

## A COMPARISON OF FORECAST PERFORMANCE BETWEEN FEDERAL RESERVE STAFF FORECASTS, SIMPLE REDUCED-FORM MODELS, AND A DSGE MODEL

ROCHELLE M. EDGE, MICHAEL T. KILEY AND JEAN-PHILIPPE LAFORTE\*

*Federal Reserve Board, Washington, DC, USA*

### SUMMARY

This paper considers the ‘real-time’ forecast performance of the Federal Reserve staff, time-series models, and an estimated dynamic stochastic general equilibrium (DSGE) model—the Federal Reserve Board’s Estimated, Dynamic, Optimization-based (Edo) model. We evaluate forecast performance using out-of-sample predictions from 1996 to 2004. We find that the Edo model can provide forecasts that are competitive with those of Federal Reserve staff. Taken together with the fact that the Edo model has also proved useful in answering a range of policy questions, this suggests a significant and broad role for richly specified DSGE models in the toolbox of central bank models. Copyright © 2010 John Wiley & Sons, Ltd.

*Received 28 March 2008; Revised 22 August 2009*

### 1. INTRODUCTION

This paper compares the forecasts of an estimated dynamic stochastic general equilibrium (DSGE) model with that of the Federal Reserve staff and reduced-form time-series models. The paper has three goals. First, much of the related literature has compared forecasts from DSGE models with simple reduced-form forecasting techniques: our comparison with Federal Reserve staff forecasts provides a potentially more stringent test, given that previous research has shown the Federal Reserve staff forecast to be of high quality relative to alternative methods.<sup>1</sup> In addition, some of the research regarding DSGE models has found strong support for DSGE specifications using Bayesian measures of fit (such as posterior odds or marginal likelihoods); however, these measures can be dependent on the analyst’s prior views and, as emphasized by Sims (2003), often appear too decisive. Given this concern, we focus on out-of-sample forecast performance.<sup>2</sup> Finally, we examine forecast performance for both top-line macroeconomic variables—that is, the state of the labor market, growth of Gross Domestic Product, inflation, and the federal funds rate—and for detailed subcategories of aggregate expenditure—that is, consumption of nondurables and services

---

\* Correspondence to: Jean-Philippe Laforte, Federal Reserve Board, 20th Street and Constitution Av. NW, Washington, DC 20551, USA. E-mail: jean-philippe.laforte@frb.gov

<sup>1</sup> See Romer and Romer (2000) and Sims (2002); Tulip (2005) more recently finds some deterioration in the relative forecast performance of Federal Reserve staff forecasts. Faust and Wright (2007) and Reifschneider and Tulip (2007) provide related evidence.

<sup>2</sup> Other research has looked at similar issues in a more limited way. Specifically, our inclusion of a DSGE model within the set of forecast models follows recent work, in particular that by Smets and Wouters (2007), suggesting that advances in Bayesian estimation methods have made DSGE models capable of providing informative forecasts. Other pieces of research have explored related issues, primarily focusing on models of the euro area, as discussed below.

and investment in consumer durables, residential capital, and non-residential capital. This detailed focus is not common in DSGE models, which typically lump several of these categories into one broad category; however, policymakers have expressed interest in such details (see, for example, Kohn, 2003), and large macroeconometric models like the Federal Reserve's FRB/US model often produce forecasts at similar, or even more disaggregated levels.

Our DSGE model is the result of the Federal Reserve Board's project on Estimated, Dynamic, Optimization-based models; that is, the Edo model. This model contains a rich description of production, expenditure, labor supply, and pricing decisions for the economy of the USA. We have presented detailed descriptions of the model's structure, our estimation strategy, and results in previous papers (see Edge *et al.*, 2007a, 2008) and so we present only a brief summary of the model's structure in Section 2. For now, we simply highlight that the model has been designed to address a broad range of policy questions, as emphasized in Edge *et al.* (2008). For example, Gali and Gertler (2007) discuss two important contributions of DSGE models to monetary policy analysis: microeconomic foundations for economic dynamics merged with rational expectations for economic agents, and the role of fluctuations in natural rates of output and interest in policy determination. The Edo model has been used to analyze these issues, especially the latter, in Edge *et al.* (2008). We have also investigated the fluctuations in the US housing market, which have been considerable over the past decade, using the Edo model (see Edge *et al.*, 2007b). Importantly, we use the same model in this other research as in the forecasting analysis herein. While many academic investigations will consider specific models that are designed to address individual questions, the large number and broad range of questions that arise under significant time pressures within a policy institution require that the core models used for policy work be capable of spanning multiple questions. Indeed, Meyer (1997) emphasizes the multiple roles of macroeconomic models in policymaking and private sector consulting work, of which forecasting is but one example.

Our period of analysis spans macroeconomic developments in the USA from mid 1996 to late 2004 (where the end-point is determined by the public availability of data for forecast evaluation at the time of this study). This period was chosen for two reasons. First, the Federal Reserve's FRB/US model—a macroeconometric model specified with a neoclassical steady state and dynamic behavior designed to address the Lucas critique through consideration of the influence of expectations and other sources of dynamics—entered operation in mid 1996. As we aim to compare a cutting-edge DSGE model with existing practices at the Federal Reserve (and, to some extent, at other central banks), we focus on the period over which current practices have been employed. Second, the structure of our DSGE model—which, as discussed below, has two production sectors that experience 'slow' and 'fast' productivity growth—requires detailed data for estimation, and we have available the relevant 'real-time' data since about mid 1996.

Of course, the period we examine is also interesting for several economic reasons. Between 1996 and 2004, the US economy experienced substantial swings in activity. From 1996 to early 2000, economic growth was rapid and inflation was low—the 'Goldilocks' economy, as dubbed by Gordon (1998). A recession followed in 2001 and, while the recession was brief, the labor market was slow to recover (see Langdon *et al.*, 2004). Our analysis over this period allows an examination of the success of our model and other techniques at capturing this business cycle. Inflation developments were also significant during this period. For example, the Federal Open Market Committee highlighted the risk of an unwelcome fall in inflation in the spring of 2003, as the rate of change in the consumer price index excluding food and energy prices dropped to around 1% that year when measured on a real-time basis. Price inflation stepped up after 2003.

Our analysis yields support to the notion that a richly specified DSGE model can contribute to the forecasting efforts at policy institutions. We find that the forecast accuracy of the Edo model is as good as, and in many cases better than, that of the forecasts of the Federal Reserve staff and the FRB/US model or projections from time series methods. Together with the fact that the Edo model has also proved useful in answering policy questions quite different from forecasting, our results suggest a significant and broad role for richly specified DSGE models in the toolbox of central bank models.

We should emphasize one important caveat to these conclusions. Although the forecasts from the vector autoregressions and the Edo model use the same *data* that were available to Federal Reserve Board staff when formulating the Greenbook and FRB/US model forecasts, because they have been carried out retrospectively they do not use the same *stock of professional knowledge* that was available to Federal Reserve Board staff when formulating the Greenbook and FRB/US model forecasts. Specifically, the vector autoregression and (more notably) the Edo model forecasts benefit, at least indirectly, from our previous research and that of others on what types of models are likely to explain the data well. In addition, our vector autoregression and Edo model forecasts have not been formulated according to the same time pressures that Board staff face in putting together the Greenbook and FRB/US model forecasts. As such, these forecasts—because they have benefited from the luxury of our being able to check that all our codes are correct and that all data are being read in correctly—are likely to be less prone to errors that could potentially weaken forecast performance. It is impossible to purge our analysis of these influences and, for these reasons, we are cautious in our final verdict.

Before turning to our analysis, we would like to highlight several pieces of related research. Smets and Wouters (2007) demonstrated that a richly specified DSGE model could fit the US macroeconomic data well and provide out-of-sample forecasts that are competitive or superior to reduced-form vector autoregressions. We build on their work in several ways. First, our model contains a more detailed description of sectoral production and household/business expenditure decisions—which, as noted earlier, appears to be a prerequisite for a policy-relevant model. Second, we measure all economic variables in a manner more consistent with the official statistics published by the US Bureau of Economic Analysis (the statistics that form the basis of policy deliberations and public discussion of economic fluctuations), whereas, in contrast, Smets and Wouters (2007) make adjustments to published figures on consumption and investment in order to match the relative price restrictions implied by their one-sector model. Finally, and most importantly, we examine out-of-sample forecast performance using real-time data and compare our DSGE model's forecast performance with Federal Reserve staff forecasts and models, thereby pushing further on the question of whether DSGE models can give policy-relevant forecast information.

Research by Adolfson *et al.* (2007) is closely related to our analysis. These authors compare the forecast performance of the DSGE model of the Riksbank to Bayesian vector autoregression (BVAR) models and, like our analysis, central bank forecasts. However, these authors do not use real-time data, and they do not compare their DSGE model to another 'structural' model as we do to the pre-existing FRB/US model. Finally, our focus on US data and Federal Reserve forecasts is of independent interest given previous analyses of the quality of the Federal Reserve's forecasts (see Romer and Romer, 2000; Sims, 2002).

Other relevant research includes Lees *et al.* (2007), who compare the forecast performance of the Reserve Bank of New Zealand's official forecasts with those from a vector autoregressive model informed by priors from a DSGE model as suggested in Del Negro and Schorfheide (2004).

Our analysis shares the idea of comparing forecasts to staff forecasts at a central bank; such a comparison seems especially likely to illuminate the relevance of such techniques for policy work. However, we focus on forecasts from a DSGE model rather than those informed by a DSGE prior. The latter approach is something of a ‘black-box’, as the connection of the DSGE structure to the resulting forecast is tenuous (and asymptotically completely absent, as the data dominate the prior). Moreover, our reliance on a DSGE model directly allows us to make economically interesting inferences regarding the aspects of the model that contribute to its successes and failures. Finally, Lees *et al.* (2007) examine a very small set of variables—specifically, output, inflation, and the policy interest rate. Our experience with larger models like FRB/US at the Federal Reserve suggests that such small systems are simply not up to the challenge of addressing the types of questions demanded of models at large central banks (as we discuss in Edge *et al.*, 2008).

Adolfson *et al.* (2006) and Christoffel *et al.* (2007) examine out-of-sample forecast performance for DSGE models of the euro area. Their investigations are very similar to ours in directly considering a fairly large DSGE model. However, the focus of each of these pieces of research is on technical aspects of model evaluation. We eschew this approach and instead attempt to identify the economic sources of the successes and failures of our model. Also, neither of these studies uses real-time data, nor do they compare forecast performance to an alternative model employed at a central bank or official staff forecasts. As discussed, we focus on real-time data and compare forecast performance to the FRB/US model and Federal Reserve Greenbook forecasts. Overall, we view both Adolfson *et al.* (2006) and Christoffel *et al.* (2007) as complementary to our analysis, but feel that the explicit comparison to ‘real-world’ central bank practices is especially valuable.

The paper is organized as follows. Section 2 provides an overview of the Edo model. Section 3 discusses the estimation and evaluation of both the Edo model as well as the alternative forecasting models used in the paper’s analysis. Section 4 introduces the alternative forecasts that the paper considers: we focus on our DSGE model (Edo) forecasts, the Federal Reserve Board’s staff projections, including those from the FRB/US model, and the forecasts from autoregressions and vector autoregressions. We also discuss our real-time data in the Section 4. Section 5 presents the comparison between the Edo model and time series models. Section 6 examines the Federal Reserve forecasts and subsample results that illustrate important economic successes and failures of our model. We discuss amendments to our DSGE model that address some of these failures and hence provide an example of the type of lesson for structural modelers that can be gleaned from forecast exercises. Section 7 concludes and points to directions for future research.

## 2. A TWO-SECTOR DSGE MODEL FOR FORECASTING

Research on policy applications of DSGE models has exploded in the last 5 years. On the policy front, the GEM project at the International Monetary Fund (see Bayoumi *et al.*, 2004) and the SIGMA project at the Federal Reserve (see Erceg *et al.*, 2006) have provided examples of richly specified models with firm microeconomic foundations that can be applied to policy questions. However, even these rich models have not had the detail on domestic economic developments, such as specifications of highly disaggregated expenditure decisions, to address the range of questions typically analyzed by large models like the Federal Reserve’s FRB/US model.<sup>3</sup> The Edo model

<sup>3</sup> See Reifschneider *et al.* (1997) for a discussion of the use of models (including the FRB/US model) in forecasting at the Federal Reserve and Brayton *et al.* (1997) for a discussion of the use of models in policy analysis.

project at the Federal Reserve has been designed to build on earlier work at policy institutions, as well as academic research such as Smets and Wouters (2007) and Altig *et al.* (2004), by expanding the modeling of domestic economic decisions while investigating the ability of such DSGE models to examine a range of policy questions. For a detailed description and discussion of previous applications, the reader is referred to Edge *et al.* (2007a,b, 2008).

Figure 1 provides a graphical overview of the economy described by the Edo model. The model possesses two final goods (good ‘CBI’ and good ‘KB’, described more fully below), which are produced in two stages by intermediate- and then final-goods producing firms (shown in the center of the figure). On the model’s demand side, there are four components of private spending (each shown in a box surrounding the producers in the figure): consumer nondurable goods and services (sold to households), consumer durable goods, residential capital goods, and non-residential capital goods. Consumer nondurable goods and services and residential capital goods are purchased (by households and residential capital goods owners, respectively) from the first of economy’s two final goods producing sectors (good ‘CBI’ producers), while consumer durable goods and non-residential capital goods are purchased (by consumer durable and residential capital goods owners, respectively) from the second sector (good ‘KB’ producers). We ‘decentralize’ the economy by assuming that residential capital and consumer durables capital are rented to households, while non-residential capital is rented to firms. In addition to consuming the nondurable goods and services that they purchase, households also supply labor to the intermediate goods-producing firms in both sectors of the economy.

Our assumption of a two-sector production structure is motivated by the trends in certain relative prices and categories of real expenditure apparent in the data. Relative prices for investment goods, especially high-tech investment goods, have fallen and real expenditure on (and production of) such goods has grown more rapidly than that for other goods and services. A one-sector model is unable to deliver long-term growth and relative price movements that are consistent with these

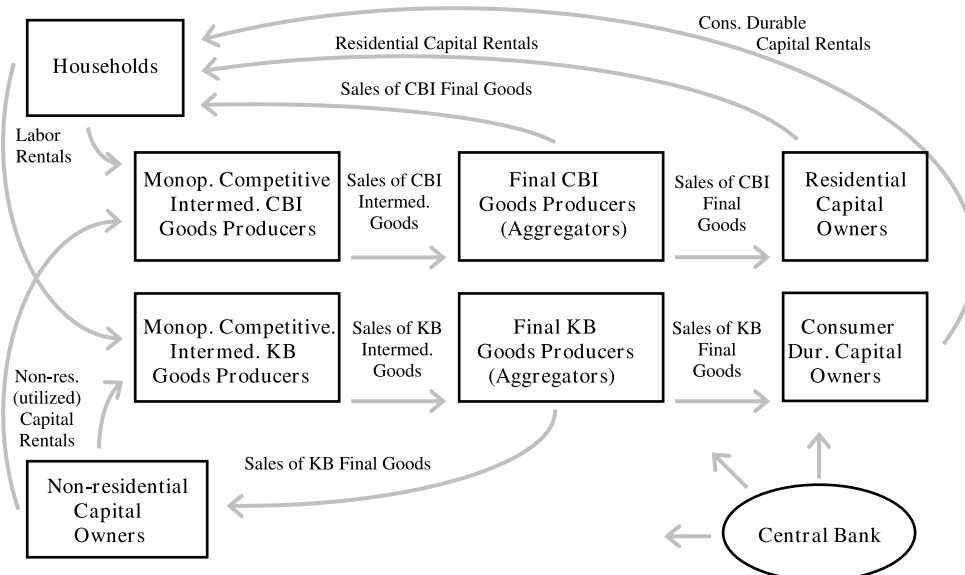


Figure 1. Model overview

stylized facts. As a result, we adopt a two-sector structure, with differential rates of technical progress across sectors. These different rates of technological progress induce secular relative price differentials, which in turn lead to different trend rates of growth across the economy's expenditure and production aggregates. We assume that the output of the slower-growing sector (denoted  $X_t^{\text{cbi}}$ ) is used for consumer nondurable goods and services and residential capital goods, and the output of a faster-growing sector (denoted  $X_t^{\text{kb}}$ ) is used for consumer durable goods and non-residential capital goods, roughly capturing the long-run properties of the data.

While differential trend growth rates are the primary motivation for our disaggregation of production, our specification of expenditure decisions is related to the well-known fact that the expenditure categories that we consider have different cyclical properties (see Edge *et al.*, 2008, for more details). Beyond the statistical motivation, our disaggregation of aggregate demand is motivated by the concerns of policymakers. A recent example relates to the divergent movements in household and business investment in the early stages of the US expansion following the 2001 recession, a topic discussed in Kohn (2003). We believe that providing a model that may explain the shifting pattern of spending through differential effects of monetary policy, technology, and preference shocks is a potentially important operational role for our disaggregated framework.

The remainder of this section provides an overview of the decisions made by each of the agents in our economy. Given some of the broad similarities between our model and others, our presentation is selective.

## 2.1. The Intermediate Goods Producer's Problem

We begin our description in the center of Figure 1. Intermediate goods producers in both sectors (specifically, sector 'CBI' and sector 'KB') produce output using a production technology that yields output (denoted  $X_t^s(j)$ ) from labor input,  $L_t^s(j)$ , capital input,  $K_t^{u,\text{nr},s}$  where the superscript 'u' denotes utilized capital and the superscript 'nr' indicates non-residential capital, and economy-wide and sector-specific productivity,  $Z_t^m$ , and  $Z_t^s$ .<sup>4</sup> Specifically:

$$X_t^s(j) = (K_t^{u,\text{nr},s}(j))^\alpha (Z_t^m Z_t^s L_t^s(j))^{1-\alpha} \text{ where } L_t^s(j) = \left( \int_0^1 L_t^s(i, j)^{\frac{\Theta_t^s-1}{\Theta_t^s}} di \right)^{\frac{\Theta_t^s}{\Theta_t^s-1}} \quad s = \text{cbi, kb} \quad (1)$$

Note that labor input is a Dixit–Stiglitz aggregate of differentiated labor inputs; this assumption will be an input in the wage Phillips curve discussed below.

The exogenous productivity terms contain a unit root; that is, they exhibit permanent movements in their levels. We assume that the stochastic processes  $Z_t^m$  and  $Z_t^{\text{kb}}$  evolve according to

$$\ln Z_t^n - \ln Z_{t-1}^n = \ln \Gamma_t^{z,n} = \ln(\Gamma_*^{z,n} \cdot \exp[\gamma_t^{z,n}]) = \ln \Gamma_*^{z,n} + \gamma_t^{z,n}, \quad n = \text{kb, } m \quad (2)$$

where  $\Gamma_*^{z,n}$  and  $\gamma_t^{z,n}$  are the steady-state and stochastic components of  $\Gamma_t^{z,n}$ . The stochastic component  $\gamma_t^{z,n}$  is assumed to evolve according to

$$\gamma_t^{z,n} = \rho^{z,n} \gamma_{t-1}^{z,n} + \varepsilon_t^{z,n} \quad n = \text{kb, } m \quad (3)$$

<sup>4</sup> We normalize  $Z_t^{\text{cbi}}$  to one, while  $Z_t^{\text{kb}}$  is not restricted.

where  $\varepsilon_t^{z,n}$  is an i.i.d. shock process, and  $\rho^{z,n}$  represents the persistence of  $\gamma_t^{z,n}$  to a shock. It is the presence of capital-specific technological progress that allows the model to generate differential trend growth rates in the economy's two production sectors. We will estimate the steady-state rates of technological progress in each sector, as described below. However, we note at this point that the data will imply a more rapid rate of technological progress in capital goods production.

Each intermediate goods producer's output enters a final-goods production technology for its sector that takes the form of the Dixit–Stiglitz aggregator. As a result, intermediate goods producers are monopolistic competitors. We further assume that the intermediate goods producers face a quadratic cost of adjusting the nominal price they charge. Consequently, an intermediate goods producing firm chooses the optimal nominal price (and the quantity it will supply consistent with that price), taking as given the marginal cost,  $MC_t^s(j)$ , of producing a unit of output,  $X_t^s(j)$ , the aggregate price level for its sector,  $P_t^s$ , and households' valuation of a unit of nominal rental income in each period,  $\Lambda_t^{\text{cnn}}/P_t^{\text{cbi}}$ , to solve

$$\begin{aligned} \max_{\{P_t^s(j), X_t^s(j), X_t^s(j)\}_{t=0}^{\infty}} \quad & \mathcal{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{\Lambda_t^{\text{cnn}}}{P_t^{\text{cbi}}} \left\{ P_t^s(j) X_t^s(j) - MC_t^s(j) X_t^s(j) \right. \\ & \left. - \frac{100 \cdot \chi^p}{2} \left( \frac{P_t^s(j)}{P_{t-1}^s(j)} - \eta^p \Pi_{t-1}^{p,s} - (1 - \eta^p) \Pi_*^{p,s} \right)^2 P_t^s X_t^s \right\} \\ \text{subject to } & X_\tau^s(j) = (P_\tau^s(j)/P_\tau^s)^{-\Theta_\tau^{x,s}} X_\tau^s \text{ for } \tau = 0, 1, \dots, \infty \text{ and } s = \text{cbi, kb} \end{aligned} \quad (4)$$

The profit function includes price-setting adjustment costs, whose magnitude depends on the parameter  $\chi^p$  and the rate of inflation relative to both the lagged and steady-state inflation rates. This type of price-setting decision delivers a new-Keynesian Phillips curve. Because adjustment costs potentially depend upon lagged inflation, the Phillips curve can take the 'hybrid' form in which inflation is linked to its own lead and lag as well as marginal cost.

The constraint against which the firm maximizes its profits is the demand curve it faces for its differentiated good, which derives from the final goods producing firm's cost minimization problem. Of particular importance for our estimation strategy and forecasting analysis is the parameter  $\Theta_t^{x,s}$ , the stochastic elasticity of substitution between the differentiated intermediate goods inputs used in the production of the consumption or capital goods sectors. We assume that

$$\theta_t^{x,s} = \varepsilon_t^{\theta,x,s} \quad (5)$$

where  $\varepsilon_t^{\theta,x,s}$  is an i.i.d. shock process. A stochastic elasticity of substitution introduces transitory markup shocks into the pricing decisions of intermediate goods producers.

A lengthier treatment of the structure of our model is provided in Edge *et al.* (2007a), which further details the cost minimization problem facing intermediate goods producers in choosing the optimal mix of factors of production. This problem determines the factors influencing marginal cost and hence pricing. At this point, we emphasize that the production and pricing decisions of the intermediate goods firms in our model economy are influenced by four 'aggregate supply' shocks: two productivity shocks, corresponding to economy-wide and capital-specific technology shocks, and two markup shocks that induce transitory fluctuations in the nominal prices in each sector.

## 2.2. The Capital Owner's Problem

We now shift from producers' decisions to spending decisions (that is, those by agents encircling our producers in Figure 1). Non-residential capital owners choose investment in non-residential capital,  $E_t^{\text{nr}}$ , the stock of non-residential capital,  $K_t^{\text{nr}}$  (which is linked to the investment decision via the capital accumulation identity), and the amount and utilization of non-residential capital in each production sector,  $K_t^{\text{nr},\text{cbi}}$ ,  $U_t^{\text{cbi}}$ ,  $K_t^{\text{nr},\text{kb}}$ , and  $U_t^{\text{kb}}$ .<sup>5</sup> (Recall that production in equation 1 depends on utilized capital  $K_t^{\text{nr},\text{nr},s} = U_t^s K_t^{\text{nr},s}$ .) This decision is described by the following maximization problem (in which the rental rate on non-residential capital,  $R_t^{\text{nr}}$ , the price of non-residential capital goods,  $P_t^{\text{kb}}$ , and households' valuation of nominal capital income in each period,  $\Lambda_t^{\text{cnn}}/P_t^{\text{cbi}}$  are taken as given):

$$\begin{aligned} & \max_{\{E_t^{\text{nr}}(k), K_{t+1}^{\text{nr}}(k), K_t^{\text{nr},\text{cbi}}(k), K_t^{\text{nr},\text{kb}}(k), U_t^{\text{cbi}}(k), U_t^{\text{kb}}(k)\}_{t=0}^{\infty}} \\ & \mathcal{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{\Lambda_t^{\text{cnn}}}{P_t^{\text{cbi}}} \left\{ R_t^{\text{nr}} U_t^{\text{cbi}}(k) K_t^{\text{nr},\text{cbi}}(k) + R_t^{\text{nr}} U_t^{\text{kb}}(k) K_t^{\text{nr},\text{kb}}(k) - P_t^{\text{kb}} E_t^{\text{nr}}(k) \right. \\ & \quad \left. - \kappa \left( \frac{U_t^{\text{cbi}}(k)^{1+\psi} - 1}{1+\psi} \right) P_t^{\text{kb}} K_t^{\text{nr},\text{cbi}} - \kappa \left( \frac{U_t^{\text{kb}}(k)^{1+\psi} - 1}{1+\psi} \right) P_t^{\text{kb}} K_t^{\text{nr},\text{kb}} \right\} \end{aligned}$$

subject to

$$\begin{aligned} K_{\tau+1}^{\text{nr}}(k) &= (1 - \delta^{\text{nr}}) K_{\tau}^{\text{nr}}(k) + A_{\tau}^{\text{nr}} E_{\tau}^{\text{nr}}(k) - \frac{100 \cdot \chi^{\text{nr}}}{2} \left( \frac{E_{\tau}^{\text{nr}}(k) - E_{\tau-1}^{\text{nr}}(k) \Gamma_t^{\text{y},\text{kb}}}{K_{\tau}^{\text{nr}}} \right)^2 K_{\tau}^{\text{nr}} \text{ and} \\ K_{\tau}^{\text{nr},\text{cbi}}(k) + K_{\tau}^{\text{nr},\text{kb}}(k) &= K_{\tau}^{\text{nr}}(k) \text{ for } \tau = 0, 1, \dots, \infty \end{aligned} \quad (6)$$

The parameter  $\delta^{\text{nr}}$  in the capital accumulation constraint denotes the depreciation rate for non-residential capital, while the parameter  $\chi^{\text{nr}}$  governs how quickly investment adjustment costs increase when  $(E_{\tau}^{\text{nr}}(k) - E_{\tau-1}^{\text{nr}}(k) \Gamma_t^{\text{y},\text{kb}})$  rises above zero. The variable  $A_t^{\text{nr}}$  is a stochastic element affecting the efficiency of non-residential investment in the capital accumulation process. Letting  $a_t^{\text{nr}} \equiv \ln A_t^{\text{nr}}$  denote the log deviation of  $A_t^{\text{nr}}$  from its steady-state value of unity, we assume that

$$a_t^{\text{nr}} = \rho^{a,\text{nr}} a_{t-1}^{\text{nr}} + \varepsilon_t^{a,\text{nr}} \quad (7)$$

The variable  $\varepsilon_t^{a,\text{nr}}$  is an i.i.d. shock process, and  $\rho^{a,\text{nr}}$  represents the persistence of  $A_t^{\text{nr}}$  from steady state following a shock to equation (7).

The problems solved by the consumer durables and residential capital owners are slightly simpler than the non-residential capital owner's problems. Because utilization rates are not variable for these types of capital, their owners make only investment and capital accumulation decisions.<sup>6</sup>

<sup>5</sup> Higher rates of utilization incur a cost (reflected in the last two terms in the capital owner's profit function). We assume that  $\kappa = R_{\text{nr}}^{\text{nr}}/P_{\text{nr}}^{\text{kb}}$ , which implies that utilization is unity in the steady state.

<sup>6</sup> Our main reason for allowing for variable capital utilization in non-residential capital—while omitting it from consumer durables and residential capital—stems from the fact that the main role that variable capital utilization plays in models like ours is that of providing a means, in addition to sticky prices and wages, of generating inertia in inflation (see



Taking as given the rental rate on consumer durables capital,  $R_t^{\text{cd}}$ , and the price of consumer-durable goods,  $P_t^{\text{kb}}$ , and households' valuation of nominal capital income,  $\Lambda_t^{\text{cnn}}/P_t^{\text{cbi}}$ , the capital owner chooses investment in consumer durables,  $I_t^{\text{cd}}$ , and its implied capital stock,  $K_t^{\text{cd}}$ , to solve

$$\max_{\{E_t^{\text{cd}}(k), K_{t+1}^{\text{cd}}(k)\}_{t=0}^{\infty}} \mathcal{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{\Lambda_t^{\text{cnn}}}{P_t^{\text{cbi}}} \{R_t^{\text{cd}} K_t^{\text{cd}}(k) - P_t^{\text{kb}} E_t^{\text{cd}}(k)\}$$

subject to

$$K_{\tau+1}^{\text{cd}}(k) = (1 - \delta^{\text{cd}}) K_{\tau}^{\text{cd}}(k) + A_{\tau}^{\text{cd}} E_{\tau}^{\text{cd}}(k) - \frac{100 \cdot \chi^{\text{cd}}}{2} \left( \frac{E_{\tau}^{\text{cd}}(k) - E_{\tau-1}^{\text{cd}}(k) \Gamma_{\tau}^{\text{x, kb}}}{K_{\tau}^{\text{cd}}} \right)^2 K_{\tau}^{\text{cd}} \quad (8)$$

for  $\tau = 0, 1, \dots, \infty$

The residential capital owner's decision is analogous:

$$\max_{\{E_t^r(k), K_{t+1}^r(k)\}_{t=0}^{\infty}} \mathcal{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{\Lambda_t^{\text{cnn}}}{P_t^{\text{cbi}}} \{R_t^r K_t^r(k) - P_t^{\text{cbi}} E_t^r(k)\}$$

subject to

$$K_{\tau+1}^r(k) = (1 - \delta^r) K_{\tau}^r(k) + A_{\tau}^r E_{\tau}^r(k) - \frac{100 \cdot \chi^r}{2} \left( \frac{E_{\tau}^r(k) - E_{\tau-1}^r(k) \Gamma_{\tau}^{\text{x, cbi}}}{K_{\tau}^{\text{cd}}} \right)^2 K_{\tau}^{\text{cd}} \quad (9)$$

for  $\tau = 0, 1, \dots, \infty$

The notation for the consumer durables and residential capital stock problems parallels that of non-residential capital. In particular, the capital-efficiency shocks,  $A_t^{\text{cd}}$  and  $A_t^r$ , follow an autoregression process similar to that given in equation (7).

We emphasize two points related to capital accumulation. First, capital accumulation is subject to adjustment costs, and hence investment responds slowly to many shocks. In addition, the 'capital accumulation technologies' are themselves subject to efficiency shocks. These three shocks to the efficiency of investment—non-residential investment, residential investment, and investment in consumer durables—enter the optimality conditions driving investment decisions as shocks to the 'intertemporal IS curves.'

### 2.3. The Household's Problem

The final private agent in the model that we will discuss is the representative household, which makes both expenditures and labor-supply decisions. The household derives utility from four

---

Christiano *et al.*, 2005). Specifically, variable utilization in non-residential capital affects the rental rate on non-residential capital paid by firms, which in turn affects marginal cost—the driving process on inflation. As a result, variable utilization in non-residential capital allows for greater inflation inertia. Variable utilization in consumer durables and residential capital would, if present, affect the rental rates on both these types of capital and in turn the shadow marginal costs of producing consumer durables and residential capital services. However, these are not variables that we can either observe or are especially critical for our model to be able to explain. For this reason, there is little purpose in allowing for variable utilization in consumer durables and residential capital; therefore, we do not allow variable utilization for these categories.

sources: its purchases of the consumer nondurable goods and non-housing services, the flow of services from its rental of consumer durable capital, the flow of services from its rental of residential capital, and its leisure time, which is equal to what remains of its time endowment after labor is supplied to the market. Preferences are separable over all arguments of the utility function.

The utility that the household derives from the three components of goods and services consumption is influenced by its habit stock for each of these consumption components, a feature that has been shown to be important for consumption dynamics in similar models. A household's habit stock for its consumption of nondurable goods and non-housing services is equal to a factor  $h^{\text{cnn}}$  multiplied by its consumption last period  $E_{t-1}^{\text{cnn}}$ . Its habit stock for the other components of consumption is defined similarly.

The household chooses its purchases of consumer nondurable goods and services,  $E_t^{\text{cnn}}$ , the quantities of residential and consumer durable capital it wishes to rent,  $K_t^r$  and  $K_t^{\text{cd}}$ , its holdings of bonds,  $B_t$ , its wage for each sector,  $W_t^{\text{cbi}}$  and  $W_t^{\text{kb}}$ , and supply of labor consistent with each wage,  $L_t^{\text{cbi}}$  and  $L_t^{\text{kb}}$ . This decision is made subject to the household's budget constraint, which reflects the costs of adjusting wages and the mix of labor supplied to each sector, as well as the demand curve it faces for its differentiated labor. Specifically, the household solves

$$\begin{aligned} & \max_{\{E_t^{\text{cnn}}, K_t^{\text{cd}}, K_t^r, \{W_t^s(i), L_t^s(i)\}_{s=\text{cbi}, \text{kb}}, B_{t+1}(i)\}_{t=0}^\infty} \\ & \mathcal{E}_0 \sum_{t=0}^\infty \beta^t \left\{ \zeta^{\text{cnn}} \Xi_t^{\text{cnn}} \ln(E_t^{\text{cnn}}(i) - h^{\text{cnn}} E_{t-1}^{\text{cnn}}(i)) + \zeta^{\text{cd}} \Xi_t^{\text{cd}} \ln(K_t^{\text{cd}}(i) - h^{\text{cd}} K_{t-1}^{\text{cd}}(i)) \right. \\ & \quad \left. + \zeta^r \Xi_t^r \ln(K_t^r(i) - h^r K_{t-1}^r(i)) - \zeta^l \Xi_t^l \frac{(L_t^{\text{cbi}}(i) + L_t^{\text{kb}}(i))^{1+\nu}}{1+\nu} \right\} \end{aligned}$$

subject to

$$\begin{aligned} R_{\tau+1}^{-1} B_{\tau+1}(i) &= B_{\tau}(i) + \sum_{s=\text{cbi}, \text{kb}} W_{\tau}^s(i) L_{\tau}^s(i) + \text{Profits}_{\tau}(i) + \text{Other transfers}_{\tau}(i) - P_{\tau}^{\text{cbi}} E_{\tau}^{\text{cnn}}(i) \\ &\quad - R_{\tau}^{\text{cd}} K_{\tau}^{\text{cd}} - R_{\tau}^r K_{\tau}^r - \sum_{s=\text{cbi}, \text{kb}} \frac{100 \cdot \chi^w}{2} \left( \frac{W_{\tau}^s(j)}{W_{\tau-1}^s(j)} - \eta^w \Pi_{\tau-1}^{w,s} - (1 - \eta^w) \Pi_{\tau}^w \right)^2 W_{\tau}^s L_{\tau}^s \\ &\quad - \frac{100 \cdot \chi^l}{2} \left( \frac{L_{\tau}^{\text{cbi}} \cdot W_{\tau}^{\text{cbi}}}{L_{\tau}^{\text{cbi}} + L_{\tau}^{\text{kb}}} + \frac{L_{\tau}^{\text{kb}} \cdot W_{\tau}^{\text{kb}}}{L_{\tau}^{\text{cbi}} + L_{\tau}^{\text{kb}}} \right) \left( \frac{L_{\tau}^{\text{cbi}}(i)}{L_{\tau}^{\text{kb}}(i)} - \eta^l \frac{L_{\tau-1}^{\text{cbi}}}{L_{\tau-1}^{\text{kb}}} - (1 - \eta^l) \frac{L_{\tau}^{\text{cbi}}}{L_{\tau}^{\text{kb}}} \right)^2 \frac{L_{\tau}^{\text{kb}}}{L_{\tau}^{\text{cbi}}}. \\ L_{\tau}^{\text{cbi}}(i) &= (W_{\tau}^{\text{cbi}}(i)/W_{\tau}^{\text{cbi}})^{-\Theta_{\tau}^{l, \text{cbi}}} L_{\tau}^{\text{cbi}}, \text{ and } L_{\tau}^{\text{kb}}(i) = (W_{\tau}^{\text{kb}}(i)/W_{\tau}^{\text{kb}})^{-\Theta_{\tau}^{l, \text{kb}}} L_{\tau}^{\text{kb}}, \\ &\text{for } \tau = 0, 1, \dots, \infty \end{aligned} \tag{10}$$

In the utility function the parameter  $\beta$  is the household's discount factor,  $\nu$  denotes its inverse labor supply elasticity, while  $\zeta^{\text{cnn}}$ ,  $\zeta^{\text{cd}}$ ,  $\zeta^r$ , and  $\zeta^l$  are scale parameters that tie down the ratios between the household's consumption components. The stationary, unit-mean, stochastic variables  $\Xi_t^{\text{cnn}}$ ,  $\Xi_t^{\text{cd}}$ ,  $\Xi_t^r$ , and  $\Xi_t^l$  represent aggregate shocks to the household's utility of its consumption components and its disutility of labor.

Letting  $\xi_t^x \equiv \ln \Xi_t^x$  denote the log-deviation of  $\Xi_t^x$  from its steady-state value of unity, we assume that

$$\xi_t^x = \rho^{\xi,x} \xi_{t-1}^x + \varepsilon_t^{\xi,x}, \quad x = \text{cnn, cd, } r, l \quad (11)$$

The variable  $\varepsilon_t^{\xi,x}$  is an i.i.d. shock process, and  $\rho^{\xi,x}$  represents the persistence of  $\Xi_t^x$  away from steady-state following a shock to equation (11).

The household's budget constraint includes wage-setting adjustment costs, which depend on the parameter  $\chi^w$  and the level of wage inflation relative to the lagged and steady-state wage inflation rates. These costs, and the monopoly power enjoyed by households in the supply of differentiated labor input to intermediate goods producers as discussed above, yield a wage Phillips curve much like the price Phillips curve discussed previously. In addition, there are costs in changing the mix of labor supplied to each sector, which depend on the parameter  $\chi^l$ . These costs incurred by the household when the mix of labor input across sectors changes may be important for sectoral comovements.

In summary, the household's optimal decisions are influenced by four structural shocks: shocks to the utility associated with nondurable and services consumption, durables consumption, housing services, and labor supply. The first three affect 'intertemporal IS curves' associated with consumption choices, while the last enters the intratemporal optimality condition influencing labor supply.

## 2.4. Monetary Authority

We now turn to the last important agent in our model, the monetary authority. It sets monetary policy in accordance with a Taylor-type interest rate feedback rule. Policymakers smoothly adjust the actual interest rate  $R_t$  to its target level  $\bar{R}_t$ :

$$R_t = (R_{t-1})^{\phi^r} (\bar{R}_t)^{1-\phi^r} \exp[\varepsilon_t^r] \quad (12)$$

where the parameter  $\phi^r$  reflects the degree of interest rate smoothing, while  $\varepsilon_t^r$  represents a monetary policy shock. The central bank's target nominal interest rate,  $\bar{R}_t$ , depends on GDP growth relative to steady-state growth,  $H_t^{\text{gdp}}/H_*^{\text{gdp}}$ , the acceleration of GDP growth,  $H_t^{\text{gdp}}/H_{t-1}^{\text{gdp}}$ , GDP inflation relative to target,  $\Pi_t^{p,\text{gdp}}/\Pi_*^{p,\text{gdp}}$ , and the acceleration of GDP inflation,  $\Pi_t^{p,\text{gdp}}/\Pi_{t-1}^{p,\text{gdp}}$ :

$$\bar{R}_t = \left( \frac{H_t^{\text{gdp}}}{H_*^{\text{gdp}}} \right)^{\phi^{h,\text{gdp}}} \left( \frac{H_t^{\text{gdp}}}{H_{t-1}^{\text{gdp}}} \right)^{\phi^{\Delta h,\text{gdp}}} \left( \frac{\Pi_t^{p,\text{gdp}}}{\Pi_*^{p,\text{gdp}}} \right)^{\phi^{\pi,\text{gdp}}} \left( \frac{\Pi_t^{p,\text{gdp}}}{\Pi_{t-1}^{p,\text{gdp}}} \right)^{\phi^{\Delta \pi,\text{gdp}}} R_* \quad (13)$$

In equation (13),  $R_*$  denotes the economy's steady-state nominal interest rate and  $\phi^{h,\text{gdp}}$ ,  $\phi^{\Delta h,\text{gdp}}$ ,  $\phi^{\pi,\text{gdp}}$ , and  $\phi^{\Delta \pi,\text{gdp}}$  denote the weights in the feedback rule.

## 2.5. Measuring Aggregate Output

We have focused on sectoral production decisions so far and have not yet discussed Gross Domestic Product (GDP). The growth rate of real GDP is defined as the Divisia (share-weighted) aggregate

of final spending in the economy, as given by the identity:

$$H_t^{\text{gdp}} = \left( \left( \frac{X_t^{\text{cbi}}}{X_{t-1}^{\text{cbi}}} \right)^{P_*^{\text{cbi}} X_*^{\text{cbi}}} \left( \frac{X_t^{\text{kb}}}{X_{t-1}^{\text{kb}}} \right)^{P_*^{\text{kb}} X_*^{\text{kb}}} \left( \frac{\Gamma_t^{\text{x,cbi}} \cdot \tilde{X}_t^{\text{gf}}}{\tilde{X}_{t-1}^{\text{gf}}} \right)^{P_*^{\text{cbi}} X_*^{\text{gf}}} \right)^{\frac{1}{P_*^{\text{cbi}} X_*^{\text{cbi}} + P_*^{\text{kb}} X_*^{\text{kb}} + P_*^{\text{cbi}} X_*^{\text{gf}}}} \quad (14)$$

In equation (14),  $\tilde{X}_t^{\text{gf}}$  represents stationary unmodeled output (that is, GDP other than  $E_t^{\text{cnn}}$ ,  $E_t^{\text{cd}}$ ,  $E_t^r$ , and  $E_t^{\text{nr}}$ ). To a first-order approximation, this definition of GDP growth is equivalent to how it is defined in the US National Income and Product Accounts.

Stationary unmodeled output is exogenous and is assumed to follow the process

$$\ln \tilde{X}_t^{\text{gf}} - \ln \tilde{X}_*^{\text{gf}} = \rho^{\text{x,gf}} (\ln \tilde{X}_t^{\text{gf}} - \ln \tilde{X}_*^{\text{gf}}) + \varepsilon^{\text{x,gf}}$$

This shock is another ‘demand’ shock, in conjunction with the shocks to capital efficiency and the utility associated with various components of consumption (excluding leisure).

The inflation rate of the GDP deflator, represented by  $\Pi_t^{p,\text{gdp}}$ , is defined implicitly by

$$\Pi_t^{p,\text{gdp}} H_t^{\text{gdp}} = \frac{P_t^{\text{gdp}} X_t^{\text{gdp}}}{P_{t-1}^{\text{gdp}} X_{t-1}^{\text{gdp}}} = \frac{P_t^{\text{cbi}} X_t^{\text{cbi}} + P_t^{\text{kb}} X_t^{\text{kb}} + P_t^{\text{cbi}} X_t^{\text{gf}}}{P_t^{\text{cbi}} X_{t-1}^{\text{cbi}} + P_{t-1}^{\text{kb}} X_{t-1}^{\text{kb}} + P_t^{\text{cbi}} X_{t-1}^{\text{gf}}}$$

## 2.6. Summary

Our presentation of the model has been brief. However, it has highlighted that our model, although it considers production and expenditure decisions in a bit more detail, shares many features with other DSGE models in the literature, including imperfect competition, nominal price and wage rigidities, and real frictions like adjustment costs and habit persistence. The rich specification of structural shocks (to productivity, preferences, capital efficiency, and mark-ups) and adjustment costs allows our model to be brought to the data with some chance of finding empirical validation.

While the fluctuations in economic variables within the Edo model reflect complex interactions between the large set of decisions made within the economy, we would also highlight a couple of structural features that may play an important role in its forecast performance. First, the model assumes a stochastic structure for productivity shocks in each sector that allows for important business cycle frequency fluctuations in technology. This view contrasts significantly with the view in early versions of the FRB/US model, where technology was modeled as a linear time trend with breaks. More recent versions of the FRB/US model have allowed for more variation in ‘trend’ total factor productivity, but the structure of the FRB/US model is not embedded in the tradition started by Kydland and Prescott (1982) and, as a result, the role of technology in fluctuations—and forecasts—of economic activity may be quite different between Edo and models or forecasting techniques similar to those embedded in the FRB/US model.

In addition, the Edo model summarizes the state of the ‘labor market’ through the behavior of hours per capita. Policy discussions will often highlight distinctions between employment and hours per worker and between employment and unemployment. We view extensions of the Edo model along these dimensions as interesting topics for future research. For now, we simply note that, over the period from the mid 1980s through 2004, the correlation between hours per capita and the unemployment rate (using currently published data) exceeded 0.85, suggesting that our

focus on hours per capita provides a reasonable first step in examining the ability of the model to capture the state of the labor market, broadly interpreted. That said, we acknowledge that the ability of the model to adequately address the welfare implications of changes in unemployment is more open to question.

Finally, we would emphasize that the behavior of prices and wages in the Edo model is governed by versions of ‘New-Keynesian’ price and wage Phillips curves. There has been a spirited debate over the empirical performance of such specifications (see Kiley, 2007; Laforte, 2007; Rudd and Whelan, 2007).

### 3. ESTIMATION AND EVALUATION OF THE DSGE AND OTHER MODELS

Before turning to our ‘real-time’ forecast exercise, it is instructive to consider an evaluation of the DSGE model that focuses on within-sample fit because such metrics have dominated recent research (see Smets and Wouters, 2007). We focus on the marginal likelihood for the DSGE model and some time series alternatives.

The DSGE model is estimated using 12 economic time series for the USA:

1. Real GDP
2. Real consumption expenditure on nondurables and services
3. Real consumption expenditure on durables
4. Real residential investment expenditure
5. Real non-residential investment expenditure, which equals real gross private domestic investment minus real residential investment<sup>7</sup>
6. GDP price inflation
7. Inflation for consumer nondurables and services
8. Inflation for consumer durables
9. Hours, which equals hours of all persons in the non-farm business sector
10. Real wage inflation, which equals the percent change in compensation per hour in the non-farm business sector deflated by the price level for consumer nondurables and services
11. The federal funds rate
12. The yield on the 10-year US Treasury Note

As is the standard practice, we estimate a log-linearized approximation to our model, which we cast in its state-space representation for the set of (in our case 12) observable variables listed above. We then use the Kalman filter to evaluate the likelihood of the observed variables, and form the posterior distribution of the parameters of interest by combining the likelihood function with a joint density characterizing some prior beliefs over parameters. Since we do not have a closed-form solution of the posterior, we rely on Markov chain Monte Carlo (MCMC) methods. We also add measurement errors processes, denoted  $\eta_t$ , for all of the observed series used in estimation except the nominal interest rate and the aggregate hours series.

Our estimation results depend upon our specification of priors and calibration of certain parameters. We use the same priors and calibration strategy for our full-sample estimation and for the out-of-sample forecast exercises we present below. A number of parameters are calibrated

<sup>7</sup> Subtraction is performed using the appropriate techniques for aggregates measured as Fisher Ideal indexes.

and held fixed throughout. As reported in Table I, we fix the household's discount factor ( $\beta$ ), the Cobb–Douglas share of capital input ( $\alpha$ ), the curvature parameter associated with costs of varying capital utilization ( $\psi$ ), the depreciation rates ( $\delta^{\text{nr}}$ ,  $\delta^{\text{cd}}$ ,  $\delta^r$ ), and the elasticities of substitution between differentiated intermediate goods and labor input ( $\Theta_{*}^{\text{x,cbi}}$ ,  $\Theta_{*}^{\text{x,kb}}$ ,  $\Theta_{*}^l$ ). Forecast performance is not very sensitive to reasonable (small) variation in these parameters. To some extent, our specifications of priors and (to a lesser degree) our calibrations of certain parameters are yet another reason why, as discussed in Section 1, our analysis is not strictly real time. Specifically, it is likely that some of our priors and calibrations are informed by research undertaken previous to this study but after the reference date of our real-time DSGE model forecast (and therefore may be based on subsequently available data). This contamination of the pure real-time forecasting exercise will likely allow the DSGE model forecast to perform better than it would had it actually been undertaken at the forecast reference date. Consequently, as discussed previously, some caution in interpreting our final results is warranted.

We also ‘calibrate’, in real time, a number of parameters important for steady-state growth and inflation. Specifically, we set the steady-state rate of inflation for nondurable and services consumption equal to the average realized over the 5 years prior to the start of the forecast period, and we estimate the steady-state rate of productivity growth in each sector to match the rate of growth of real GDP and real wages implied by the model to the corresponding values in the data from the fourth quarter of 1984 to the start of the forecast period. These choices determine the parameters  $\Pi_{*}^{p,\text{gdp}}$ ,  $\Gamma_{*}^{z,m}$ , and  $\Gamma_{*}^{z,\text{kb}}$ . Therefore, these parameters vary as we move our forecast window forward, but are not based on information from the relevant forecast period.

The remainder of the model parameters are estimated in real time.<sup>8</sup> The priors placed over the model parameters are reported in Tables II and III. We highlight the following: the parameters governing habit persistence ( $h^{\text{cnn}}$ ,  $h^{\text{cd}}$ ,  $h^r$ ) have prior distributions spanning the interval 0 to 1 that are centered on 0.5 and relatively uninformative; the parameters determining the indexation of price and wage inflation to lagged inflation are centered on 0, consistent with the ‘theory’ of the New-Keynesian Phillips curve that often implies no indexation (that is, indexation is typically added as an *ad hoc* adjustment to fit the data); and the parameters governing the autocorrelation in the structural shocks have prior distributions that span 0 to 1 and typically are centered on moderate to high degrees of persistence.<sup>9</sup>

Table I. Calibrated parameters

$\beta$	$\alpha$	$\psi$	$\delta^{\text{nr}}$	$\delta^{\text{cd}}$	$\delta^r$	$\Theta_{*}^{\text{x,cbi}}$	$\Theta_{*}^{\text{x,kb}}$	$\Theta_{*}^l$
0.998	0.260	5	0.030	0.055	0.004	7.000	7.000	7.000

<sup>8</sup> We draw from the posterior distribution of the DSGE model using MCMC methods. More specifically, we opted for a close version of the random-walk Metropolis algorithm presented in An and Schorfheide (2007). We simulated 75,000 draws from the posterior using the diagonal of the inverse Hessian evaluated at the posterior to calibrate the covariance matrix for the multivariate normal specification of the jump distribution. We scaled the covariance matrix in order to achieve an acceptance rate of about 20%. Finally, we excluded the first 5000 observations as a burn-in period and picked out every eighth draw from the chain. A numerical approximation of the marginal density was obtained using the modified harmonic mean estimator of Geweke (1999).

<sup>9</sup> These choices are consistent with other treatments in the literature and our earlier work; some researchers have disagreed with priors for exogenous structural shocks that assume substantial persistence. We have investigated such alternative priors. Our results regarding forecast accuracy hold (in all cases) for such alternative priors. However, other model

Table II. Prior distributions for model parameters

Param.	Prior type	Prior mean	Prior SD	Param.	Prior type	Prior mean	Prior SD
$h^{\text{cnn}}$	Beta	0.500	0.015	$r^{\Delta\pi}$	Normal	0.000	0.250
$h^{\text{cd}}$	Beta	0.500	0.015	$r^{h,\text{gdp}}$	Normal	0.500	0.250
$h^r$	Beta	0.500	0.015	$r^{\Delta h,\text{gdp}}$	Normal	0.000	0.250
$v$	Gamma	2.000	1.000	$\rho^f$	Beta	0.750	0.013
$\chi^p$	Gamma	2.000	1.000	$\rho^{a,\text{nr}}$	Beta	0.500	0.023
$\eta^p$	Normal	0.000	0.250	$\rho^{a,\text{cd}}$	Beta	0.750	0.013
$\chi^w$	Gamma	2.000	1.000	$\rho^{a,r}$	Beta	0.500	0.013
$\eta^w$	Normal	0.000	0.250	$\rho^{\xi,\text{cnn}}$	Beta	0.750	0.013
$\chi^{\text{nr}}$	Gamma	2.000	1.000	$\rho^{\xi,\text{cd}}$	Beta	0.750	0.013
$\chi^{\text{cd}}$	Gamma	2.000	1.000	$\rho^{\xi,r}$	Beta	0.750	0.013
$\chi^r$	Gamma	6.000	1.000	$\rho^{\xi,l}$	Beta	0.750	0.013
$\chi^l$	Gamma	2.000	1.000	$\rho^{\gamma,m}$	Beta	0.500	0.023
$\eta^l$	Normal	0.000	0.250	$\rho^{\gamma,\text{kb}}$	Beta	0.750	0.013
$r^\pi$	Normal	1.500	0.250	$\rho^{x,\text{gf}}$	Beta	0.750	0.013

Table III. Prior distributions for standard deviations.  $\text{ME}_j$  refers to the standard deviation of the measurement error associated with observable variable  $j$ 

Param.	Prior type	Prior mean	Prior SD	Param.	Prior type	Prior mean	Prior SD
$\sigma_{\xi,\text{cnn}}$	Inv. Gamma	3.000	2.000	$\text{ME}_1$	Inv. Gamma	0.500	2.000
$\sigma_{\xi,\text{cd}}$	Inv. Gamma	3.000	2.000	$\text{ME}_2$	Inv. Gamma	0.500	2.000
$\sigma_{\xi,r}$	Inv. Gamma	3.000	2.000	$\text{ME}_3$	Inv. Gamma	0.500	2.000
$\sigma_{\xi,l}$	Inv. Gamma	3.000	2.000	$\text{ME}_4$	Inv. Gamma	0.500	2.000
$\sigma_{a,\text{cd}}$	Inv. Gamma	2.000	2.000	$\text{ME}_5$	Inv. Gamma	0.500	2.000
$\sigma_{a,r}$	Inv. Gamma	4.000	2.000	$\text{ME}_6$	Inv. Gamma	0.500	2.000
$\sigma_{a,\text{nr}}$	Inv. Gamma	4.000	2.000	$\text{ME}_7$	Inv. Gamma	0.500	2.000
$\sigma_{\gamma,m}$	Inv. Gamma	0.500	2.000	$\text{ME}_8$	Inv. Gamma	0.500	2.000
$\sigma_{\gamma,\text{kb}}$	Inv. Gamma	0.500	2.000	$\text{ME}_9$	Inv. Gamma	0.500	2.000
$\sigma_{\theta,m}$	Inv. Gamma	0.500	2.000	$\text{ME}_{10}$	Inv. Gamma	0.500	2.000
$\sigma_{\theta,\text{kb}}$	Inv. Gamma	0.500	2.000	$\text{ME}_{11}$	Inv. Gamma	0.500	2.000
$\sigma_r$	Inv. Gamma	0.200	2.000	$\text{ME}_{12}$	Inv. Gamma	0.500	2.000

In addition to the DSGE model, we consider a number of other reduced-form models in our forecasting exercises below. Before turning to those exercises, the remainder of this section reports different measures of fit for our DSGE model and different specifications of (Bayesian) vector autoregressions (BVAR) in the 12 variables we consider. Given the relatively large size of our system, our BVAR follows the approach of Banbura *et al.* (2008) for the specification of the prior distribution over parameters.<sup>10</sup> However, their prior distribution of the covariance matrix of the

properties are sensitive to such choices. Such sensitivity is unavoidable in these types of analyses, where the data do not provide much information in some cases.

<sup>10</sup> Specifically, we adopted their dummy observation approach where the data of the VAR system are augmented with artificial observations in order to impose the prior distribution. The scale coefficients used to set up the prior were obtained by running univariate autoregressions of order 4 for each variable of the system. A key parameter in this approach is a tightness parameter. We experimented with different values. The only relevant results are those with moderate to significant shrinkage; loose prior distributions produced results that are similar to those associated with unrestricted VARs. We report the case for moderate shrinkage (a tightness parameter of 0.2 in the framework of Banbura *et al.*, 2008). In addition, one technical issue arose in our forecasting exercises, which involve a large number of different samples: some vintages

errors is improper, thereby making it impossible for us to compute the marginal likelihood. As an alternative for the model comparison exercise, we compute the marginal data density of a Bayesian VAR model using a normal Wishart prior distribution (as in Fernandez-Villaverde and Rubio-Ramirez, 2004; Kadiyala and Karlsson, 1997). The key difference between the two prior distributions is the characterization of the covariance matrix.

Table IV reports the marginal likelihood of the DSGE model and Bayesian VAR models as well as the Akaike (AIC) and Schwarz (SIC) information criteria for different lag orders of the VAR model (all for the sample period 1984Q4 to 2004Q4). The Bayesian model comparison exercise indicates the DSGE model outperforms the BVAR models. The good performance of the structural model can be explained by the presence of measurement errors—in addition to structural shocks—in its state-space form representation. Despite some differences in the characterization of the prior distributions, we rely on these results to motivate our pick of a lag order of two for our subsequent out-of-sample forecasting analysis of BVAR models. For the VAR model the statistical criteria present a different take on the optimal lag order. The AIC prefers a more complex specification of the model, while the SIC criterion favors a more parsimonious parametrization. These results are consistent with previous findings based on simulated and historical data (see Koehler and Murphree, 1988). Because the literature commonly suggests that the SIC is preferable, we have opted for a lag order of one in the case of the VAR model. A small lag order also reduced the frequency with which the VAR parameter estimates implied a non-stationary system, which we deemed invalid for forecasting. When the least-square parameter estimates of the VAR specification were inconsistent with a stationary representation, we estimated instead a Bayesian specification with a minimal amount of shrinkage that delivered a stationary representation.

#### 4. ALTERNATIVE FORECASTS

We compare the forecasts from our DSGE model with four alternatives: the Federal Reserve Board's staff's judgemental projection for FOMC meetings, commonly called the Greenbook projection, the FRB/US model projection, and two reduced-form vector-autoregressive models.

##### 4.1. The Greenbook Forecast

The first set of forecasts that we compare with our DSGE model projection are those produced by the staff at the Federal Reserve Board. The Federal Open Market Committee (FOMC) meets

Table IV. Statistical criteria for optimal lag length

Model	Marginal likelihood	Model	AIC	SIC
DSGE	−753.5			
BVAR(1)	−824.5	VAR(1)	867.9	943.5
BVAR(2)	−818.7	VAR(2)	836.2	981.7
BVAR(3)	−819.2	VAR(3)	797.5	1012.8

would deliver explosive posterior estimates for a given calibration of the tightness parameter. When this happened, we iteratively reduced the value of the tightness parameter by small increments until the weight on the prior distribution was large enough that the estimation of the model delivers a stationary specification.



eight times a year at slightly irregularly spaced intervals. In the lead-up to each of these meetings, the staff at the Board of Governors put together a detailed forecast of the economic outlook that is published (usually a bit less than a week before the FOMC meeting) in a document unofficially known as the Greenbook. The Greenbook forecast, which is most readily available on the web site of the Federal Reserve Bank of Philadelphia, reflects the views of the staff and not the committee members.

As we now explain, the maximum projection horizon for the Greenbook forecast vintages is not constant and varies from ten to six quarters during the course of a year. In September of each year, the staff extend the forecast to include the year following the next in the projection period. Since the third quarter is not yet finished at the time of the September forecast, that quarter is included in the Greenbook projection horizon, generating a maximum horizon of ten quarters. The end-point of the projection horizon remains fixed for subsequent forecasts as the starting point moves forward. As a result, by the July/August forecast round of the following year the projection period extends out only six quarters. In our analysis, we consider a maximum forecast horizon of eight quarters because the number of observations for nine and ten quarters is very small. Note also that the nature of the Greenbook forecast horizon implies that the number of observations for a forecast horizon of eight quarters will be smaller than the number of observations for horizons of six quarters and less. Specifically, of the eight Greenbook projections prepared each year, only five—that is, those prepared for the September, November, December, January, and March FOMC meeting—include forecasts that extend for eight quarters. In contrast, all eight projections prepared each year include forecasts that extend six quarters or less. In the exercises that we undertake in Section 5 (and present in Tables V–IX), when comparing eight-quarter-ahead projections, we only consider forecasts (and forecast errors) generated by the alternative models that correspond to the September, November, December, January, and March Greenbook. We also report the number of forecast observations that we are using in each case.

We use the forecasts produced for the FOMC meetings starting in September 1996 and ending in December 2002; this period includes the beginning of the period over which the FRB/US model (discussed below) has been employed. We choose December 2002 as the end-point because Greenbook forecasts are made public only with a 5-year lag, so forecasts through 2002 are the most recent vintage that is publicly available. The Appendix provides detailed information on the dates of Greenbook forecasts we use and the horizons covered in each forecast. One important aspect of our analysis is that we link our forecast timing to the timing of FOMC meetings. As a result, we will compare eight forecasts a calendar year, and the ‘real-time’ jumping-off point for these forecasts is somewhat irregular. All of our model and forecast comparisons will use the databases employed by the Federal Reserve staff in ‘real time’; this includes our comparison to time series methods, which we can extend through forecasts generated with data available as of November 2004.

#### **4.2. The FRB/US Model Forecast**

The Greenbook projection is a judgmental projection that is derived from quantitative data, qualitative information, different economic models, and various forecasting techniques; it is not the output of any one model. The second set of forecasts that we compare with our DSGE model projection are those produced by the Federal Reserve’s FRB/US model, which is one of the tools that is used as an input into the deliberations that lead to the Greenbook projection. These model

Table V. RSMEs of models: September 1996–November 2004

Model		1Q	2Q	3Q	4Q	8Q
<i>Real GDP growth</i>						
AR(2)		0.470	0.521	0.497	0.547	0.551
Relative RMSE:	DSGE/Edo	0.953	0.920**	0.948	0.917*	0.976
	VAR(1)	1.125*	1.052	1.121**	1.018	1.073
	BVAR(2)	1.096**	1.031	1.071	1.002	1.078
<i>GDP price inflation</i>						
AR(2)		0.276	0.258	0.243	0.281	0.288
Relative RMSE:	DSGE/Edo	1.064	1.065*	1.061	1.038	0.925
	VAR(1)	1.139	1.175**	1.222**	1.231**	1.258
	BVAR(2)	1.088	1.137*	1.150	1.192*	1.177
<i>Hours per capita</i>						
AR(2)		3.128	3.599	4.099	4.442	5.974
Relative RMSE:	DSGE/Edo	1.007	0.997	0.985	0.957	0.846
	VAR(1)	0.994	1.019	1.036	1.055	1.033
	BVAR(2)	1.012	1.029	1.035	1.038	1.004
<i>Nominal funds rate</i>						
AR(2)		0.170	0.260	0.339	0.426	0.618
Relative RMSE:	DSGE/Edo	1.224**	1.208**	1.153	1.092	0.941
	VAR(1)	0.986	1.068	1.116*	1.144*	1.248*
	BVAR(2)	1.006	1.054	1.084	1.104	1.058
Number of observations		66.000	66.000	66.000	66.000	43.000

Note: Asterisks denote significance at the

\*\* 1% and \* 5% level, respectively.

forecasts are prepared at the same time as each Greenbook forecast is prepared and also have the same projection horizon as each Greenbook forecast. The FRB/US model forecasts conditions on the same path for the federal funds rate used in the Greenbook projection, so all statistics related to the federal funds rate in our comparisons are identical between the Greenbook and FRB/US forecasts.

With regard to model structure, the FRB/US model differs significantly from Edo and similar DSGE models. First, while the FRB/US model's equations for most economic decisions are related to those based on explicit optimization as in the Edo model, *ad hoc* elements are introduced to the model to improve model fit in many cases. In addition, the specification of FRB/US has largely proceeded along an 'equation-by-equation' route, with only a small role in estimation for full-model dynamics—a feature that has been criticized for an insufficient attention to system properties and econometric rigor (see Sims, 2002, 2008). Finally, expectations in forecasting exercises using FRB/US are not 'rational' or 'model-consistent', but instead are based upon least-squares projection rules estimated using data realizations over the last several decades.

#### 4.3. Forecasts Generated by Reduced-Form Models

We consider the forecasts generated by two variants of reduced-form vector-autoregressive (VAR) models. The first model is a one-lag VAR system in the 12 variables used to estimate the Edo model. The second model is a two-lag BVAR that introduces onto the coefficients a modified

Table VI. Mean bias of models: September 1996–November 2004

Model	1Q	2Q	3Q	4Q	8Q
<i>Real GDP growth</i>					
AR(2)	−0.037	−0.116	−0.056	−0.140	−0.175
DSGE/Edo	0.024	−0.050	−0.011	−0.104	−0.165
VAR(1)	−0.074	−0.091	−0.021	−0.088	−0.122
BVAR(2)	−0.040	−0.088	−0.021	−0.094	−0.124
<i>GDP price inflation</i>					
AR(2)	0.075	0.087	0.104	0.126	0.092
DSGE/Edo	0.072	0.094	0.108	0.125	0.079
VAR(1)	0.044	0.023	0.024	0.026	−0.038
BVAR(2)	0.032	0.020	0.016	0.021	−0.052
<i>Hours per capita</i>					
AR(2)	2.593	2.806	3.100	3.143	3.348
DSGE/Edo	2.592	2.834	3.146	3.193	3.558
VAR(1)	2.639	2.923	3.327	3.479	4.096
BVAR(2)	2.669	2.971	3.361	3.501	4.106
<i>Nominal funds rate</i>					
AR(2)	0.050	0.089	0.129	0.178	0.309
DSGE/Edo	0.086	0.152	0.202	0.251	0.336
VAR(1)	0.048	0.087	0.127	0.177	0.282
BVAR(2)	0.057	0.099	0.138	0.185	0.255
Number of observations	66.000	66.000	66.000	66.000	43.000

Note: A positive value indicates that a variable's forecast is on average over-predicting its realized value.

version of the dummy-observation prior outlined earlier. The key features of their specifications were motivated in Section 3. We re-estimate these models each forecast.

Readers will likely recognize that the data in our model are released with different delays. For example, interest rate data are available daily and immediately while GDP data—or more specifically, the NIPA—in the USA are first released about a month after a quarter ends. For the most part, we do not account for these differences and ignore any data that are available at the time of a NIPA release but pertain to periods subsequent to the reference quarter of the NIPA data.<sup>11</sup> As a result, our information set in the estimation of the reduced-form models is, in this regard, sparser than the true real-time data actually used in the Federal Reserve staff forecasts.<sup>12</sup>

<sup>11</sup> The exceptions to this, however, are all January Greenbooks and the October 2003 Greenbook. All January Greenbooks close before the BEA has released its advance estimate of the previous year's fourth-quarter NIPA; the October 2003 Greenbook also closed before the BEA released its advance estimate of the third quarter 2003 NIPA. In these cases we use the Federal Reserve staff's best estimate of the preceding quarter's NIPA because almost all of the source data used for the advance release of the NIPA is available at the time that Greenbook closes.

<sup>12</sup> An example may help the reader understand the timing conventions used in our forecast exercises. The September 1996 FOMC meeting occurred on September 24, and the Greenbook (staff forecast) prepared for that meeting was published on September 18. At that time, the last quarter with a complete historical dataset was the second quarter of 1996; we call this the 'last quarter of history'. As a result, the first forecast quarter in our forecast exercise is the third quarter of 1996 (that is, the first forecast quarter, 1996Q3, is the 'current quarter', as September falls in the third quarter). The complete forecast horizon for the September 1996 Greenbook is 1996Q3 to 1998Q4 (ten quarters), and we call this the 'Greenbook forecast horizon'. The Appendix on Greenbook forecasts and real-time data presents a complete description of the dates of FOMC meetings, staff forecast (Greenbook) publication, the last quarter of history, the estimation sample for our models, and the forecast horizon.

Table VII. RSMEs of models, disaggregated variables: September 1996–November 2004

Model		1Q	2Q	3Q	4Q	8Q
<i>Real cons. growth, nondurables</i>						
AR(2)		0.343	0.348	0.371	0.363	0.326
Relative RMSE:	DSGE/Edo	1.000	1.062	0.967	1.029	1.159**
	VAR(1)	1.211	1.237**	1.138**	1.121*	1.166
	BVAR(2)	1.035	1.067*	1.048	1.114	1.183
<i>Real cons. growth, durables</i>						
AR(2)		2.773	2.642	2.379	2.897	3.076
Relative RMSE:	DSGE/Edo	0.975	0.988	0.981	0.979	0.989
	VAR(1)	1.018	1.040	1.128*	1.033	1.026
	BVAR(2)	0.999	1.043	1.057	1.012	1.021
<i>Real inv. growth, non-residential</i>						
AR(2)		3.976	3.805	3.586	3.804	3.534
Relative RMSE:	DSGE/Edo	0.904*	0.898**	0.892*	0.902*	0.952
	VAR(1)	1.050	1.043	1.010	1.019	1.047
	BVAR(2)	0.984	1.059*	1.008	1.006	1.039
<i>Real inv. growth, residential</i>						
AR(2)		2.080	1.932	1.967	2.057	2.380
Relative RMSE:	DSGE/Edo	1.149**	1.471**	1.705**	1.772**	1.650**
	VAR(1)	0.996	1.224*	1.496*	1.463*	1.405*
	BVAR(2)	0.914	1.123	1.379*	1.336*	1.212
Number of observations		66.000	66.000	66.000	66.000	43.000

Note: Asterisks denote significance at the

\*\* 1% and \* 5% level, respectively.

#### 4.4. Generating Real-Time Forecasts

An accurate comparison of the performance of different forecasts requires the use of real-time data. The Federal Reserve Board's Greenbook and FRB/US model projections are real-time forecasts as they are archived when the Greenbook forecast is closed.

Since March 1996 the staff have stored the Greenbook projection from each FOMC forecasting round in readable electronic databases that contain the level of detail needed for a rich DSGE model like Edo. Importantly for the purposes of this research, these databases also include historical data for the data series the staff forecast that extend back to about 1975. Because these databases were archived at the time that each particular Greenbook forecast was closed, the historical data from these databases represent the real-time data available to the staff at the time that they were preparing their forecast. Consequently, we estimate our DSGE and time series models with historical data from these historical Greenbook databases, on the assumption that the Greenbook forecasts were generated using the same information sets. Constructing real-time datasets on which to estimate our DSGE and atheoretic models simply involves pulling the relevant series, reported earlier in our description of the series used to estimate our DSGE model, from the Greenbook database. As with the reduced-form models, we do not account for differences in model data availability schedules such that the information set for estimating the DSGE model is sparser than the 'real-time' data used in the Federal Reserve staff forecasts.

In principle, the construction of real-time forecasts from the DSGE model presents no additional difficulties. In practice, however, some issues arise. The DSGE model involves modeling the joint

Table VIII. RSMES of models: September 1996–December 2002

Model		1Q	2Q	3Q	4Q	8Q
<i>Real GDP growth</i>						
AR(2)		0.478	0.480	0.556	0.607	0.602
Relative RMSE:	DSGE/Edo	0.943	0.874**	0.938	0.901*	0.970
	VAR(1)	1.093*	1.127*	1.133**	1.031	1.100
	BVAR(2)	1.097**	1.084*	1.082*	1.013	1.105
	Greenbook	1.153	1.299*	1.165*	1.189*	1.104
	FRB/US	1.066	1.408**	1.163**	1.158*	1.138*
<i>GDP price inflation</i>						
AR(2)		0.289	0.280	0.256	0.301	0.303
Relative RMSE:	DSGE/Edo	1.028	1.041	1.043	1.052	0.916
	VAR(1)	1.114	1.110	1.122	1.109	1.024
	BVAR(2)	1.066	1.097	1.059	1.093	0.981
	Greenbook	1.063	0.835	0.752**	0.701**	0.934
	FRB/US	0.941	1.023	0.947	0.918	0.865
<i>Hours per capita</i>						
AR(2)		3.039	3.744	4.386	4.847	6.843
Relative RMSE:	DSGE/Edo	0.991	0.981	0.967	0.947	0.850
	VAR(1)	0.995	1.010	1.030	1.047	1.038
	BVAR(2)	1.005	1.015	1.018	1.026	1.007
	Greenbook	1.004	1.018	1.020	1.024	0.983
	FRB/US	0.992	0.998	1.001	1.009	0.929
<i>Nominal funds rate</i>						
AR(2)		0.186	0.283	0.368	0.464	0.687
Relative RMSE:	DSGE/Edo	1.204*	1.173*	1.124	1.073	0.963
	VAR(1)	0.961	1.049	1.100*	1.120*	1.167
	BVAR(3)	1.015	1.063	1.100	1.114	1.023
	Greenbook	0.743**	0.721**	0.812**	0.888	0.983
	FRB/US	0.743**	0.721**	0.812**	0.888	0.983
Number of observations		51.000	51.000	51.000	51.000	32.000

Note: Asterisks denote significance at the  
 \*\* 1% and \* 5% level, respectively.

stochastic process followed by a large number of variables, which may improve the estimates of underlying structural parameters and hence forecast accuracy. In addition, the solution and estimation of the DSGE model is somewhat more involved than that associated with simple time series regressions (which can be estimated almost instantly in virtually any software package, including even simple spreadsheets). As a result, estimation in the DSGE model is performed using the real-time datasets once per year, specifically in the July/August round in which an annual rebenchmarking of the NIPA data takes place. This contrasts with the approach followed for the VAR forecasts, where re-estimation is performed for each forecast. Parameter estimates for the Edo model are then held constant for the forecasts generated in subsequent rounds until the following July/August, at which point the model is re-estimated using the four additional quarters of data.<sup>13</sup> Note, however, that it is only the data used to estimate the model that remain

<sup>13</sup> With regard to the parameter estimates used to generate our DSGE model forecasts we note that we do not take into account the uncertainty associated with these parameters, which our Bayesian estimation technique would actually leave us very well situated to consider. While we certainly view the uncertainty aspect of the DSGE model's forecast as very important, we do not pursue it in the ensuing analysis.

Table IX. RSMEs of models, disaggregated variables: September 1996–December 2002

Model		1Q	2Q	3Q	4Q	8Q
<i>Real cons. growth, nondurables</i>						
AR(2)		0.352	0.371	0.406	0.397	0.369
Relative RMSE:	DSGE/Edo	0.978	0.964	0.916*	0.958	1.117**
	VAR(1)	1.265	1.245**	1.154**	1.139*	1.178*
	BVAR(2)	1.003	1.058	1.061	1.132*	1.196
	Greenbook	1.031	0.914	0.836**	0.980	1.155*
	FRB/US	0.972	1.048	0.947	1.203*	1.301**
<i>Real cons. growth, durables</i>						
AR(2)		2.936	2.686	2.635	3.099	3.260
Relative RMSE:	DSGE/Edo	0.921*	0.960	0.980	0.973	0.970
	VAR(1)	1.001	1.062	1.099	1.020	1.024
	BVAR(2)	0.951*	1.094*	1.036	0.997	1.016
	Greenbook	1.367**	1.187	1.107*	1.052*	1.029
	FRB/US	1.341**	1.105	0.978	0.959	1.044
<i>Real inv. growth, non-residential</i>						
AR(2)		4.191	4.065	3.610	4.094	3.792
Relative RMSE:	DSGE/Edo	0.879**	0.881**	0.902*	0.909*	0.980
	VAR(1)	1.083	1.070	1.067	1.047	1.092
	BVAR(2)	1.022	1.082*	1.066	1.029	1.082
	Greenbook	1.075	1.135	1.218**	1.157**	1.024
	FRB/US model	1.040	1.234**	1.093*	0.982	1.056
<i>Real inv. growth, residential</i>						
AR(2)		2.098	1.941	1.874	2.227	2.178
Relative RMSE:	DSGE/Edo	1.046	1.241**	1.445**	1.446**	1.605**
	VAR(1)	1.044	1.278*	1.736**	1.493*	1.509*
	BVAR(2)	0.921	1.197*	1.603**	1.370*	1.266
	Greenbook	1.027	1.388**	1.295**	1.191*	1.077
	FRB/US	0.916	1.141	1.073	0.953	1.146
Number of observations		51.000	51.000	51.000	51.000	32.000

Note: Asterisks denote significance at the

\*\* 1% and \* 5% level, respectively.

constant across the forecasts for the year. The ‘jumping-off’ period that is used for each forecast generated by the DSGE model is the staff’s estimate of the last quarter of history taken from the corresponding Greenbook database.

We compute statistics on forecast accuracy by comparing the forecasts based on real-time data to the realizations of these series contained in the data’s ‘first final’ release.<sup>14</sup>

## 5. COMPARISON TO REDUCED-FORM MODEL FORECASTS

We focus on two distinct sets of variables. The first are the ‘top-line’ macroeconomic aggregates—specifically, the percent change in real GDP per capita, GDP price inflation, detrended

<sup>14</sup> In earlier drafts of the paper we compared our real-time forecasts to the realizations of series contained in the most recent vintage of data. We updated these comparisons (not shown) for this version of the paper; in this case using data available at the the October 2007 FOMC meeting. Using for comparison the ‘first final’ realization of the data instead of the most recent vintage alters very little the relative performance of the different forecast models, although it does alter the magnitude of the forecast biases.

hours per capita, and the federal funds rate. The second are the disaggregated categories of expenditure—the percent changes in real personal consumption expenditures on nondurables and services, real personal consumption expenditures on durables, real non-residential investment, and real residential investment. We evaluate forecast accuracy along two dimensions: the absolute size of errors and the bias in errors. We measure the absolute size of errors using the root mean squared error (RMSE), while bias is measured by the mean average error. We use forecasts generated by an AR(2) model as the benchmark against which we compare our model forecasts; this serves as a challenging point of comparison given that univariate models have been documented by several authors to have more accurate forecasting ability than multivariate models (see Atkeson and Ohanian, 2001; D’Agostino *et al.*, 2006; Marcellino *et al.*, 2006).

## 5.1. The Main Macroeconomic Aggregates

The main macroeconomic aggregates examined are real GDP growth, GDP price inflation, detrended hours per capita, and the federal funds rate. This set captures aggregate activity and is the focus of many small modeling efforts. In addition, the focus on this set of variables will link directly to some of the main macroeconomic developments over the 1996–2005 period.

### 5.1.1 Real GDP Growth

The first set of results presented in Table V focuses on the forecasts for real GDP growth. The line labeled AR(2) reports the RMSE at various forecast horizons (specifically, one through four quarters out and eight quarters out), in percentage points, for the forecast of the percent change in GDP per capita generated by the AR(2) model. The RMSEs equal about 1/2 percentage point (not an annual rate). The remainder of the reported figures for GDP growth report the relative RMSEs for the other forecast methods; values below 1 indicate that the model performs better than AR(2) model. The relative RMSE for the DSGE/Edo model is below 1 at each reported horizon, although the only significant differences (at the 5% level) according to the Diebold and Mariano (1995) statistic are associated with the two- and four-quarter-ahead forecasts.<sup>15</sup> (Note that at the one- to four-quarter-ahead horizons there are 66 forecasts, despite the fact that our sample spans less than 10 years, because we produce forecasts on an FOMC meeting basis and there are eight FOMC meetings a year). The VAR(1) and BVAR(2) models tend to perform worse than the AR(2) at the reported horizons, with relative RMSEs exceeding 1; for both models, the difference at the initial horizon is statistically significant.

The results on bias for GDP growth are shown in the upper part of Table VI (positive value shown in the table implies that a variable’s forecast is over-predicting its realized value). The biases associated with each method are small, typically well below 0.1 percentage point (in absolute value), and are negative, indicating a general tendency by all forecast methods to under-predict real GDP growth.

### 5.1.2 GDP Price Inflation

The next set of results reported in Table V focuses on the forecasts for GDP prices. The line for the AR(2) indicates that the RMSE for GDP price inflation is about 0.25 percentage point at each horizon. The comparison across forecast methods indicates that the DSGE/Edo model, the

<sup>15</sup> The standard errors for the Diebold–Mariano statistics are adjusted for heteroskedasticity and autocorrelation (up to the forecast horizon minus one) using the method of Newey and West (1987).

VAR(1), and the BVAR(2) all tend to forecast worse than the AR(2) model; these differences are mostly statistically significant at horizons of four quarters or less for the VAR methods, and the DSGE/Edo deterioration is significant at the two-quarter horizon. The DSGE/Edo model has slightly better relative RMSEs than the VAR methods.

The average biases for GDP price inflation, in Table VI, are quite small and are mostly positive, indicating a tendency to over-predict inflation.

### 5.1.3 *Hours per Capita*

As noted earlier, the state of the labor market in DSGE/Edo is summarized by detrended hours worked per capita, which is calculated by removing the mean of actual hours worked per capita observed over the estimation period. As the state of the labor market is one side of the Federal Reserve's dual mandate of full employment and price stability, the ability of the DSGE/Edo model to forecast hours per capita, relative to the ability of other models, is an important metric for model evaluation. We focus on detrended hours per capita as it is the closest analogue in our model to a notion of slack, such as deviation of the unemployment rate from its natural rate. As will be apparent, errors in forecasts of detrended hours per capita can stem from errors in estimates of the trend or in the forecast going forward.

The third set of lines in Table V reports the RMSE statistics related to detrended hours per capita. These errors are large, between 3 and 6 percentage points at the reported horizons. The large size of these errors reflects the real-time nature of the exercise: historical data on hours per capita can be revised substantially, making forecasting difficult.

The remaining lines in the panel referring to hours per capita report the RMSE for the DSGE/Edo model, the VAR(1), and the BVAR(2) relative to the RMSE for the AR(2) model. Several results are apparent. First, the DSGE/Edo model seems to perform better than the AR(2) almost uniformly, while the BVAR(2) has RMSEs essentially identical to the AR(2). The VAR(1) model performs a bit worse than the AR(2) beyond the two-quarter horizon. However, according to the Diebold and Mariano (1995) test, none of the reported poorer performance of the alternative forecasts reported for hours per capita in Table V are significant in the statistical sense.

Table VI presents the bias of each forecast of hours per capita. In all cases, the bias is positive and large—about 2 percentage points at the one-quarter horizon. The bias in the model forecasts over this period suggests that each of the forecasts tended to systematically overstate the 'tightness' in the labor market over this period. In order to gauge whether the size of the bias and revisions in detrended hours per capita from revisions to trend are reasonable, it is useful to look at related statistics for estimates of the natural rate of unemployment (which has been studied more than hours). For example, the CBO's revision to its 1997Q1 estimate of the natural rate of unemployment (NAIRU) from the January 1997 estimate to the estimate of January 2008 is  $-0.7$  percentage point (from  $5.8\%$  to  $5.1\%$ ) (see Congressional Budget Office, 1997, 2008). An ordinary least squares regression of the unemployment rate on hours per capita yields a coefficient of  $0.3$ ; using this  $0.3$  coefficient to translate the revision in the CBO's natural rate into a revision for the trend in hours yields a figure around 2 percentage points (that is,  $0.7/0.3$ ), suggesting that the revision noted above ( $1.9$  percentage points) and the bias statistics (2 percentage points at the one-quarter horizon) are reasonable. These results on detrended hours per capita highlight again the difficulty that detrending creates for gauging the state of the economy in monetary policy applications, a point forcefully emphasized by Orphanides and van Norden (2003).



#### 5.1.4 *The Federal Funds Rate*

The final set of results reported in Table V focuses on the forecasts for the federal funds rate. The line for AR(2) indicates that the RMSE for the federal funds rate is about 0.2 percentage point at the one-quarter horizon and rises to 0.6 percentage point at the eight-quarter horizon. The DSGE/Edo model performs worse than AR(2) for horizons out to four quarters, but only to a statistically significant degree for the first two quarters. The VAR(1) model outperforms AR(2) at short horizons, but its performance deteriorates at long horizons—where the VAR(1) RMSE exceeds that of the DSGE model.

### 5.2. Disaggregated Measures of Expenditure

Looking underneath aggregate GDP growth provides further insight into forecast performance. Policymakers are often interested in developments within individual sectors, such as the strength of business investment, the state of the housing market, or diverging trends in consumer and business spending (see Kohn, 2003).

We consider the forecast performance of the various methods under consideration for the percent changes in real personal consumption expenditures on nondurables and services, real personal consumption expenditures on durables, real non-residential investment, and real residential investment in Table VII. The structure of the reported statistics is the same as in Table V. We take away two summary points. First, the forecasts of the DSGE/Edo model, as summarized by the RMSE, are more accurate than those of the AR(2), VAR(1) and BVAR(2) for the components of durables consumption and non-residential investment at nearly all horizons, and in some cases (especially for non-residential investment) more accurate by large and statistically significant margins. Second, all the multivariate methods have difficulty forecasting residential investment, and the relative RMSEs of the DSGE/Edo model are larger than those of other models over this period; we return to this finding in the next section.

### 5.3. Summary of Empirical Results

Overall, we have found that our DSGE model provides forecasts for activity and inflation that are competitive with those from univariate and multivariate time series models for a broad range of variables. Nonetheless, the forecasts from our DSGE model, and from the multivariate VAR models, are most often not superior to forecasts from a univariate autoregression by a statistically significant margin even in those cases when the RMSEs are somewhat lower from these multivariate alternatives. As a result, it is not clear from these exercises whether our DSGE model provides information that would prove useful in a policy context. To address this question, the next section examines the policy relevance question by comparing the accuracy of forecasts from our DSGE model relative to the accuracy of Federal Reserve staff forecasts.

## 6. COMPARISON TO FEDERAL RESERVE STAFF FORECASTS

We now present the forecast performance of the DSGE/Edo model along with that of Federal Reserve staff forecasts and forecasts from the Federal Reserve's FRB/US model. We have two goals. First, a comparison to existing methods at the Federal Reserve is more policy relevant than a comparison to AR and VAR methods, in part because Federal Reserve forecasts have not

placed much weight on projections from these types of models. Second, we attempt to identify what features of our model or the data contribute to the successes and failures recorded by the DSGE/Edo model along the forecast dimension from 1996 to 2004, with an eye toward future changes in specification or research projects that attempt to incorporate additional features into our DSGE framework in order to improve its forecast performance and its utility as a policy tool more generally. The public availability of Federal Reserve staff forecasts have led us to focus on comparisons of forecasts using data for FOMC meetings from September 1996 to December 2002.

### 6.1. Forecast Performance

Tables VIII and IX present statistics on forecast accuracy for the projections generated using the data from the September 1996 FOMC meeting to the December 2002 FOMC meeting.

With regard to the labor market, it is apparent that the staff projections in the Greenbook and from the FRB/US model for detrended hours per capita share the dominant feature reported earlier: the errors are very large, exceeding 3 percentage points even at the one-quarter horizon. As we emphasized earlier, this reflects the difficulty of detrending in 'real time'.<sup>16</sup> While the errors are uniformly large across methods, the forecast performance of the DSGE/Edo model dominates that of the Greenbook and FRB/US model at all horizons.

The results for other measures of economic activity are also revealing. The forecast accuracy (in an RMSE sense) of DSGE/Edo is better than the Greenbook projections for GDP growth (Table VIII), the growth of the components of consumption expenditures, and growth of non-residential investment for nearly all horizons (Table IX). However, few of these forecasts are better than the AR(2) forecast in the statistical sense (except for non-residential investment, where the DSGE/Edo forecasts fare well).<sup>17</sup>

The result that DSGE/Edo (or AR(2) models) have similar or lower 'real-time' out-of-sample RMSEs for many real activity measures may be surprising, especially at short horizons, where the Federal Reserve staff devote significant resources to assessing near-term developments (see Romer and Romer, 2000; Sims, 2002). We think this is a significant finding. As we have emphasized in previous work (for example Edge *et al.*, 2008), the ability of a structural model like our DSGE model to tell economically meaningful stories can make such models more attractive in a policy context than time series alternatives, and the additional result that forecast performance may be acceptable as well adds further support to the consideration of such tools.

It is also interesting to note that the DSGE/Edo model and the Greenbook made large errors in their forecasts of residential investment, with the DSGE/Edo model doing poorly relative to the Greenbook at the eight-quarter horizon. This performance will be one of the topics discussed in the next subsection.

Returning to Table VIII, the results for GDP price inflation continue to suggest that the DSGE/Edo model is competitive with best practices. In particular, the DSGE/Edo model has RMSEs for inflation that are lower than those of the Greenbook at some (that is, the one- and eight-quarter) horizons and greater than those of the Greenbook at other. Sims (2002) reported that the near-term inflation forecasts in the Greenbook were very good, so the competitive performance

<sup>16</sup> Greenbook forecasts do not include a projection for detrended hours per capita; as detrending simply involves removing the mean for this series, we compute the Greenbook and FRB/US model forecasts by removing the mean from 1984Q4 to the last period that is forecast.

<sup>17</sup> Note that the Diebold–Mariano statistics for each row use the AR(2) as the baseline model.

of the DSGE model even at such short horizons provides a signal that this type of model may provide valuable additional information in the inflation forecasts at the Federal Reserve. And such forecasts may be quite important: the dual mandate has price stability as one objective, and many discussions of monetary policy emphasize the importance of inflation forecasts in the setting of monetary policy.

Finally, the results for the federal funds rate show that the staff Greenbook projection is better than the DSGE/Edo or AR(2) projections by a large and statistically significant margin (with relative RMSEs around 0.7) for the period spanning the September 1996 to December 2002 Greenbooks. We will discuss this finding in the next subsection.

## 6.2. Implications

We interpret the entire set of results in three ways. First, the performance of the Edo model in explaining labor market developments seems competitive with other approaches. Nonetheless, the forecast errors for hours per capita are large and the bias has been significant over the 1996–2004 period. As a result, we view as a high priority efforts to model the labor market in a more nuanced way. Such efforts include an allowance for factors that may improve detrending—such as factors that allow for permanent shocks to households' supply of hours that would be estimable via the Kalman filter and may reflect economic factors like demographics—as well as including in the Edo model features (such as those considered by Gertler *et al.*, 2008) that allow for the adjustment of hours along both the intensive and extensive margins.

It is also interesting that the Edo model forecasts poorly the nominal federal funds rate, especially relative to the staff projections from the Federal Reserve. The performance relative to the Federal Reserve staff projection may reflect an information advantage; that is, that the staff has insight into the likely course of policy from interactions with policymakers. However, the Edo model also had a somewhat worse forecast performance for the federal funds rate than the VAR methods, and perusal of full-sample VAR parameter estimates suggests a possible reason: lags of (log) hours per capita have sizable and statistically significant coefficients in a reduced-form (VAR model) federal funds rate regression over our sample period, and this relationship appears tighter than that to real GDP growth. Such a relationship seems reasonable, since, as we emphasized earlier, the dual mandate of the Federal Reserve includes full employment, and hours per capita are the closest analogue in the first generation of the Edo model to the deviation of unemployment from its natural rate. As a result of these findings, we have explored different policy rule specifications and, in particular, we have emphasized the role for hours per capita in subsequent research (see Edge *et al.*, 2007b).

The most notable other aspect of the results for economic activity is the poor forecast performance for residential investment. The pace of growth in residential investment over 2001–2004 was extraordinary (although, of course, residential investment weakened substantially after this period). Indeed, even around the recession of 2001, residential investment was not as weak as is typical during economic downturns. Factors such as greater availability of mortgage finance may have been one factor influencing residential investment over this period, as may also have been behavioral factors, such as speculative investment. We view structural investigations of these issues in general equilibrium models as an interesting topic for research.

In addition, the relatively poor performance in forecasting residential investment, compared to the very good performance forecasting non-residential investment, may reflect the attention to the

modeling of non-residential investment in the Edo model. As we emphasized in the Introduction, we adopted a two-sector growth model with fast technological progress in the sector producing non-residential investment (and consumer durable) goods in order to match the steady-state growth facts across different expenditure categories. In doing so, we built on a literature developed in the second half of the 1990s (see Greenwood *et al.*, 2000; Whelan, 2003). It is certainly possible, as we emphasized in the Introduction, that our use of ‘real-time’ data has insufficiently controlled for the influence of developments leading up to the Edo model and that our forecast performance is aided by the fact that we have specified our model after the data have been realized. While it is impossible to remove the effect of such influences from our analysis, the fact that residential investment is forecast less well and non-residential investment is forecast well may reflect the focus on non-residential investment in our model development.

## 7. CONCLUSIONS

Our goal has been to provide a comparison between forecasts from a richly specified DSGE model with those from time series alternatives and the staff forecasts of the Federal Reserve. Our analysis has demonstrated that DSGE models with rich detail on domestic production and spending decisions can provide forecasts of macroeconomic aggregates that are competitive with the approaches used in central banks.

We take several lessons from these findings both for policy-related analyses and future research. Most importantly, the finding that a richly specified DSGE model is competitive with reasonable forecast alternatives provides support for the use of such models in macroeconomic forecasting. Taken together with the fact that the Edo model has also proved useful in answering policy questions, this result suggests a significant and broad role for richly specified DSGE models in the toolbox of central bank models. We also suspect that our findings provide interesting clues regarding the structure of the economy that may help inform monetary policy. For example, DSGE models like Edo have a structure that implies a very important role for fluctuations in technology, or productivity, in the business cycle, whereas more traditional models at central banks like the FRB/US model give fluctuations in technology a smaller role. Another example may relate to inflation, where the Edo model provides good forecasts. Some research has been very critical of New-Keynesian models of the Phillips curve (notably Rudd and Whelan, 2007), but the forecast success reported herein suggests a dimension of empirical validation for such models that has not been previously emphasized. Finally, our discussion has highlighted two areas where results suggest that further research and perhaps amendments to the structure of models like Edo are warranted. The first was the structure of the labor market, including modeling of the intensive and extensive margin and, perhaps, an explicit role for search or other frictions. The second was the role of financial innovation in the rise, and subsequent fall, of residential investment after 2001.

## ACKNOWLEDGEMENTS

Rochelle M. Edge Michael T. Kiley, and Jean-Philippe Laforte are affiliated with the Board of Governors of the Federal Reserve System. For helpful comments we would like to thank participants at conferences and workshops hosted by the Bank of Canada, the Central Bank of Chile, the European Central Bank, the Bank of Finland, the Reserve Bank of New Zealand, the Swiss National Bank, CIRANO and the University of Toronto Institute for Policy Analysis,

Dynare, the Euro Area Business Cycle Network, the NBER Summer Institute and the Society for Computational Economics. This version of the paper has benefited substantially from the comments and suggestions of two anonymous referees. The views expressed do not necessarily reflect those of the Board of Governors of the Federal Reserve System or its staff. All errors are our own.

## REFERENCES

- Adolfson M, Lindé J, Villani M. 2006. *Forecasting performance of an open economy dynamic stochastic general equilibrium model*. Sveriges Riksbank Working Paper Series No. 190.
- Adolfson M, Andersson M, Lindé J, Villani M, Vredin A. 2007. Modern forecasting models in action: improving macroeconomic analyses at central banks. *International Journal of Central Banking* **3**: 111–144.
- Altig D, Christiano L, Eichenbaum M, Linde J. 2005. Firm-Specific Capital, Nominal Rigidities and the Business Cycle, NBER Working Papers 11034, National Bureau of Economic Research, Inc.
- An S, Schorfheide F. 2007. Bayesian analysis of DSGE models. *Econometric Reviews* **26**: 205–210.
- Atkeson A, Ohanian L. 2001. Are Phillips curves useful for forecasting inflation? *Federal Reserve Bank of Minneapolis Quarterly Review* **2**: 2–11.
- Banbura M, Giannone D, Reichlin L. 2008. *Large Bayesian VARs*. ECB Working Paper Series No. 966.
- Bayoumi T, Laxton D, Faruquee H, Hunt B, Karam P, Lee J, Rebucci A, Tchakarov I. 2004. *GEM: a new international macroeconomic model*. IMF occasional paper No. 239.
- Brayton F, Levin A, Tryon R, Williams J. 1997. The evolution of macro models at the Federal Reserve Board. *Carnegie-Rochester Conference Series on Public Policy* **47**: 43–81.
- Christiano L, Eichenbaum M, Evans C. 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy* **113**: 1–45.
- Christoffel K, Coenen G, Warne A. 2007. *The new area-wide model of the Euro Area: specification, estimation results and properties*. Manuscript, European Central Bank.
- Congressional Budget Office. 1997. *The economic and budget outlook: fiscal years 1998–2007*. A CBO Report.
- Congressional Budget Office. 2008. *The economic and budget outlook: fiscal years 2008–2018*. A CBO Report.
- D'Agostino A, Giannone D, Surico P. 2006. *(Un)predictability and macroeconomic stability*. ECB Working Paper Series No. 605.
- Del Negro M, Schorfheide F. 2004. Priors from general equilibrium models for VARs. *International Economic Review* **45**: 643–673.
- Diebold F, Mariano R. 1995. Comparing predictive accuracy. *Journal of Business and Economic Statistics* **13**: 253–263.
- Edge R, Kiley M, Laforde JP. 2007a. *Documentation of the Research and Statistics Division's estimated DSGE model of the U.S. economy*. Federal Reserve Board FEDS Working Paper Series, No. 2007-53.
- Edge R, Kiley M, Laforde JP. 2007b. *The sources of fluctuations in residential investment*. Manuscript Federal Reserve Board, Washington, DC.
- Edge R, Kiley M, Laforde JP. 2008. An estimated DSGE model of the U.S. economy with an application to natural rate measures. *Journal of Economic Dynamics and Control* **32**: 2512–2535.
- Erceg C, Guerrieri L, Gust C. 2006. SIGMA: a new open economy model for policy analysis. *International Journal of Central Banking* **2**: 1–50.
- Faust J, Wright J. 2007. *Comparing Greenbook and reduced form forecasts using a large realtime dataset*. NBER Working Paper Series No. W13397.
- Fernandez-Villaverde J, Rubio-Ramirez J. 2004. Comparing dynamic equilibrium models to data: a Bayesian approach. *Journal of Econometrics* **123**: 153–187.
- Gali J, Gertler M. 2007. Macroeconomic modeling for monetary policy evaluation. *Journal of Economic Perspectives* **21**: 25–45.
- Gertler M, Sala L, Trigari A. 2008. An estimated monetary DSGE model with unemployment and staggered nominal wage bargaining. *Journal of Money, Credit and Banking* **40**: 1713–1764.
- Geweke J. 1999. Using simulation methods for Bayesian econometric models: inference, development and communication. *Econometric Reviews* **18**: 1–126.

- Gordon R. 1998. Foundations of the Goldilocks economy: supply shocks and the time-varying NAIRU. *Brookings Papers on Economic Activity* No. 2 297–333.
- Greenwood J, Hercowitz Z, Krusell P. 2000. The role of investment-specific technological change in the business cycle. *European Economic Review* **44**: 91–115.
- Kadiyala R, Karlsson S. 1997. Numerical methods for estimation and inference in Bayesian VAR-models. *Journal of Applied Econometrics* **12**: 99–132.
- Kiley M. 2007. A quantitative comparison of sticky-price and sticky-information models of price setting. *Journal of Money, Credit, and Banking* **39**: 101–125.
- Koehler A, Murphree E. 1988. A comparison of the Akaike and Schwarz criteria for selecting model order. *Applied Statistics* **37**: 187–195.
- Kohn D. 2003. The strength in consumer durables and housing: policy stabilization or a problem in the making? Remarks at the *Conference on Finance and Macroeconomics*, sponsored by the Federal Reserve Bank of San Francisco and Stanford Institute for Economic Policy Research, San Francisco, CA, 28 February.
- Kydland F, Prescott E. 1982. Time-to-build and aggregate fluctuations. *Econometrica* **50**: 1345–1370.
- Laforte JP. 2007. Pricing models: a Bayesian DSGE approach to the U.S. economy. *Journal of Money, Credit, and Banking* **39**: 127–154.
- Langdon D, Krantz R, Strole M. 2004. Post-recession trends in nonfarm employment and related economic indicators. *Monthly Labor Review* September: 49–56.
- Lees K, Matheson T, Smith C. 2007. *Open economy DSGE-VAR forecasting and policy analysis: head-to-head with the RBNZ published forecasts*. Reserve Bank of New Zealand Discussion Paper Series No. 2007-01.
- Marcellino M, Stock J, Watson M. 2006. A comparison of direct and iterated multistep AR methods for forecasting macroeconomic time series. *Journal of Econometrics* **135**: 499–526.
- Meyer L. 1997. The role for structural macroeconomic models. Remarks at the *American Economic Association Panel on Monetary and Fiscal Policy*, New Orleans, LA.
- Newey W, West K. 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* **55**: 703–708.
- Orphanides A, van Norden S. 2003. The unreliability of output gap estimates in real time. *Review of Economics and Statistics* **84**: 569–583.
- Reifschneider D, Tulip P. 2007. *Gauging the uncertainty of the economic outlook from historical forecasting errors*. Federal Reserve Board FEDS Working Paper Series No. 2007-60.
- Reifschneider D, Stockton D, Wilcox D. 1997. Econometric models and the monetary policy process. *Carnegie-Rochester Conference Series on Public Policy* **47**: 1–37.
- Romer C, Romer D. 2000. Federal Reserve information and the behavior of interest rates. *American Economic Review* **90**: 429–457.
- Rudd J, Whelan K. 2007. Modeling inflation dynamics: a critical survey of recent research. *Journal of Money, Credit, and Banking* **39**: 155–170.
- Sims C. 2002. The role of models and probabilities in the monetary policy process. *Brookings Papers on Economic Activity* **33**: 1–62.
- Sims C. 2003. *Probability models for monetary policy decisions*. Manuscript Princeton University, Princeton, NJ.
- Sims C. 2008. Improving monetary policy models. *Journal of Economic Dynamics and Control* **32**: 2460–2475.
- Smets F, Wouters R. 2007. Shocks and frictions in the U.S. business cycles: a Bayesian DSGE approach. *American Economic Review* **97**: 586–606.
- Tulip P. 2005. *Has output become more predictable? Changes in Greenbook forecast accuracy*. Federal Reserve Board FEDS Working Paper Series No. 2005-31.
- Whelan K. 2003. A two-sector approach to modelling U.S. NIPA data. *Journal of Money, Credit, and Banking* **35**: 627–656.

## APPENDIX: GREENBOOK FORECASTS AND REAL-TIME DATA

## A.I. Greenbook and NIPA release dates (September 1996–December 1997)

GB name	GB closed	FOMC meeting	Estim. period, DSGE model	Last qtr of history	GB Forecast horizon	Interim NIPA releases
Sep. 96	9/18/96	9/24/96	85:Q1–96:Q2	96:Q2	96:Q3–98:Q4	96:Q2 Fin.: 9/27/96 96:Q3 Adv.: 10/30/96
Nov. 96	11/6/96	11/13/96	85:Q1–96:Q2	96:Q3	96:Q4–98:Q4	96:Q3 Pre.: 11/27/96
Dec. 96	12/12/96	12/17/96	85:Q1–96:Q2	96:Q3	96:Q4–98:Q4	96:Q3 Fin.: 12/20/96
Jan. 97	1/29/97	2/4–5/97	85:Q1–96:Q2	96:Q4	97:Q1–98:Q4	96:Q4 Adv.: 1/31/97 96:Q4 Pre.: 2/28/97
Mar. 97	3/19/97	3/25/97	85:Q1–96:Q2	96:Q4	97:Q1–98:Q4	96:Q4 Fin.: 3/28/97 97:Q1 Adv.: 4/30/97
May 97	5/15/97	5/20/97	85:Q1–96:Q2	97:Q1	97:Q2–98:Q4	97:Q1 Pre.: 5/30/97
Jun. 97	6/25/97	7/1–2/97	85:Q1–96:Q2	97:Q1	97:Q2–98:Q4	97:Q1 Fin.: 6/27/97 97:Q2 Adv. & 94–96 Ann. Rev.: 7/31/97
Aug. 97	8/14/97	8/19/97	85:Q1–97:Q2	97:Q2	97:Q3–98:Q4	97:Q2 Pre.: 8/28/97
Sep. 97	9/24/97	9/30/97	85:Q1–97:Q2	97:Q2	97:Q3–99:Q4	97:Q2 Fin.: 9/26/97 97:Q3 Adv.: 10/31/97
Nov. 97	11/6/97	11/12/97	85:Q1–97:Q2	97:Q3	97:Q4–99:Q4	97:Q3 Pre.: 11/26/97
Dec. 97	12/11/97	12/16/97	85:Q1–97:Q2	97:Q3	97:Q4–99:Q4	97:Q3 Fin.: 12/23/97

## A.II. Greenbook and NIPA release dates (March 98–May 1999)

GB name	GB closed	FOMC meeting	Estim. period, DSGE model	Last qtr of history	GB forecast horizon	Interim NIPA releases
Jan. 98	1/28/98	2/3–4/98	85:Q1–97:Q2	97:Q4	98:Q1–99:Q4	97:Q4 Adv.: 1/30/98 97:Q4 Pre.: 2/27/98
Mar. 98	3/19/98	3/25/98	85:Q1–97:Q2	97:Q4	98:Q1–99:Q4	97:Q4 Fin.: 3/26/98 98:Q1 Adv.: 4/30/98
May 98	5/14/98	5/19/98	85:Q1–97:Q2	98:Q1	98:Q2–99:Q4	98:Q1 Pre.: 5/28/98
Jun. 98	6/24/98	6/30/98 & 7/1/98	85:Q1–97:Q2	98:Q1	98:Q2–99:Q4	98:Q1 Fin.: 6/25/98 98:Q2 Adv. & 95–97 Ann. Rev.: 7/31/98

## A.II. (Continued)

GB name	GB closed	FOMC meeting	Estim. period, DSGE model	Last qtr of history	GB forecast horizon	Interim NIPA releases
Aug. 98	8/13/98	8/18/98	85:Q1–98:Q2	98:Q2	98:Q3–99:Q4	98:Q2 Pre.: 8/27/98
Sep. 98	9/23/98	8/29/98	85:Q1–98:Q2	98:Q2	98:Q3–00:Q4	98:Q2 Fin.: 9/24/98 98:Q3 Adv.: 10/30/98
Nov. 98	11/13/98	11/17/98	85:Q1–98:Q2	98:Q3	98:Q4–00:Q4	98:Q3 Pre.: 11/24/98
Dec. 98	12/16/98	12/22/98	85:Q1–98:Q2	96:Q3	98:Q4–00:Q4	98:Q3 Fin.: 12/23/98
Jan. 99	1/28/99	2/2–3/99	85:Q1–98:Q2	98:Q4	99:Q1–00:Q4	98:Q4 Adv.: 1/29/99 98:Q4 Pre.: 2/26/99
Mar. 99	3/24/99	3/30/99	85:Q1–98:Q2	98:Q4	99:Q1–00:Q4	98:Q4 Fin.: 3/31/99 98:Q1 Adv.: 4/30/99
May 99	5/13/99	5/18/99	85:Q1–98:Q2	99:Q1	99:Q2–00:Q4	99:Q1 Pre.: 5/27/99

## A.III. Greenbook and NIPA release dates (June 1999–September 2000)

GB name	GB closed	FOMC meeting	Estim. period, DSGE model	Last qtr of history	GB forecast horizon	Interim NIPA releases
Jun. 99	6/23/99	6/29–30/99	85:Q1–98:Q2	99:Q1	99:Q2–00:Q4	99:Q1 Fin.: 6/25/99 99:Q2 Adv.: 7/29/99
Aug. 99	8/18/99	8/24/99	85:Q1–99:Q2	99:Q2	99:Q3–00:Q4	99:Q2 Pre.: 8/26/99
Sep. 99	9/29/99	10/5/99	85:Q1–99:Q2	99:Q2	99:Q3–01:Q4	99:Q2 Fin.: 9/30/99 99:Q3 Adv. & Comp. Rev.: 10/28/99
Nov. 99	11/10/99	11/16/99	85:Q1–99:Q3	96:Q3	99:Q4–01:Q4	99:Q3 Pre.: 11/24/99
Dec. 99	12/15/99	12/21/99	85:Q1–99:Q3	99:Q3	99:Q4–01:Q4	99:Q3 Fin.: 12/22/99
Jan. 00	1/27/00	2/1–2/00	85:Q1–99:Q3	99:Q4	00:Q1–01:Q4	99:Q4 Adv.: 1/28/00 99:Q4 Pre.: 2/25/00
Mar. 00	3/15/00	3/21/00	85:Q1–99:Q3	99:Q4	00:Q1–01:Q4	99:Q4 Fin.: 3/30/00 99:Q1 Adv.: 4/27/00
May 00	5/11/00	5/16/00	85:Q1–99:Q3	00:Q1	00:Q2–01:Q4	00:Q1 Pre.: 5/25/00
Jun. 00	6/21/00	6/27–28/00	85:Q1–99:Q3	00:Q1	00:Q2–01:Q4	00:Q1 Fin.: 6/27/00 00:Q2 Adv. & 97–99 Ann. Rev.: 7/28/00



A.III. (*Continued*)

GB name	GB closed	FOMC meeting	Estim. period, DSGE model	Last qtr of history	GB forecast horizon	Interim NIPA releases
Aug. 00	8/16/00	8/22/00	85:Q1–00:Q2	00:Q2	00:Q3–01:Q4	00:Q2 Pre.: 8/25/00 00:Q2 Fin.: 9/28/00 00:Q3 Adv.: 10/27/00
Sep. 00	9/27/00	9/3/00	85:Q1–00:Q2	00:Q2	00:Q3–02:Q4	

## A.IV. Greenbook and NIPA release dates (Nov. 2000–January 2002)

GB name	GB closed	FOMC meeting	Estim. period, DSGE model	Last qtr of history	GB forecast horizon	Interim NIPA releases
Nov. 00	11/8/00	11/15/00	85:Q1–00:Q2	00:Q3	00:Q4–02:Q4	00:Q3 Pre.: 11/29/00 00:Q3 Fin.: 12/21/00
Dec. 00	12/13/00	12/19/00	85:Q1–00:Q2	00:Q3	00:Q4–02:Q4	
Jan. 01	1/25/01	1/30–31/01	85:Q1–00:Q2	00:Q4	01:Q1–02:Q4	00:Q4 Adv.: 1/31/01 00:Q1 Pre.: 2/28/01
Mar. 01	3/14/01	3/20/01	85:Q1–00:Q2	00:Q4	01:Q1–02:Q4	
May 01	5/9/01	5/15/01	85:Q1–00:Q2	01:Q1	01:Q2–02:Q4	00:Q4 Fin.: 3/20/01 01:Q1 Adv.: 4/27/01
Jun. 01	6/20/01	6/26–27/01	85:Q1–00:Q2	01:Q1	01:Q2–02:Q4	
Aug. 01	8/15/01	8/21/01	85:Q1–01:Q2	01:Q2	01:Q3–02:Q4	01:Q1 Fin.: 6/28/01 01:Q2 Adv. & 98–00 Ann. Rev.: 7/27/01
						01:Q2 Pre.: 8/17/01 01:Q2 Fin.: 9/24/01
						01:Q3 Adv.: 10/31/01
Nov. 01	10/31/01	11/6/01	85:Q1–01:Q2	01:Q3	01:Q4–03:Q4	01:Q3 Pre.: 11/20/01
Dec. 01	12/5/01	12/11/01	85:Q1–01:Q2	01:Q3	01:Q4–03:Q4	01:Q3 Fin.: 12/19/01
Jan. 02	1/23/02	1/29–30/02	85:Q1–01:Q2	01:Q4	02:Q1–03:Q4	01:Q4 Adv.: 1/30/02 01:Q4 Pre.: 2/28/02

## A.V. Greenbook and NIPA release dates (March 2002–June 2003)

GB name	GB closed	FOMC meeting	Estim. period, DSGE model	Last qtr of history	GB forecast horizon	Interim NIPA releases
Mar. 02	3/13/02	3/19/02	85:Q1–01:Q2	01:Q4	02:Q1–03:Q4	01:Q4 Fin.: 3/19/02 02:Q1 Adv.: 4/26/02

## A.V. (Continued)

GB name	GB closed	FOMC meeting	Estim. period, DSGE model	Last qtr of history	GB forecast horizon	Interim NIPA releases
May 02	5/1/02	5/7/02	85:Q1–01:Q2	02:Q1	02:Q2–03:Q4	02:Q1 Pre.: 5/17/02
Jun. 02	6/19/02	6/25–26/02	85:Q1–01:Q2	02:Q1	02:Q2–03:Q4	02:Q1 Fin.: 6/20/02 02:Q2 Adv. & 99–01 Ann. Rev.: 7/31/02
Aug. 02	8/7/02	8/13/02	85:Q1–02:Q2	02:Q2	02:Q3–03:Q4	02:Q2 Pre.: 8/20/02
Sep. 02	9/18/02	9/24/02	85:Q1–02:Q2	02:Q2	02:Q3–04:Q4	02:Q2 Fin.: 9/23/02 02:Q3 Adv.: 10/31/02
Nov. 02	10/30/02	11/6/02	85:Q1–02:Q2	02:Q3	02:Q4–04:Q4	02:Q3 Pre.: 11/19/02
Dec. 02	12/4/02	12/10/02	85:Q1–02:Q2	02:Q3	02:Q4–04:Q4	02:Q3 Fin.: 12/18/02
Jan. 03	1/22/03	1/28–29/03	85:Q1–02:Q2	02:Q4	03:Q1–04:Q4	02:Q4 Adv.: 1/30/03 02:Q4 Pre.: 2/28/03
Mar. 03	3/12/03	3/18/03	85:Q1–02:Q2	02:Q4	03:Q1–04:Q4	02:Q4 Fin.: 3/14/03 03:Q1 Adv.: 4/25/03
May 03	4/30/03	5/6/03	85:Q1–02:Q2	03:Q1	03:Q2–04:Q4	03:Q1 Pre.: 5/22/03
Jun. 03	6/18/03	6/24–25/03	85:Q1–02:Q2	03:Q1	03:Q2–04:Q4	03:Q1 Fin.: 6/19/03 03:Q2 Adv.: 7/31/03

## A.VI. Greenbook and NIPA release dates (August 2003–November 2004)

GB name	GB closed	FOMC meeting	Estim. period, DSGE model	Last qtr of history	GB forecast horizon	Interim NIPA releases
Aug. 03	8/6/03	8/12/03	85:Q1–03:Q2	03:Q2	03:Q3–04:Q4	03:Q2 Pre.: 8/14/03
Sep. 03	9/10/03	9/16/03	85:Q1–03:Q2	03:Q2	03:Q3–05:Q4	03:Q2 Fin.: 9/15/03
Oct. 03	10/22/03	10/28/03	85:Q1–03:Q2	03:Q3	03:Q4–05:Q4	03:Q3 Adv.: 10/30/03 03:Q3 Pre.: 11/13/03
Dec. 03	12/3/03	12/9/03	85:Q1–03:Q2	03:Q3	03:Q4–05:Q4	Comp. Rev.: 12/10/03 03:Q3 Fin.: 12/16/03
Jan. 04	1/21/04	1/27–28/04	85:Q1–03:Q2	03:Q4	04:Q1–05:Q4	03:Q4 Adv.: 1/30/04 03:Q4 Pre.: 2/27/04
Mar. 04	3/10/04	3/16/04	85:Q1–03:Q2	03:Q4	04:Q1–05:Q4	03:Q4 Fin.: 3/25/03 04:Q1 Adv.: 4/29/04

A.VI. (*Continued*)

GB name	GB closed	FOMC meeting	Estim. period, DSGE model	Last qtr of history	GB forecast horizon	Interim NIPA releases
May 04	4/28/04	5/4/04	85:Q1–03:Q2	04:Q1	04:Q2–05:Q4	04:Q1 Pre.: 5/27/04
Jun. 04	6/23/04	6/29–30/04	85:Q1–03:Q2	04:Q1	04:Q2–05:Q4	04:Q1 Fin.: 6/25/04 04:Q2 Adv. & 01–03 Ann. Rev.: 7/30/04
Aug. 04	8/4/04	8/10/04	85:Q1–04:Q2	04:Q2	04:Q3–05:Q4	04:Q2 Pre.: 8/27/04
Sep. 04	9/15/04	9/21/04	85:Q1–04:Q2	04:Q2	04:Q3–06:Q4	04:Q2 Fin.: 9/29/04 04:Q3 Adv.: 10/29/04
Nov. 04	11/3/04	11/10/04	85:Q1–04:Q2	04:Q3	04:Q4–06:Q4	