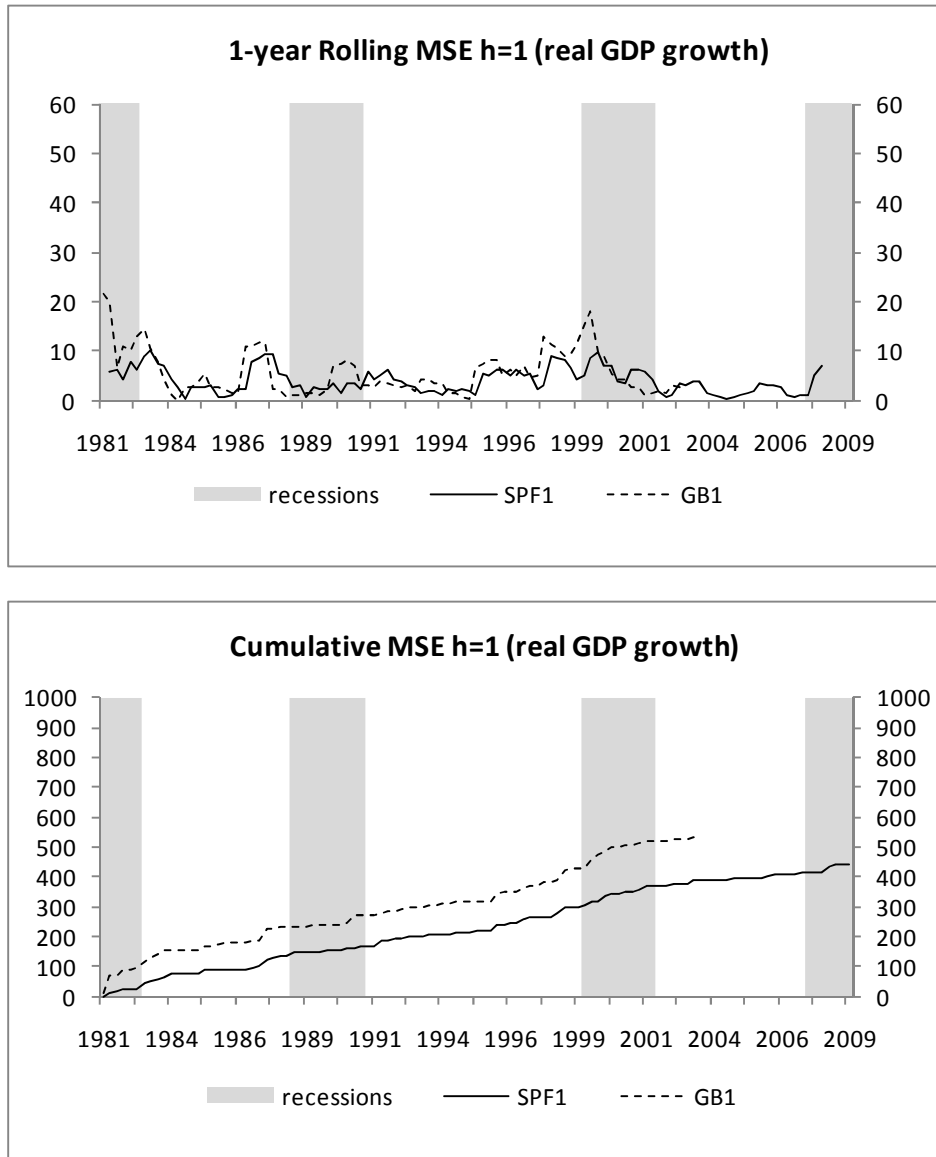
Variables	Model	horizon													
		1	2	3	4	5	6	8	10	12	14	16	18	20	
Real GDP growth	SPF	0.66	0.78	0.91	0.96	0.98	-	-	-	-	-	-	-	-	
	RBC	0.79	0.92	0.96	0.96	0.97	0.97	0.97	0.98	0.98	0.98	0.99	0.99	0.99	
	NK	0.55	0.69	0.78	0.85	0.89	0.92	0.96	0.97	0.98	0.98	0.99	0.99	1.00	
Consumption	SPF	0.68	0.74	0.78	0.84	0.83	-	-	-	-	-	-	-	-	
	RBC	0.64	0.83	0.92	0.94	0.94	0.95	0.96	0.97	0.98	0.99	0.99	0.99	0.99	
	NK	0.41	0.59	0.72	0.82	0.88	0.91	0.95	0.97	0.97	0.97	0.97	0.97	0.98	
RNR investment	SPF	0.65	0.73	0.79	0.90	0.94	-	-	-	-	-	-	-	-	
	RBC	0.63	0.82	0.86	0.87	0.88	0.89	0.93	0.96	0.99	1.00	1.01	1.02	1.02	
	NK	0.43	0.59	0.65	0.68	0.71	0.73	0.78	0.81	0.84	0.85	0.86	0.87	0.87	
Hours	SPF (Unemployment)	0.20	0.22	0.24	0.30	0.36	-	-	-	-	-	-	-	-	
	RBC	0.18	0.31	0.40	0.47	0.51	0.54	0.59	0.62	0.65	0.67	0.69	0.72	0.74	
	NK	0.01	0.02	0.03	0.05	0.07	0.10	0.16	0.24	0.31	0.39	0.46	0.53	0.59	
Wages	SPF	-	-	-	-	-	-	-	-	-	-	-	-	-	
	RBC	0.53	0.99	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	
	NK	0.83	0.88	0.89	0.90	0.90	0.90	0.91	0.92	0.92	0.93	0.94	0.94	0.95	
Inflation	SPF	0.67	0.70	0.72	0.75	0.77	-	-	-	-	-	-	-	-	
	Univariate (NK-based)	0.14	0.29	0.40	0.49	0.55	0.60	0.68	0.73	0.77	0.80	0.82	0.85	0.86	
	NK	0.13	0.27	0.38	0.46	0.53	0.58	0.66	0.72	0.76	0.79	0.82	0.84	0.86	
Interest Rates	SPF	0.09	0.22	0.25	0.28	0.33	-	-	-	-	-	-	-	-	
	Univariate (NK-based)	0.14	0.29	0.40	0.49	0.55	0.60	0.68	0.73	0.77	0.80	0.82	0.85	0.86	
	NK	0.04	0.08	0.13	0.18	0.22	0.26	0.33	0.39	0.44	0.48	0.52	0.55	0.58	

**PANEL B - EMPIRICAL**

Variables	Model	horizon													
		1	2	3	4	5	6	8	10	12	14	16	18	20	
Real GDP growth	SPF	0.66	0.78	0.91	0.96	0.98	-	-	-	-	-	-	-	-	
	RBC	0.84	0.86	0.99	1.02	1.13	1.16	1.20	1.23	1.23	1.20	1.18	1.14	1.13	
	NK	1.66	1.71	1.76	1.62	1.73	1.58	1.36	1.21	1.13	1.11	1.09	1.09	1.09	
Consumption	SPF	0.68	0.74	0.78	0.84	0.83	-	-	-	-	-	-	-	-	
	RBC	1.19	1.31	1.47	1.67	1.73	1.79	1.68	1.59	1.43	1.26	1.20	1.09	1.06	
	NK	2.00	2.05	1.76	1.42	1.36	1.28	1.02	0.95	0.90	0.91	0.93	0.95	1.00	
RNR investment	SPF	0.65	0.73	0.79	0.90	0.94	-	-	-	-	-	-	-	-	
	RBC	0.66	0.93	1.07	1.10	1.06	1.02	0.95	0.96	0.99	1.03	1.10	1.13	1.15	
	NK	1.23	2.15	2.75	2.96	2.91	2.76	2.22	1.84	1.50	1.31	1.23	1.18	1.17	
Hours	SPF (Unemployment)	0.20	0.22	0.24	0.30	0.36	-	-	-	-	-	-	-	-	
	RBC	0.96	0.96	0.97	0.98	0.98	0.99	1.02	1.02	1.06	1.09	1.14	1.17	1.19	
	NK	0.09	0.26	0.47	0.66	0.79	0.92	1.11	1.26	1.30	1.26	1.27	1.23	1.29	
Wages	SPF	-	-	-	-	-	-	-	-	-	-	-	-	-	
	RBC	0.98	0.99	1.01	1.02	1.02	1.03	1.04	1.02	1.05	1.06	1.09	1.12	1.14	
	NK	1.08	0.97	0.97	0.98	0.99	1.00	1.02	1.03	1.07	1.09	1.13	1.16	1.18	
Inflation	SPF	0.67	0.70	0.72	0.75	0.77	-	-	-	-	-	-	-	-	
	NK	4.13	4.63	4.71	4.33	4.00	3.70	3.06	2.38	2.22	2.19	2.13	2.21	2.22	
Interest Rates	SPF	0.09	0.22	0.25	0.28	0.33	-	-	-	-	-	-	-	-	
	NK	0.06	0.19	0.39	0.64	0.90	1.11	1.48	1.68	1.85	2.00	2.09	2.15	2.21	

Notes: This table presents the relative (to the standard deviation of the variables) RMSFE at different forecast horizons, from h=1,...,20 for the SPF, theoretical and empirical NK and RBC models. Panel A also includes the univariate version of the NK SW07 model for both inflation and interest rates.

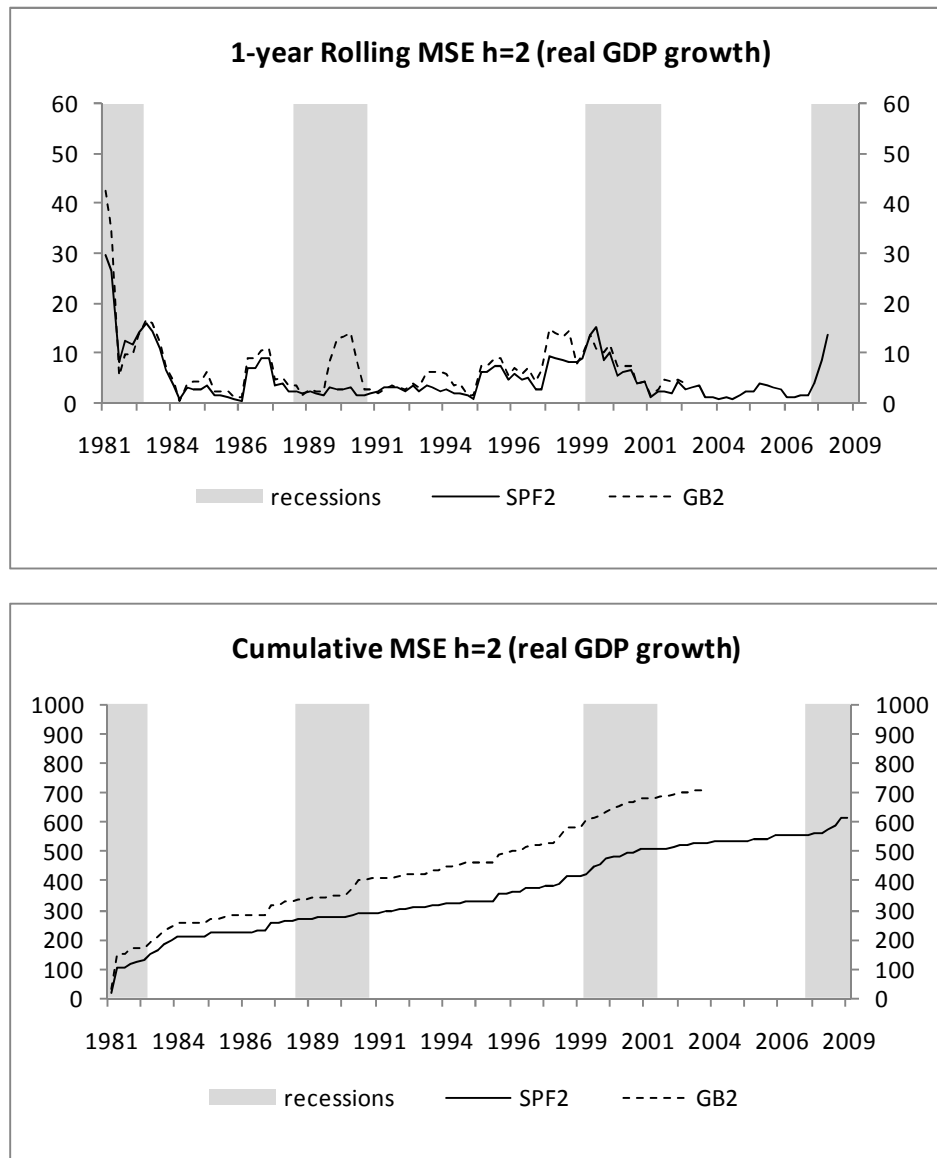
Figure 1.a



Source: Authors' calculations.

Note: This figure presents 1 year-rolling MSE and cumulative MSE for both the SPF's and Fed Green Book's predictions for real GDP growth for  $h=1$ . Shaded grey areas highlight NBER recession periods.

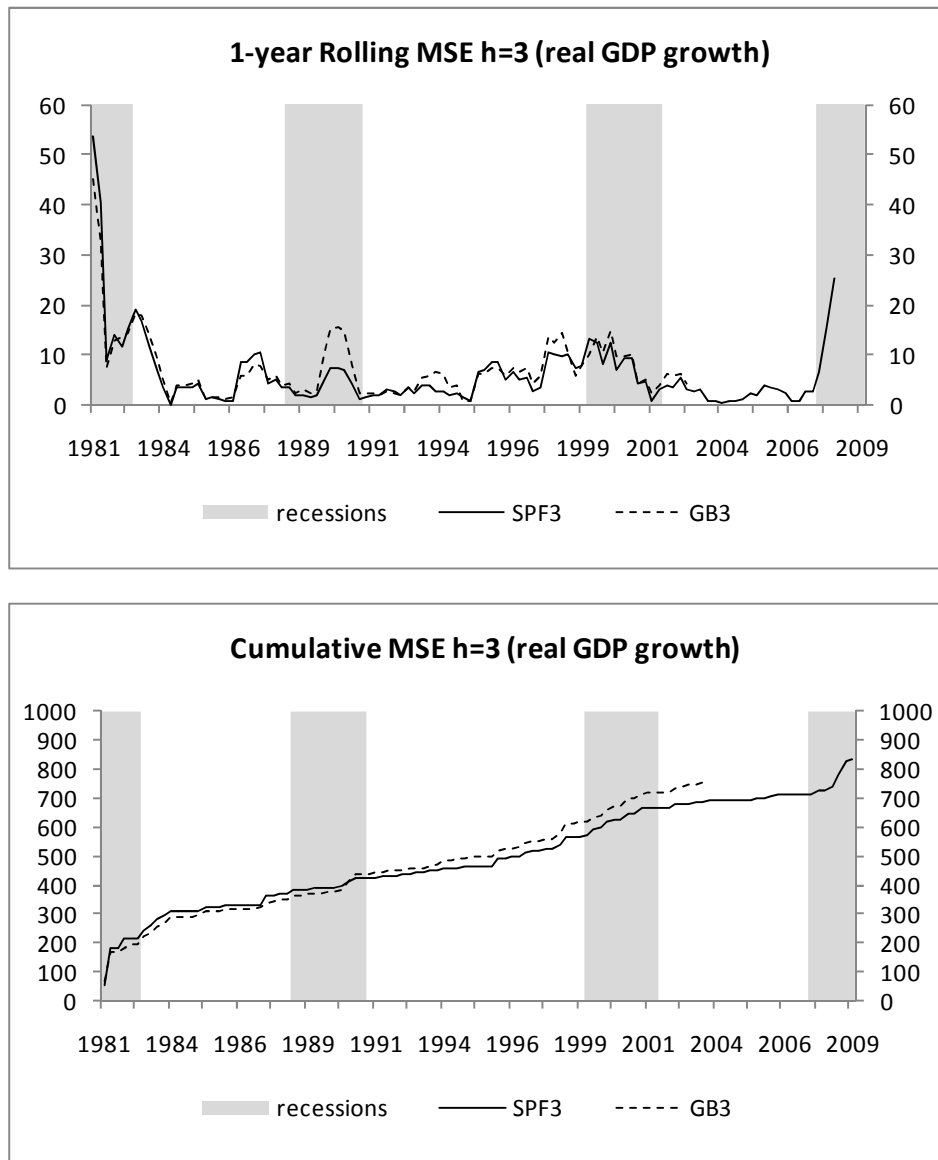
Figure 1.b



Source: Authors' calculations.

Note: This figure presents 1 year-rolling MSE and cumulative MSE for both the SPF's and Fed Green Book's predictions for real GDP growth for  $h=2$ . Shaded grey areas highlight NBER recession periods.

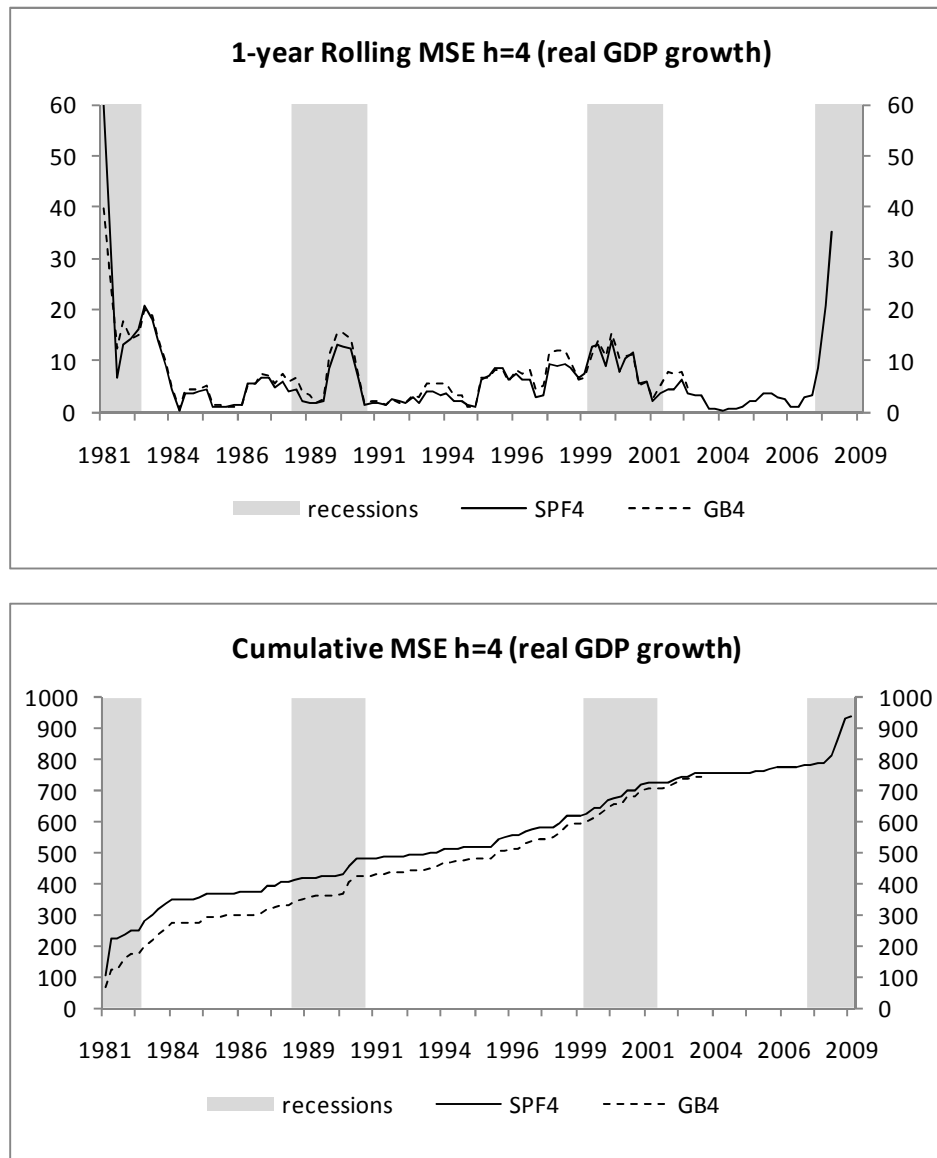
Figure 1.c



Source: Authors' calculations.

Note: This figure presents 1 year-rolling MSE and cumulative MSE for both the SPF's and Fed Green Book's predictions for real GDP growth for  $h=3$ . Shaded grey areas highlight NBER recession periods.

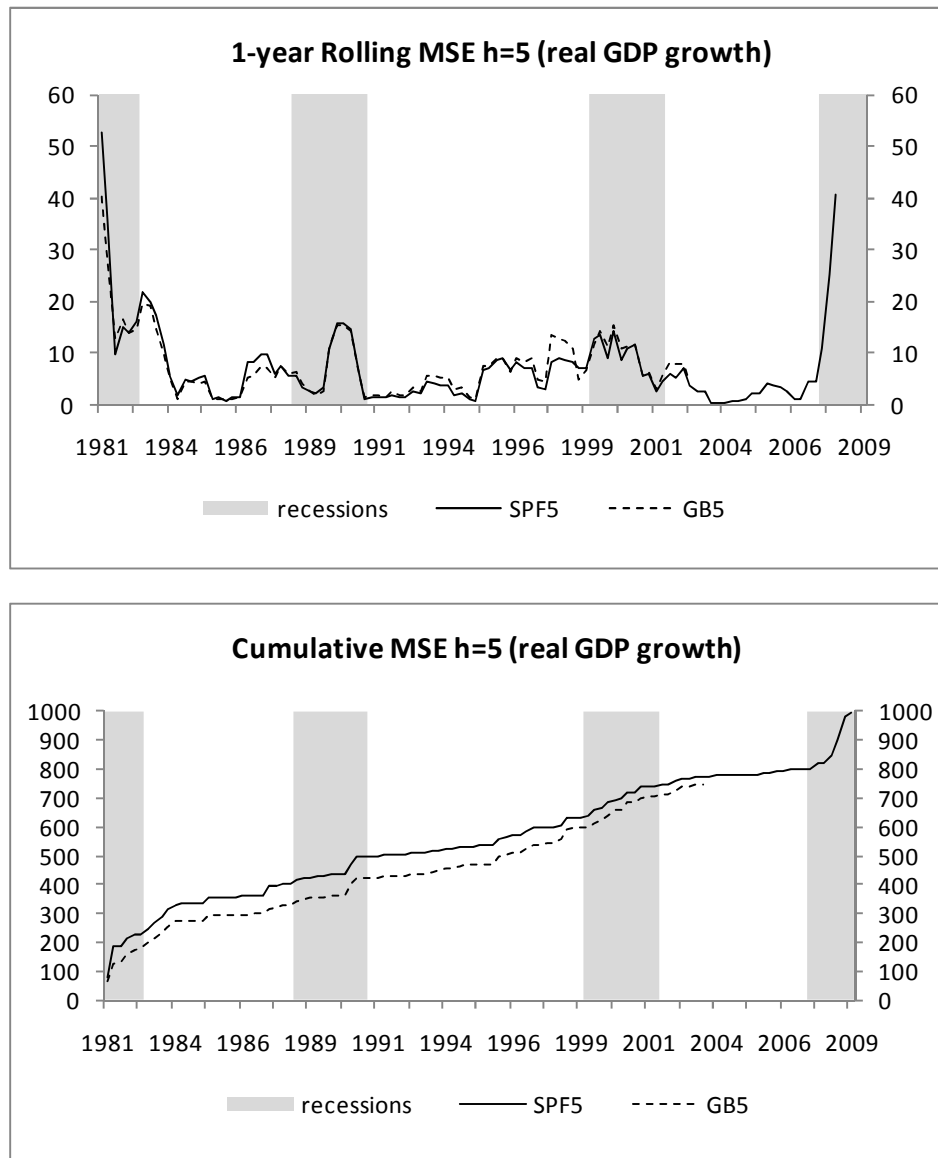
Figure 1.d



Source: Authors' calculations.

Note: This figure presents 1 year-rolling MSE and cumulative MSE for both the SPF's and Fed Green Book's predictions for real GDP growth for  $h=4$ . Shaded grey areas highlight NBER recession periods.

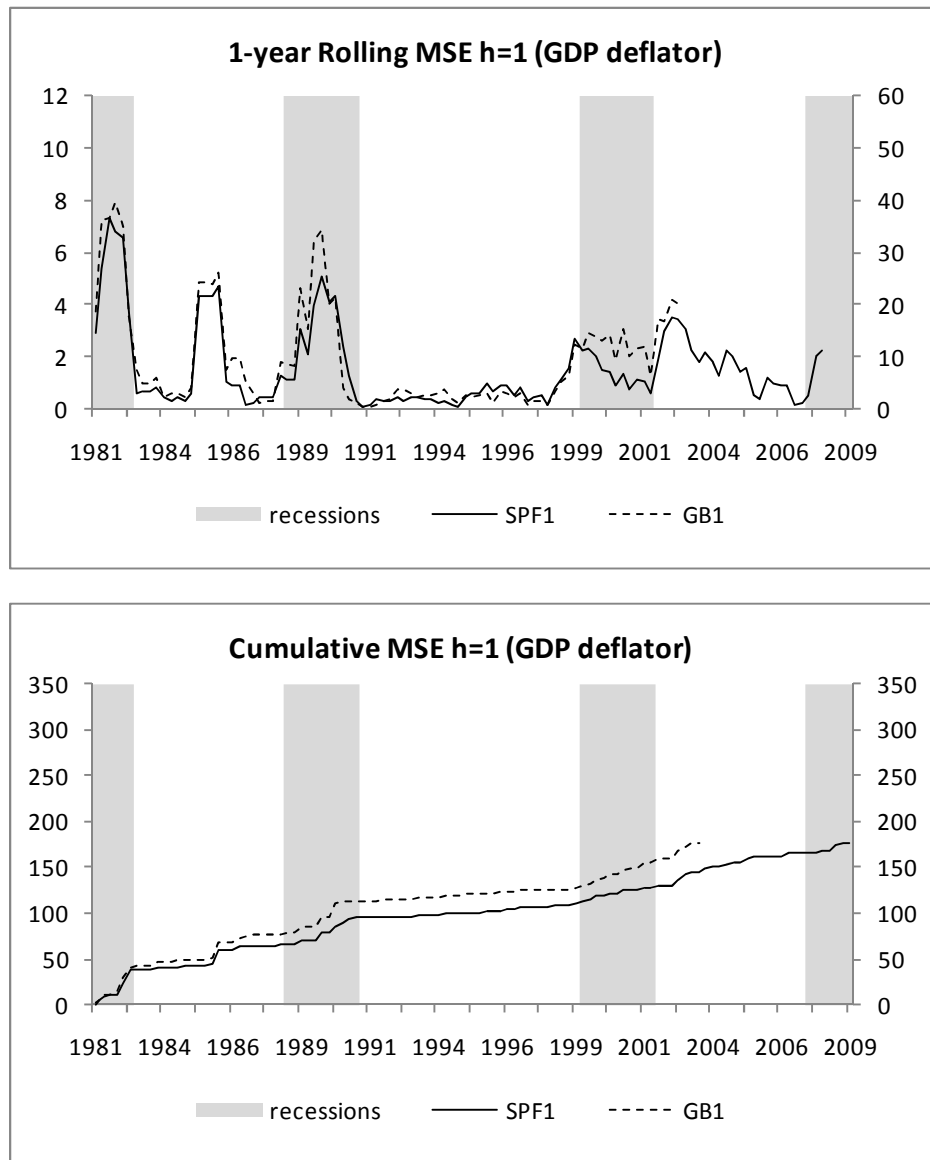
Figure 1.d



Source: Authors' calculations.

Note: This figure presents 1 year-rolling MSE and cumulative MSE for both the SPF's and Fed Green Book's predictions for real GDP growth for  $h=5$ . Shaded grey areas highlight NBER recession periods.

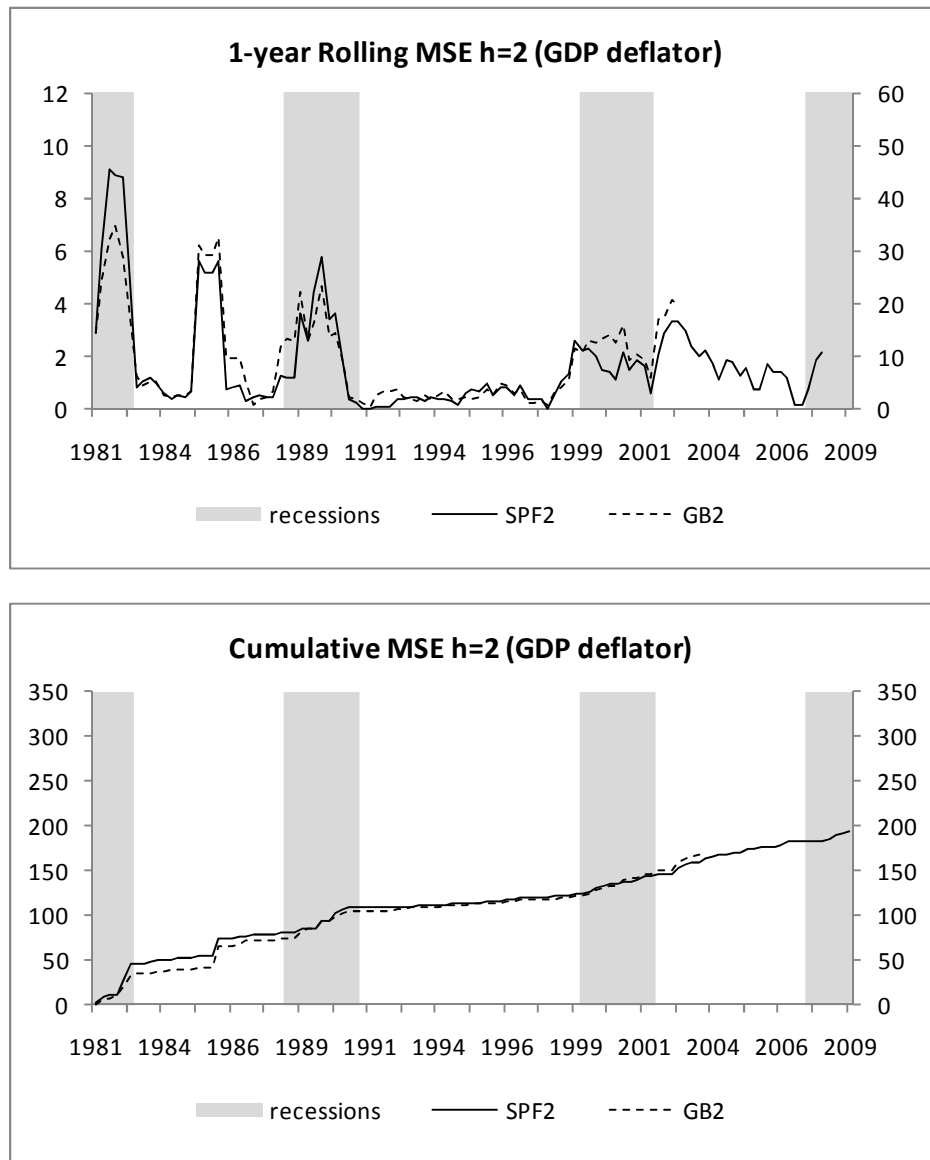
Figure 1.e



Source: Authors' calculations.

Note: This figure presents 1 year-rolling MSE and cumulative MSE for both the SPF's and Fed Green Book's predictions for the GDP deflator for  $h=1$ . Shaded grey areas highlight NBER recession periods.

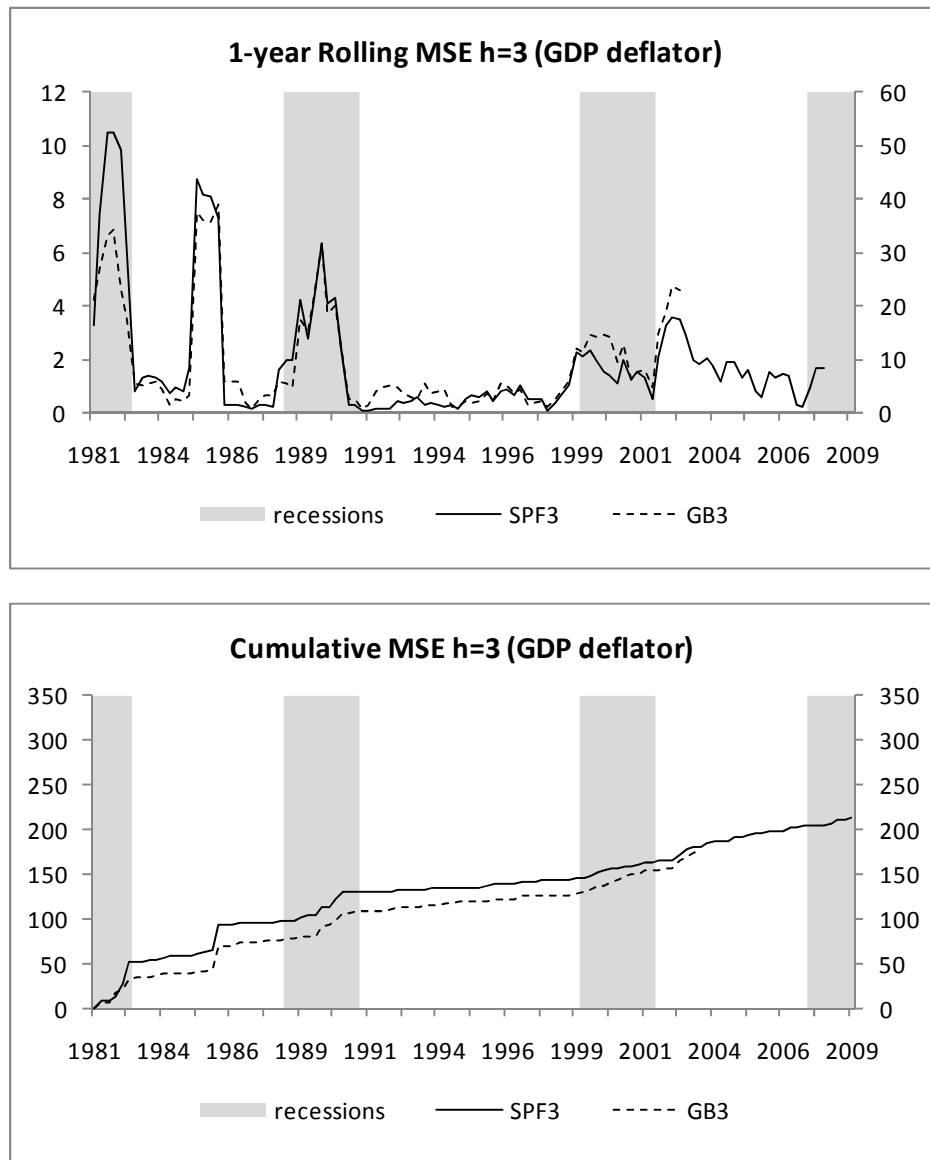
Figure 1.f



Source: Authors' calculations.

Note: This figure presents 1 year-rolling MSE and cumulative MSE for both the SPF's and Fed Green Book's predictions for the GDP deflator for  $h=2$ . Shaded grey areas highlight NBER recession periods.

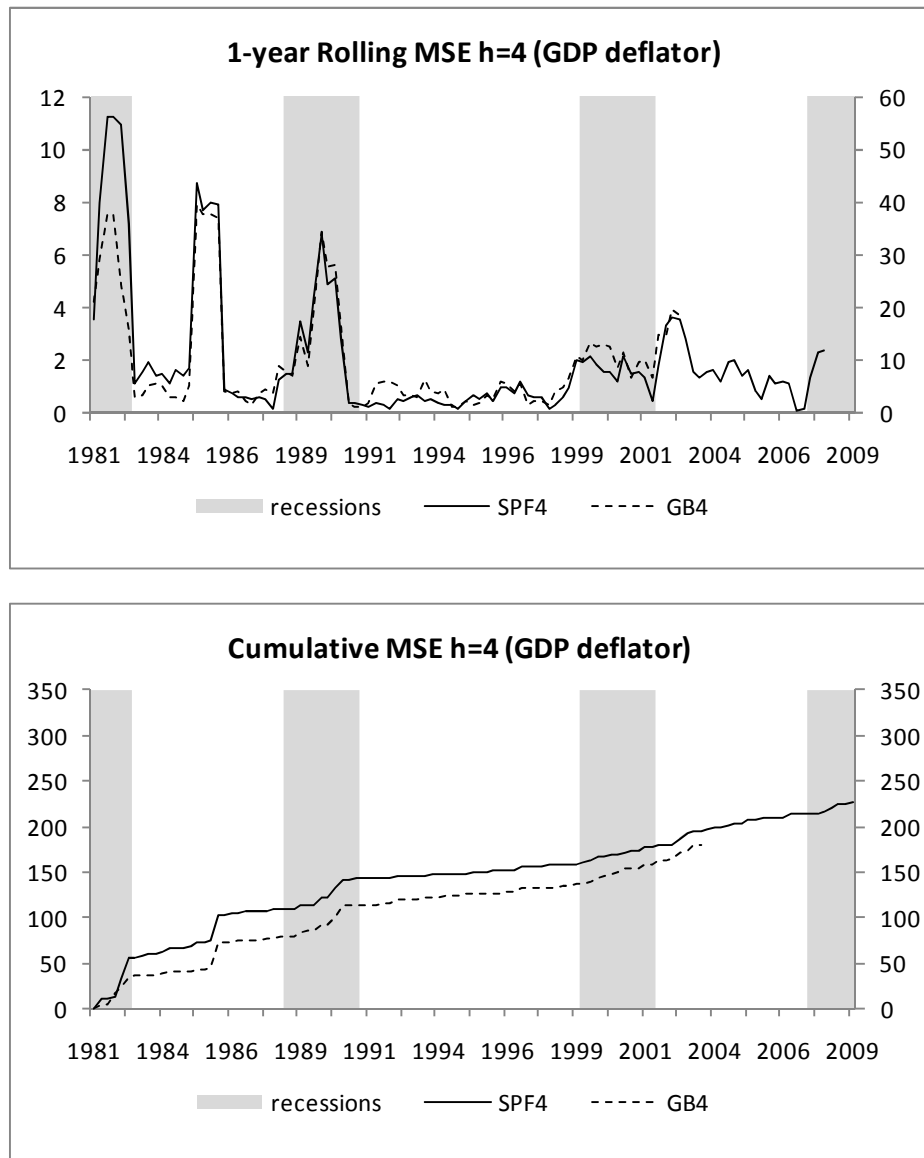
Figure 1.g



Source: Authors' calculations.

Note: This figure presents 1 year-rolling MSE and cumulative MSE for both the SPF's and Fed Green Book's predictions for the GDP deflator for  $h=3$ . Shaded grey areas highlight NBER recession periods.

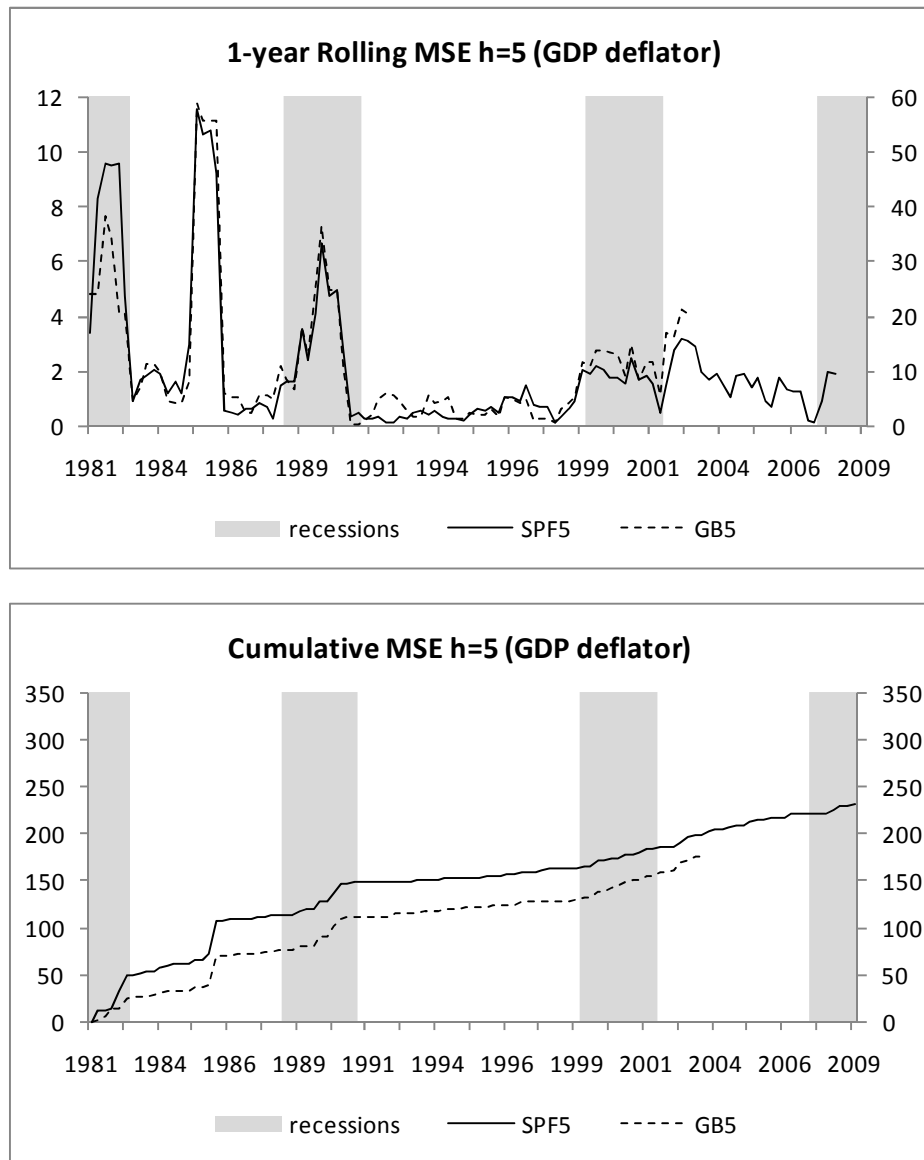
Figure 1.h



Source: Authors' calculations.

Note: This figure presents 1 year-rolling MSE and cumulative MSE for both the SPF's and Fed Green Book's predictions for the GDP deflator for  $h=4$ . Shaded grey areas highlight NBER recession periods.

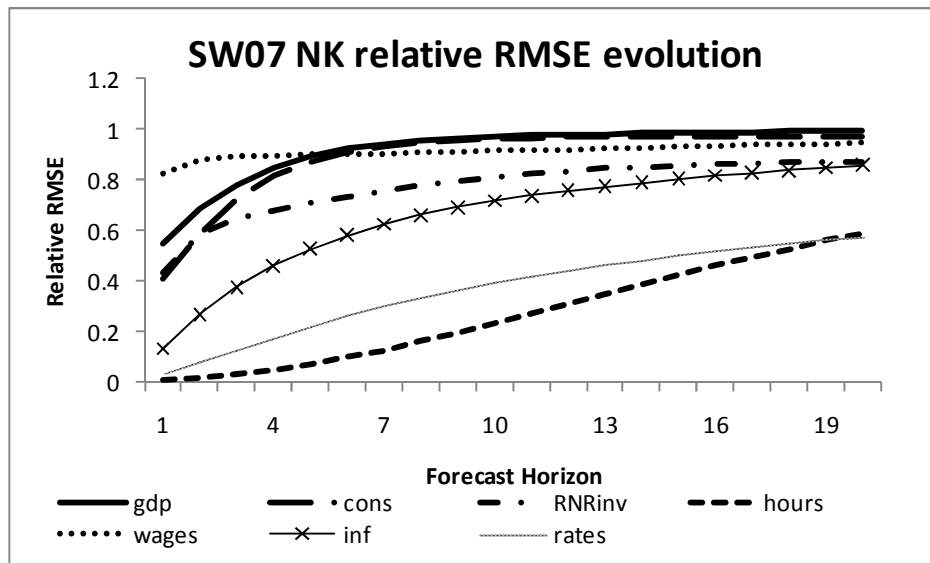
Figure 1.i



Source: Authors' calculations.

Note: This figure presents 1 year-rolling MSE and cumulative MSE for both the SPF's and Fed Green Book's predictions for the GDP deflator for  $h=5$ . Shaded grey areas highlight NBER recession periods.

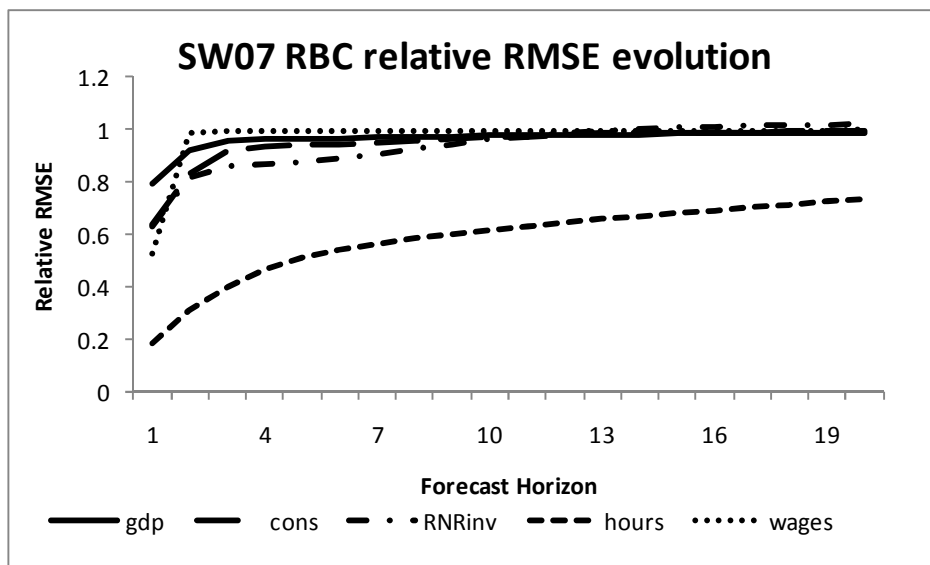
Figure 2



Source: Authors' calculations.

Note: This figure presents the relative RMSFE for each macro variable of interest at different forecast horizons (up to 20).

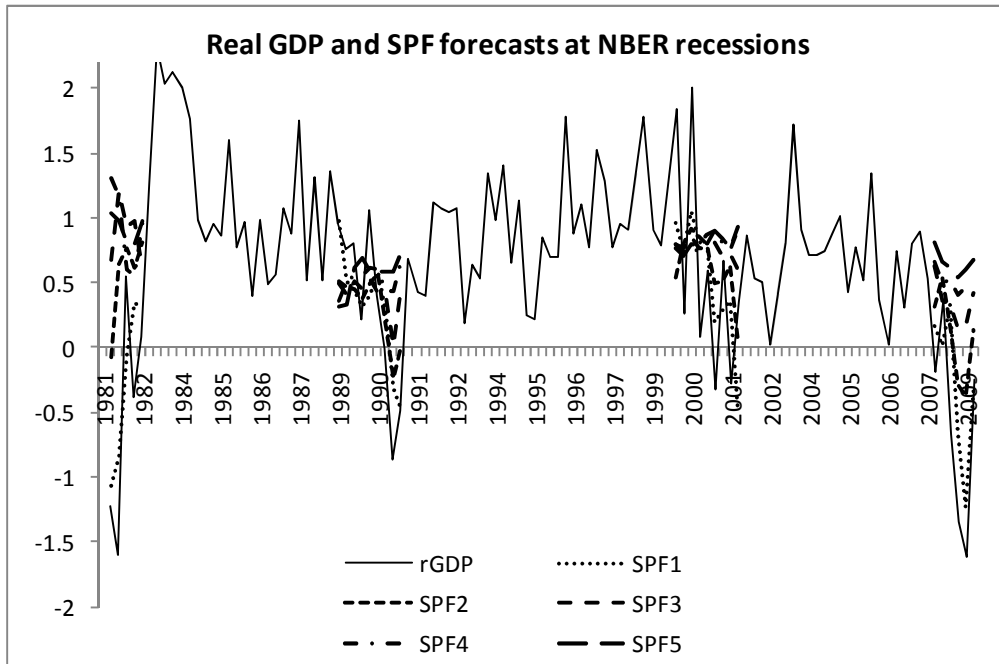
Figures 3



Source: Authors' calculations.

Note: This figure presents the relative RMSFE for each macro variable of interest at different forecast horizons (up to 20).

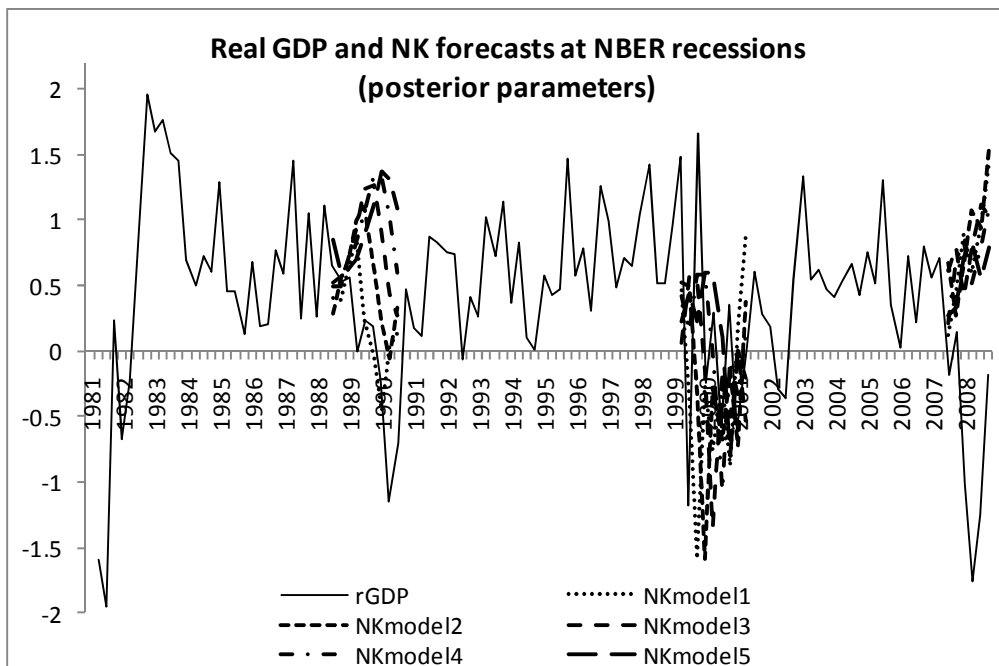
Figure 4.a



Source: Authors' calculations.

Note: This figure presents both the actual realization and predicted SPF values for real GDP growth between 1981 and 2009, the latter set plotting only observations at different horizons (1 to 5) for recession periods as identified by the NBER.

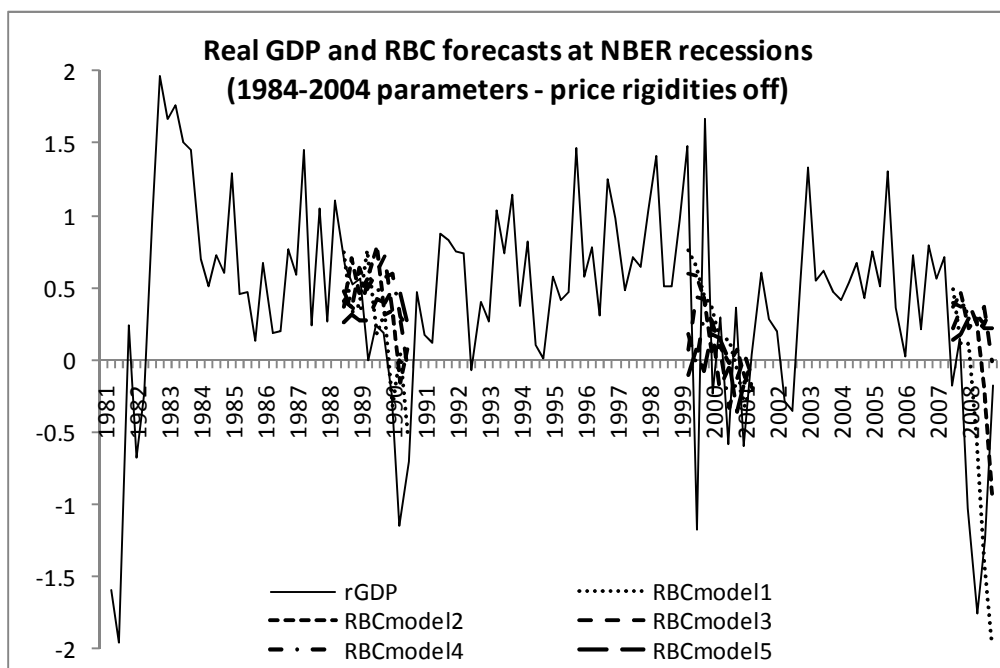
Figure 4.b



Source: Authors' calculations.

Note: This figure presents both the actual realization and predicted NK values for real GDP growth between 1981 and 2009, the latter set plotting only observations at different horizons (1 to 5) for recession periods as identified by the NBER.

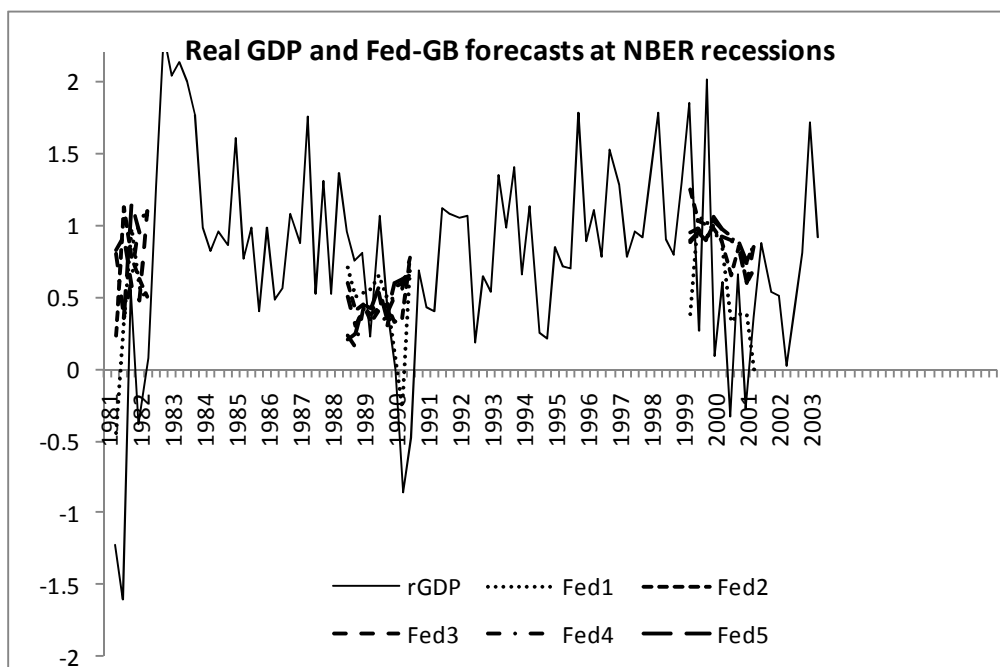
Figure 4.c



Source: Authors' calculations.

Note: This figure presents both the actual realization and predicted RBC values for real GDP growth between 1981 and 2009, the latter set plotting only observations at different horizons (1 to 5) for recession periods as identified by the NBER.

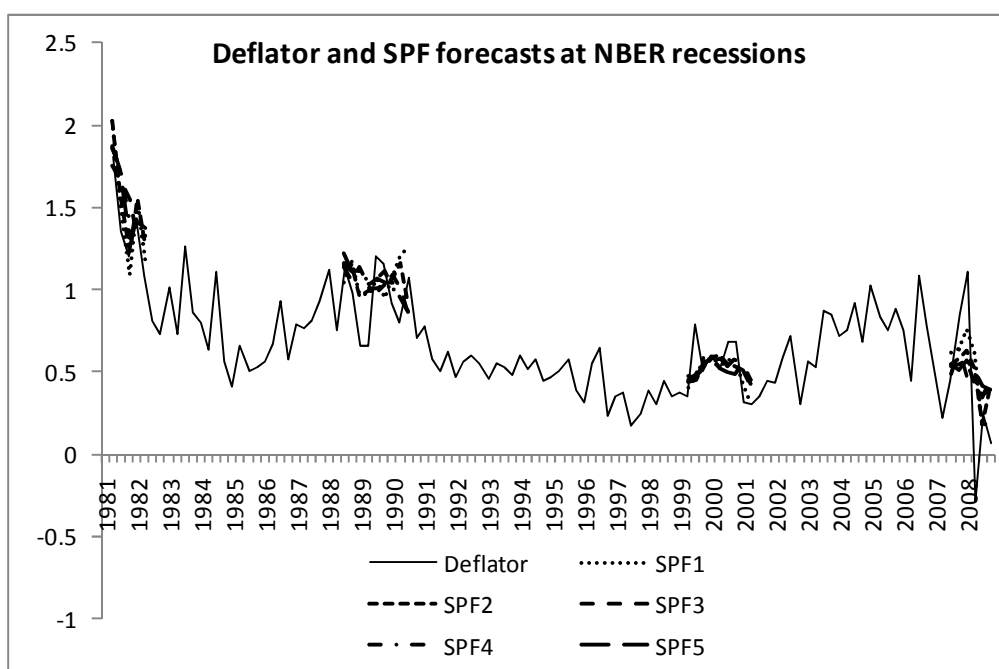
Figure 4.d



Source: Authors' calculations.

Note: This figure presents both the actual realization and predicted Fed Green Book values for real GDP growth between 1981 and 2009, the latter set plotting only observations at different horizons (1 to 5) for recession periods as identified by the NBER.

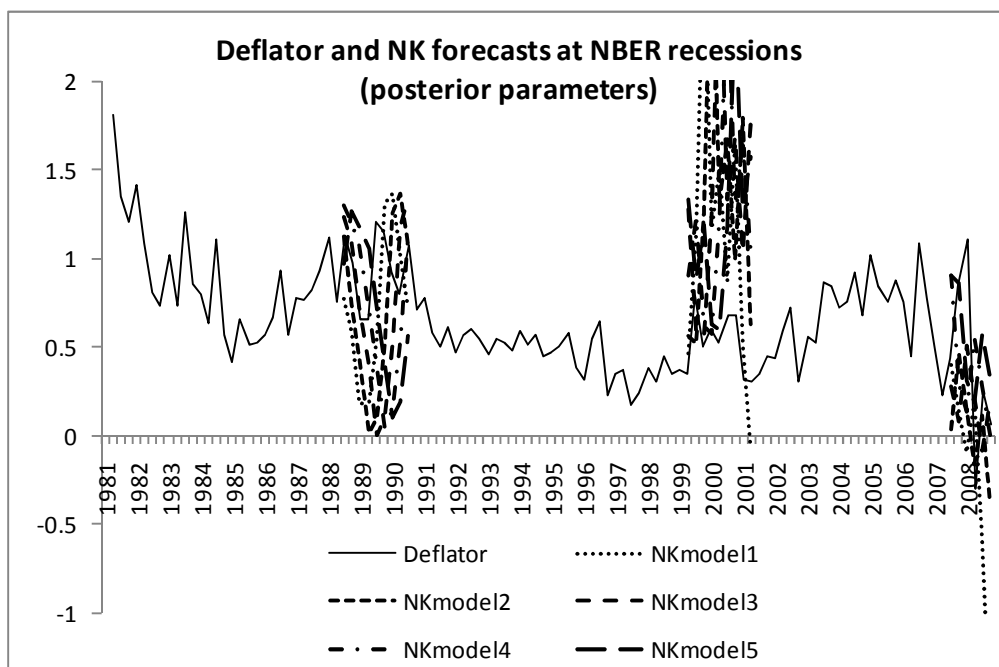
Figure 4.e



Source: Authors' calculations.

Note: This figure presents both the actual realization and predicted SPF values for the GDP-deflator between 1981 and 2009, the latter set plotting only observations at different horizons (1 to 5) for recession periods as identified by the NBER.

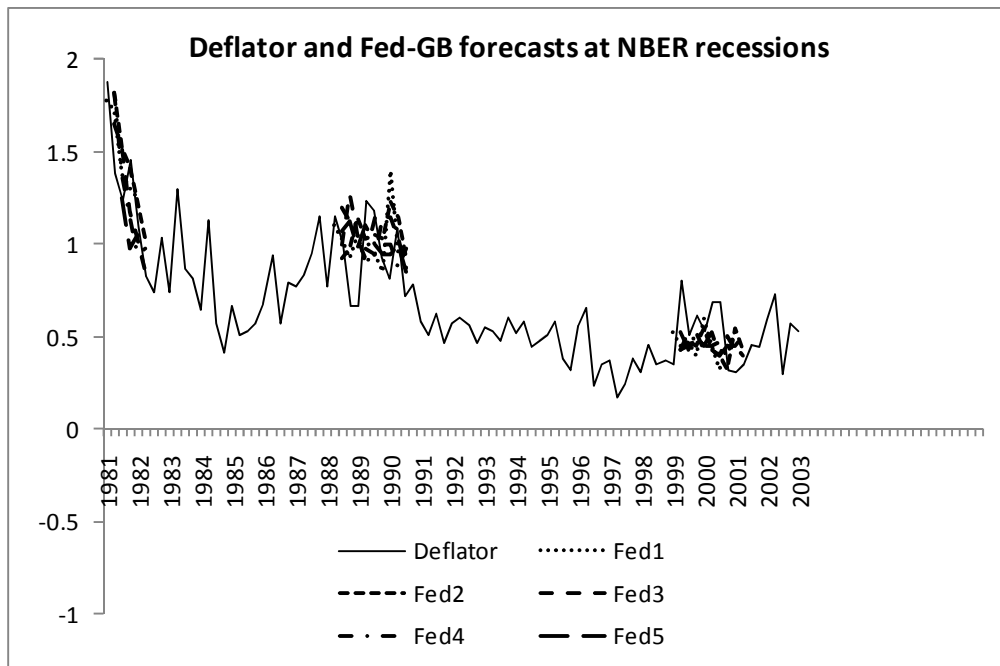
Figure 4.f



Source: Authors' calculations.

Note: This figure presents both the actual realization and predicted NK values the GDP-deflator between 1981 and 2009, the latter set plotting only observations at different horizons (1 to 5) for recession periods as identified by the NBER.

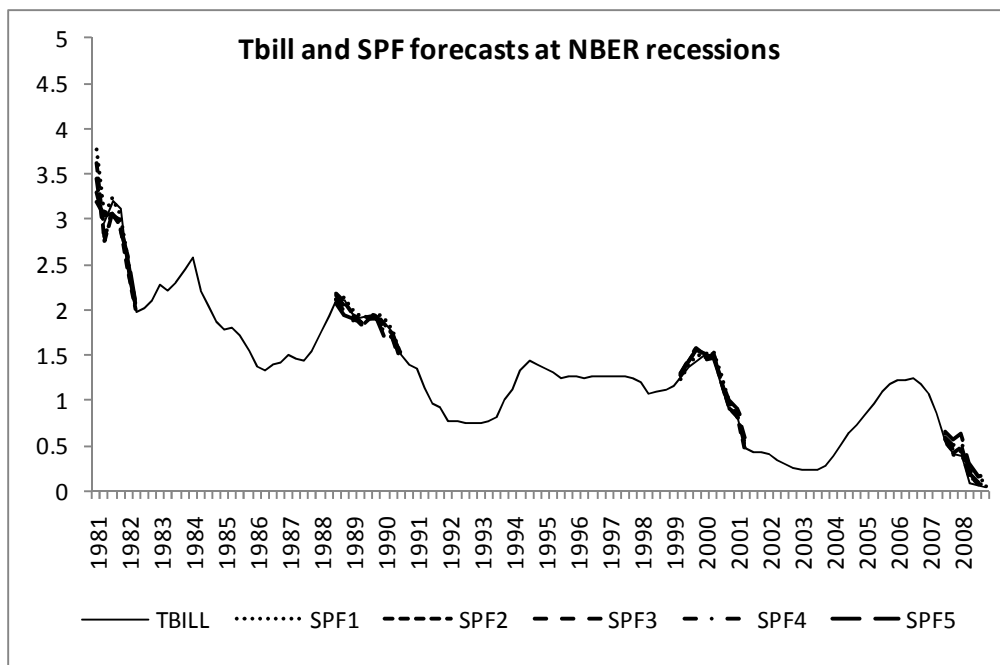
Figure 4.g



Source: Authors' calculations.

Note: This figure presents both the actual realization and predicted Fed Green Book values the GDP-deflator between 1981 and 2003, the latter set plotting only observations at different horizons (1 to 5) for recession periods as identified by the NBER.

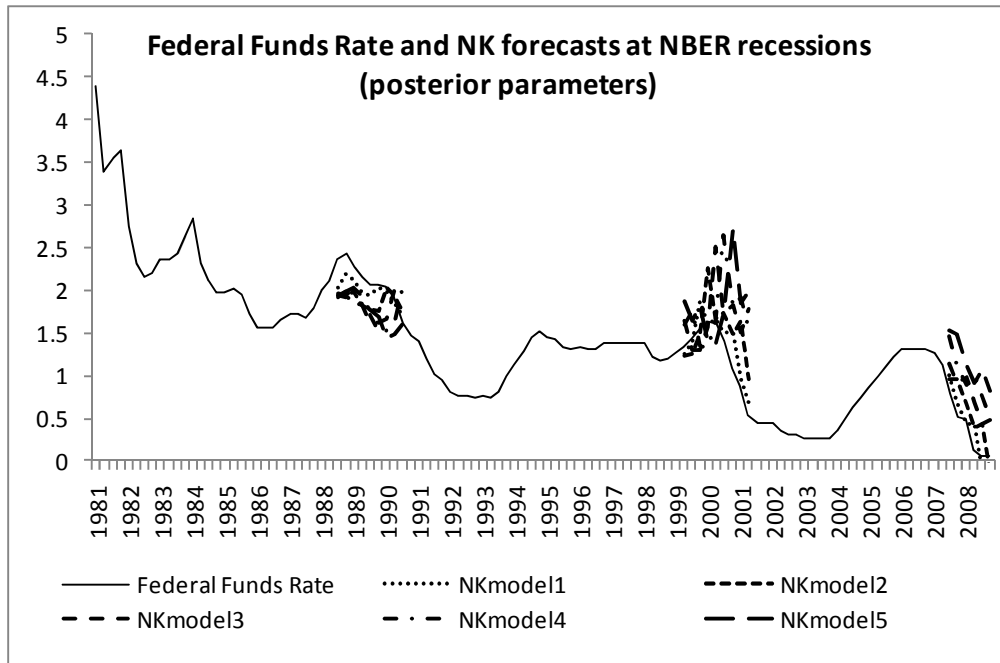
Figure 4.h



Source: Authors' calculations.

Note: This figure presents both the actual realization and predicted SPF values for interest rates (T bill-based) between 1981 and 2009, the latter set plotting only observations at different horizons (1 to 5) for recession periods as identified by the NBER.

Figure 4.i



Source: Authors' calculations.

Note: This figure presents both the actual realization and predicted NK values for interest rates (Federal Funds Rate) between 1981 and 2009, the latter set plotting only observations at different horizons (1 to 5) for recession periods as identified by the NBER.

## WORKING PAPERS

### 2010

- 1/10 MEASURING COMOVEMENT IN THE TIME-FREQUENCY SPACE  
— *António Rua*
- 2/10 EXPORTS, IMPORTS AND WAGES: EVIDENCE FROM MATCHED FIRM-WORKER-PRODUCT PANELS  
— *Pedro S. Martins, Luca David Opromolla*
- 3/10 NONSTATIONARY EXTREMES AND THE US BUSINESS CYCLE  
— *Miguel de Carvalho, K. Feridun Turkman, António Rua*
- 4/10 EXPECTATIONS-DRIVEN CYCLES IN THE HOUSING MARKET  
— *Luisa Lambertini, Caterina Mendicino, Maria Teresa Punzi*
- 5/10 COUNTERFACTUAL ANALYSIS OF BANK MERGERS  
— *Pedro P. Barros, Diana Bonfim, Moshe Kim, Nuno C. Martins*
- 6/10 THE EAGLE. A MODEL FOR POLICY ANALYSIS OF MACROECONOMIC INTERDEPENDENCE IN THE EURO AREA  
— *S. Gomes, P. Jacquinot, M. Pisani*
- 7/10 A WAVELET APPROACH FOR FACTOR-AUGMENTED FORECASTING  
— *António Rua*
- 8/10 EXTREMAL DEPENDENCE IN INTERNATIONAL OUTPUT GROWTH: TALES FROM THE TAILS  
— *Miguel de Carvalho, António Rua*
- 9/10 TRACKING THE US BUSINESS CYCLE WITH A SINGULAR SPECTRUM ANALYSIS  
— *Miguel de Carvalho, Paulo C. Rodrigues, António Rua*
- 10/10 A MULTIPLE CRITERIA FRAMEWORK TO EVALUATE BANK BRANCH POTENTIAL ATTRACTIVENESS  
— *Fernando A. F. Ferreira, Ronald W. Spahr, Sérgio P. Santos, Paulo M. M. Rodrigues*
- 11/10 THE EFFECTS OF ADDITIVE OUTLIERS AND MEASUREMENT ERRORS WHEN TESTING FOR STRUCTURAL BREAKS IN VARIANCE  
— *Paulo M. M. Rodrigues, Antonio Rubia*
- 12/10 CALENDAR EFFECTS IN DAILY ATM WITHDRAWALS  
— *Paulo Soares Esteves, Paulo M. M. Rodrigues*
- 13/10 MARGINAL DISTRIBUTIONS OF RANDOM VECTORS GENERATED BY AFFINE TRANSFORMATIONS OF INDEPENDENT TWO-PIECE NORMAL VARIABLES  
— *Maximiano Pinheiro*
- 14/10 MONETARY POLICY EFFECTS: EVIDENCE FROM THE PORTUGUESE FLOW OF FUNDS  
— *Isabel Marques Gameiro, João Sousa*
- 15/10 SHORT AND LONG INTEREST RATE TARGETS  
— *Bernardino Adão, Isabel Correia, Pedro Teles*
- 16/10 FISCAL STIMULUS IN A SMALL EURO AREA ECONOMY  
— *Vanda Almeida, Gabriela Castro, Ricardo Mourinho Félix, José Francisco Maria*
- 17/10 FISCAL INSTITUTIONS AND PUBLIC SPENDING VOLATILITY IN EUROPE  
— *Bruno Albuquerque*

- 18/10 GLOBAL POLICY AT THE ZERO LOWER BOUND IN A LARGE-SCALE DSGE MODEL  
— *S. Gomes, P. Jacquinot, R. Mestre, J. Sousa*
- 19/10 LABOR IMMOBILITY AND THE TRANSMISSION MECHANISM OF MONETARY POLICY IN A MONETARY UNION  
— *Bernardino Adão, Isabel Correia*
- 20/10 TAXATION AND GLOBALIZATION  
— *Isabel Correia*
- 21/10 TIME-VARYING FISCAL POLICY IN THE U.S.  
— *Manuel Coutinho Pereira, Artur Silva Lopes*
- 22/10 DETERMINANTS OF SOVEREIGN BOND YIELD SPREADS IN THE EURO AREA IN THE CONTEXT OF THE ECONOMIC AND FINANCIAL CRISIS  
— *Luciana Barbosa, Sónia Costa*
- 23/10 FISCAL STIMULUS AND EXIT STRATEGIES IN A SMALL EURO AREA ECONOMY  
— *Vanda Almeida, Gabriela Castro, Ricardo Mourinho Félix, José Francisco Maria*
- 24/10 FORECASTING INFLATION (AND THE BUSINESS CYCLE?) WITH MONETARY AGGREGATES  
— *João Valle e Azevedo, Ana Pereira*
- 25/10 THE SOURCES OF WAGE VARIATION: AN ANALYSIS USING MATCHED EMPLOYER-EMPLOYEE DATA  
— *Sónia Torres, Pedro Portugal, John T. Addison, Paulo Guimarães*
- 26/10 THE RESERVATION WAGE UNEMPLOYMENT DURATION NEXUS  
— *John T. Addison, José A. F. Machado, Pedro Portugal*
- 27/10 BORROWING PATTERNS, BANKRUPTCY AND VOLUNTARY LIQUIDATION  
— *José Mata, António Antunes, Pedro Portugal*
- 28/10 THE INSTABILITY OF JOINT VENTURES: LEARNING FROM OTHERS OR LEARNING TO WORK WITH OTHERS  
— *José Mata, Pedro Portugal*
- 29/10 THE HIDDEN SIDE OF TEMPORARY EMPLOYMENT: FIXED-TERM CONTRACTS AS A SCREENING DEVICE  
— *Pedro Portugal, José Varejão*
- 30/10 TESTING FOR PERSISTENCE CHANGE IN FRACTIONALLY INTEGRATED MODELS: AN APPLICATION TO WORLD INFLATION RATES  
— *Luís F. Martins, Paulo M. M. Rodrigues*
- 31/10 EMPLOYMENT AND WAGES OF IMMIGRANTS IN PORTUGAL  
— *Sónia Cabral, Cláudia Duarte*
- 32/10 EVALUATING THE STRENGTH OF IDENTIFICATION IN DSGE MODELS. AN A PRIORI APPROACH  
— *Nikolay Iskrev*
- 33/10 JOBLESSNESS  
— *José A. F. Machado, Pedro Portugal, Pedro S. Raposo*

## 2011

- 1/11 WHAT HAPPENS AFTER DEFAULT? STYLIZED FACTS ON ACCESS TO CREDIT  
— *Diana Bonfim, Daniel A. Dias, Christine Richmond*
- 2/11 IS THE WORLD SPINNING FASTER? ASSESSING THE DYNAMICS OF EXPORT SPECIALIZATION  
— *João Amador*

- 3/11 UNCONVENTIONAL FISCAL POLICY AT THE ZERO BOUND  
— *Isabel Correia, Emmanuel Farhi, Juan Pablo Nicolini, Pedro Teles*
- 4/11 MANAGERS' MOBILITY, TRADE STATUS, AND WAGES  
— *Giordano Mion, Luca David Opromolla*
- 5/11 FISCAL CONSOLIDATION IN A SMALL EURO AREA ECONOMY  
— *Vanda Almeida, Gabriela Castro, Ricardo Mourinho Félix, José Francisco Maria*
- 6/11 CHOOSING BETWEEN TIME AND STATE DEPENDENCE: MICRO EVIDENCE ON FIRMS' PRICE-REVIEWING STRATEGIES  
— *Daniel A. Dias, Carlos Robalo Marques, Fernando Martins*
- 7/11 WHY ARE SOME PRICES STICKIER THAN OTHERS? FIRM-DATA EVIDENCE ON PRICE ADJUSTMENT LAGS  
— *Daniel A. Dias, Carlos Robalo Marques, Fernando Martins, J. M. C. Santos Silva*
- 8/11 LEANING AGAINST BOOM-BUST CYCLES IN CREDIT AND HOUSING PRICES  
— *Luisa Lambertini, Caterina Mendicino, Maria Teresa Punzi*
- 9/11 PRICE AND WAGE SETTING IN PORTUGAL LEARNING BY ASKING  
— *Fernando Martins*
- 10/11 ENERGY CONTENT IN MANUFACTURING EXPORTS: A CROSS-COUNTRY ANALYSIS  
— *João Amador*
- 11/11 ASSESSING MONETARY POLICY IN THE EURO AREA: A FACTOR-AUGMENTED VAR APPROACH  
— *Rita Soares*
- 12/11 DETERMINANTS OF THE EONIA SPREAD AND THE FINANCIAL CRISIS  
— *Carla Soares, Paulo M. M. Rodrigues*
- 13/11 STRUCTURAL REFORMS AND MACROECONOMIC PERFORMANCE IN THE EURO AREA COUNTRIES: A MODEL-BASED ASSESSMENT  
— *S. Gomes, P. Jacquinot, M. Mohr, M. Pisani*
- 14/11 RATIONAL VS. PROFESSIONAL FORECASTS  
— *João Valle e Azevedo, João Tovar Jalles*