EFSF
7/11/08
4:30 pm

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*

March 25, 2008

Abstract

This paper considers the "real-time" forecast performance of the Federal Reserve staff, times-series models, and an estimated dynamic stochastic general equilibrium (DSGE) model – the Federal Reserve Board's new Estimated, Dynamic, Optimization-based (Edo) model. We evaluate forecast performance using out-of-sample predictions from 1996 through 2005 – examining over 70 forecasts presented to the Federal Open Market Committee (FOMC). Our analysis builds on previous real-time forecasting exercises along two dimensions. First, we consider time-series models, a structural DSGE model that has been employed to answer policy questions quite different from forecasting, and the forecasts produced by the staff at the Federal Reserve Board. In addition, we examine forecasting performance of our DSGE model at a relatively detailed level by separately considering the forecasts for various components of consumer expenditures and private investment. The results provide significant support to the notion that richly specified DSGE models belong in the forecasting toolbox of a central bank. We also identify important areas for further research.

*Rochelle M. Edge (rochelle.m.edge@frb.gov), Michael T. Kiley (michael.t.kiley@frb.gov), and Jean-Philippe Laforte (jean-philippe.laforte@frb.gov) are affiliated with the Board of Governors of the Federal Reserve System. We would like to thank seminar participants at the Dynare 2006 conference, the European Central Bank, and the 2007 meetings of the Society for Computational Economics for helpful comments. The views expressed do not necessarily reflect those of the Board of Governors of the Federal Reserve System or it staff. All errors are our own.

1 Introduction

This paper considers the forecast performance since 1996 of the Federal Reserve staff, reduced-form time-series models, and an estimated dynamic stochastic general equilibrium (DSGE) model. We have 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[1] In addition, much of the research regarding DSGE models has emphasized within-sample measures of fit (such as Bayesian posterior odds or marginal likelihoods) – frequently finding strong support for DSGE specifications. As emphasized by Sims [2003a], these measures can be dependent on the analyst's prior views and often appear "too" decisive; in response to this concern, we focus on out-of-sample forecast performance.[2] Finally, we examine forecast performance for both top-line macroeconomic variables – the state of the labor market, growth of Gross Domestic Product, inflation, and the federal funds rate – and for detailed subcategories of aggregate expenditure – e.g., consumption of nondurables and services and investment in consumer durables, residential capital, and business 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 (e.g., 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 – the Edo model. This model contains a rich description of production, expenditure, labor supply, and pricing decisions for the economy of the United States. We have presented detailed descriptions of the model's structure, our estimation strategy, and results in previous papers (Edge *et al.* [2007b] and Edge *et al.* [2007a]).

[1] See, Romer and Romer [2000] and Sims [2003b]; Tulip [2005] finds some deterioration in the relative forecast performance of Federal Reserve staff forecasts. Faust and Wright [2007] provides related evidence.

[2] Other research has looked at similar issues in a more limited way, e.g. our inclusion of a DSGE model within the set of forecast models follows recent work (in particular by Smets and Wouters [2007]) suggesting that advances in Bayesian estimation methods have made DSGE models capable of providing informative forecasts. Adolfson *et al.* [2005] and Christoffel *et al.* [2007] have explored related issues for models of the Euro area, as discussed below.

We present a 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.* [2007b]. 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.* [2007b]. We have also investigated the fluctuations in the U.S. housing market – which have been considerable over the past decade – using the Edo model (Edge *et al.* [2007c]). Significantly, we use the same model in this other research and 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. Meyer [1997] emphasizes the multiple roles, including forecasting and emphasizing "storytelling", of macroeconomic models in policymaking and private-sector consulting work.

Our period of analysis spans macroeconomic developments in the United States from mid-1996 to mid-2005 (where the end-point is determined by the availability of data for forecast evaluation). 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 2005, the U.S. economy experienced substantial swings in activity. From 1996 to early 2000, economic growth was rapid and inflation was low – a "Goldilocks" economy as dubbed by Gordon [1998]. A recession followed in 2001. While the recession was brief, the labor market was slow to recover (e.g., Langdon *et al.* [2004]), and our analysis through this period allows an examination of the success of our model and other techniques at capturing this period. 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 rate of change in the consumer price index excluding food and energy prices dropped to around 1 percent that year. Overall price inflation stepped up after 2003.

Our analysis yields very strong support for the notion that a richly-specified DSGE model can provide useful information to support the forecasting efforts at policy institutions. We find that the forecast accuracy of the Edo model is as good as, and in many cases much better than, that of the forecasts of the Federal Reserve staff and the FRB/US model or projections from time-series methods. This finding is fairly uniform across top-line measures of economic activity and inflation and for detailed components of aggregate expenditure.

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 U.S. 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 U.S. Bureau of Economic Analysis – the statistics that influence policy deliberations and public discussion of economic fluctuations; 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.

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 Negro and Schorfheide [2002]. 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, as we discuss in section 5 through 7. Finally, Lees *et al.* [2007] examine a very small set of variables – 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 and large central banks (as we discuss in Edge *et al.* [2007b]).

Adolfson *et al.* [2005] 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; 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.* [2005] 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.

Section 2 provides an overview of the Edo model. Section 3 presents how we construct forecasts and a comparison of the in-sample fit of our DSGE model to that of a vector-autoregression; we highlight some difficulties with such a comparison that provide some of the motivation for our real-time forecast evaluation. Section 4 introduces the alternative forecasts against which we compare the Edo model. We focus on the Federal Reserve Board's staff projections, including those from the FRB/US model, and the forecasts from vector autoregessions. For our purposes, an accurate comparison of the performance of different forecasts necessitates the use of real-time data, which is also discussed in the fourth section. Section 5 presents the comparison between Edo and time-series models results, and section 6 expands the analysis to include Federal Reserve forecasts while examining subsample results to illustrate important economic successes and failures of our model. Section 7 concludes and points to directions for future research.

2 A Two-Sector DSGE Model for Forecasting

Research on policy applications of dynamic. stochastic, general-equilibrium (DSGE) models has exploded in the last five years. On the policy front, the GEM project at the International Monetary Fund (e.g., IMF [2004]) and the SIGMA project at the Federal Reserve (e.g., 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.[3] The Estimated, Dynamic, Optimization-based (Edo) model project at the Federal Reserve has been designed to build on earlier work a 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.* [2007b], Edge *et al.* [2007a], and Edge *et al.* [2007c].

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, respec-

[3] Reifschneider *et al.* [1997] discuss the use of the FRB/US model at the Federal Reserve.

tively) 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 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^{cbi}) is used for consumer nondurable goods and services and residential capital goods and the output of a faster growing sector (denoted X_t^{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.* [2007b] 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 U.S. 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 (e.g., 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, $K_t^{u,nr,s}$ where the superscript "u" denotes utilized capital and the superscript "nr" indicates nonresidential capital, and economywide and sector-specific productivity, Z_t^m, and Z_t^s.[4]

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

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^{kb} evolve according to

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

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

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

where $\epsilon_t^{z,n}$ 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, e.g. $\Gamma_*^{z,kb} > 1$.

Each intermediate-good producers' output enters a final-goods production technology for its sector that takes the Dixit-Stiglitz form. 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

[4] We normalize Z_t^{cbi} to one, while Z_t^{kb} is not restricted.

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, $\Lambda_t^{cnn}/P_t^{cbi}$, to solve:

$$\max_{\{P_t^s(j),X_t^s(j),X_t^s(j)\}_{t=0}^{\infty}} \mathcal{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{\Lambda_t^{cnn}}{P_t^{cbi}} \Bigg\{ P_t^s(j) X_t^s(j) - MC_t^s(j) X_t^s(j) - \frac{100 \cdot \chi^p}{2} \left(\frac{P_t^s(j)}{P_{t-1}^s(j)} - \eta^p \Pi_{t-1}^{p,s} - (1-\eta^p) \Pi_*^{p,s} \right)^2 P_t^s X_t^s \Bigg\}$$

$$\text{subject to } X_\tau^s(j) = (P_\tau^s(j)/P_\tau^s)^{-\Theta_\tau^{x,s}} X_\tau^s \text{ for } \tau = 0, 1, ..., \infty \text{ and } s = cbi, kb, \tag{4}$$

The profit function reflects price-setting adjustment costs (the size which depend on the parameter χ^p and the lagged and steady-state inflation rate). 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 $\Theta_t^{x,s}$, e.g., 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

$$\theta_t^{x,s} = \epsilon_t^{\theta,x,s} \tag{5}$$

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

Our lengthy treatment of the structure of our model (Edge *et al.* [2007a]) provides further details on 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^{nr}, the stock of non-residential capital, K_t^{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, $K_t^{nr,cbi}$, U_t^{cbi}, $K_t^{nr,kb}$, and U_t^{kb}. (Recall, that production in equation 1 depends on utilized capital $K_t^{u,nr,s} = U_t^s K_t^{nr,s}$.)[5]

The mathematical representation of this decision is described by the following maximization problem (in which capital owners take as given the rental rate on non-residential capital, R_t^{nr}, and the price of non-residential capital goods, P_t^{kb}, and households' valuation of nominal capital income in each period, $\Lambda_t^{cnn}/P_t^{cbi}$):

$$\max_{\{E_t^{nr}(k),K_{t+1}^{nr}(k),K_t^{nr,cbi}(k),K_t^{nr,kb}(k)U_t^{cbi}(k),U_t^{kb}(k)\}_{t=0}^{\infty}}$$

$$\mathcal{E}_0 \sum_{t=0}^{\infty} \beta^t \frac{\Lambda_t^{cnn}}{P_t^{cbi}} \Bigg\{ R_t^{nr} U_t^{cbi}(k) K_t^{nr,cbi}(k) + R_t^{nr} U_t^{kb}(k) K_t^{nr,kb}(k) - P_t^{kb} E_t^{nr}(k)$$
$$-\kappa \left(\frac{U_t^{cbi}(k)^{1+\psi} - 1}{1+\psi} \right) P_t^{kb} K_t^{nr,cbi} - \kappa \left(\frac{U_t^{kb}(k)^{1+\psi} - 1}{1+\psi} \right) P_t^{kb} K_t^{nr,kb} \Bigg\}$$

subject to

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

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

$$a_t^{nr} = \rho^{a,nr} a_{t-1}^{nr} + \epsilon_t^{a,nr}. \tag{7}$$

The problems solved by the consumer durables and residential capital owners are slightly simpler than the nonresidential capital owner's problems. Since utilization rates are not

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

variable for these types of capital, their owners make only investment and capital accumulation decisions. Taking as given the rental rate on consumer durables capital, R_t^{cd}, and the price of consumer-durable goods, P_t^{kb}, and households' valuation of nominal capital income, $\Lambda_t^{cnn}/P_t^{cbi}$, the capital owner chooses investment in consumer durables, I_t^{cd}, and its implied capital stock, K_t^{cd}, to solve:

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

subject to

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

for $\tau = 0, 1, ..., \infty$. (8)

The residential capital owner's decision is analogous:

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

subject to

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

for $\tau = 0, 1, ..., \infty$. (9)

The notation for the consumer durables and residential capital stock problems paralells that of non-residential capital. In particular, the capital-efficiency shocks, A_t^{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 – business investment, residential investment, and investment in consumer durables – enter the optimality conditions driving investment decisions as shocks to the "intertemporal IS curves" (to borrow a phrase from the New-Keynesian literature) driving investment.

2.3 The Household's Problem

The final private agent in the model that we will discuss is the household who makes both expenditures and labor-supply decisions. Households derive utility from four sources:

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

The utility that households derive 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 non-durable goods and non-housing services is equal to a factor h^{cnn} multiplied by its consumption last period E_{t-1}^{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^{cnn}, the quantities of residential and consumer durable capital it wishes to rent, K_t^r and K_t^{cd}, its holdings of bonds, B_t, its wage for each sector, W_t^{cbi} and W_t^{kb}, and supply of labor consistent with each wage, L_t^{cbi} and L_t^{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:

$$\max_{\left\{E_t^{cnn}(i),K_t^{cd}(i),K_t^r(i),\{W_t^s(i),L_t^s(i)\}_{s=cbi,kb},B_{t+1}(i)\right\}_{t=0}^{\infty}}$$

$$\mathcal{E}_0\sum_{t=0}^{\infty}\beta^t\Big\{\varsigma^{cnn}\Xi_t^{cnn}\ln(E_t^{cnn}(i)-h^{cnn}E_{t-1}^{cnn}(i))+\varsigma^{cd}\Xi_t^{cd}\ln(K_t^{cd}(i)-h^{cd}K_{t-1}^{cd}(i))$$
$$+\varsigma^r\Xi_t^r\ln(K_t^r(i)-h^rK_{t-1}^r(i))-\varsigma^l\Xi_t^l\frac{(L_t^{cbi}(i)+L_t^{kb}(i))^{1+\nu}}{1+\nu}\Big\}.$$

subject to

$$R_\tau^{-1}B_{\tau+1}(i)=B_\tau(i)+\sum_{s=cbi,kb}W_\tau^s(i)L_\tau^s(i)+Profits_\tau(i)+Other\ \ Transfers_\tau(i)-P_\tau^{cbi}E_\tau^{cnn}(i)$$
$$-R_\tau^{cd}K_\tau^{cd}-R_\tau^rK_\tau^r-\sum_{s=cbi,kb}\frac{100\cdot\chi^w}{2}\left(\frac{W_\tau^s(j)}{W_{\tau-1}^s(j)}-\eta^w\Pi_{\tau-1}^{w,s}-(1-\eta^w)\Pi_*^w\right)^2W_\tau^sL_\tau^s$$
$$-\frac{100\cdot\chi^l}{2}\left(\frac{L_*^{cbi}\cdot W_\tau^{cbi}}{L_*^{cbi}+L_*^{kb}}+\frac{L_*^{kb}\cdot W_\tau^{kb}}{L_*^{cbi}+L_*^{kb}}\right)\left(\frac{L_\tau^{cbi}(i)}{L_\tau^{kb}(i)}-\eta^l\frac{L_{\tau-1}^{cbi}}{L_{\tau-1}^{kb}}-(1-\eta^l)\frac{L_*^{cbi}}{L_*^{kb}}\right)^2\frac{L_\tau^{kb}}{L_\tau^{cbi}}.$$

$$L_\tau^{cbi}(i)=\left(W_\tau^{cbi}(i)/W_\tau^{cbi}\right)^{-\Theta_\tau^{l,cbi}}L_\tau^{cbi},\text{ and }L_\tau^{kb}(i)=\left(W_\tau^{kb}(i)/W_\tau^{kb}\right)^{-\Theta_\tau^{l,kb}}L_\tau^{kb},$$

$$\text{for }\tau=0,1,...,\infty. \tag{10}$$

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

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

$$\xi_t^x = \rho^{\xi,x}\xi_{t-1}^x + \epsilon_t^{\xi,x}, \qquad\qquad x = cnn, cd, r, l. \tag{11}$$

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

The household's budget constraint reflects wage setting adjustment costs, which depend on the parameter χ^w and the lagged and steady-state wage inflation rate; 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 χ^l. The costs incurred by household when the mix of labor input across sectors changes may be inportant for sectoral comovements.

In summary, the households' 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 an Taylor-type interest-rate feedback rule. Policymakers smoothly adjust the actual interest rate R_t to its target level $\bar{R}_t$

$$R_t = (R_{t-1})^{\phi^r}\left(\bar{R}_t\right)^{1-\phi^r}\exp\left[\epsilon_t^r\right], \tag{12}$$

where the parameter ϕ^r reflects the degree of interest rate smoothing, while ϵ_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^{gdp}/H_*^{gdp}, the acceleration of GDP growth, H_t^{gdp}/H_{t-1}^{gdp}, GDP inflation relative to target, $\Pi_*^{p,gdp}/\Pi_t^{p,gdp}$, and the acceleration of GDP inflation, $\Pi_t^{p,gdp}/\Pi_{t-1}^{p,gdp}$:

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

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

2.5 Measuring Aggregate Output

We have focsued on sectoral production decisions so far; Gross Domestic Product (GDP) has not yet been discussed. 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:

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

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

Stationary un-modeled output is exogenous and is assumed to follow the process:

$$\ln \widetilde{X}_t^{gf} - \ln \widetilde{X}_*^{gf} = \rho^{x,gf}\left(\ln \widetilde{X}_t^{gf} - \ln \widetilde{X}_*^{gf}\right) + \epsilon^{x,gf}.$$

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 $\Pi_t^{p,gdp}$, is defined implicitly by:

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

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 Edo reflect the 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 will allow 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 2005, 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.

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 (e.g., Kiley [2007], Laforte [2007], and Rudd and Whelan [2007]).

3 A Traditional Evaluation of the DSGE model

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, as such metrics have dominated recent research (e.g., 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 (twelve) economic time series for the United States:

1. Real gross domestic product;
2. Real consumption expenditure on nondurables and services;
3. Real consumption expenditure on durables;
4. Real residential investment expenditure;
5. Real business investment expenditure, which equals real gross private domestic investment minus real residential investment;[6]
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;[7]
10. Real wage inflation, which equals teh 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 ten-year U.S. 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

[6] Subtraction is performed using the appropriate techniques for aggregates measured as Fisher Ideal indexes.

[7] We scale nonfarm business hours by the ratio of nominal spending in our model to nominal non-farm business sector output in order to model a level of hours more appropriate for the total economy.

Markov-Chain Monte Carlo (MCMC) methods. We also add measurement errors processes, denoted η_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 parmaters are calibrated and held fixed throughout. As reported in table 1, we fix the household's discount factor(β), the Cobb-Douglas share of capital input (α), the curvature parameter associated with costs of varying capital utilization (ψ), the depreciation rates (δ^{nr}, δ^{cd}, δ^{r}), and the elasticities of substition between differentiated intermediate goods and labor input ($\Theta_*^{x,cbi}$, $\Theta_*^{x,kb}$, Θ_*^{l}). Forecast performance is not very sensitive to reasonable (small) variation in these paramters.

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 five years prior to the end of the data, 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 end of the available data. These choices determine the parameters $\Pi_*^{p,gdp}$, $\Gamma_*^{z,m}$, and $\Gamma_*^{z,kb}$.

The remainder of the model parameters are estimated. The priors placed over the model parameters are reported in table 2 and table 3. We highlight the following: the parameters governing habit persistence (h^{cnn}, h^{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 or 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 – i.e., 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.[8]

[8]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 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.

In addition to the DSGE model, we compute the marginal likelihood for a vector autoregression (VAR) in the same 12 variables mentioned above, with one lag. (Given the large number of included variables, additional lags tend to lower the forecast performance of the model, likely from overfitting). The results are extreme: for the period from 1984Q2 to 2004Q2, the marginal (log) likelihood for the VAR model is -1119.1; the comparable figure for the DSGE/Edo model is -753.5. Such a large difference in marginal likelihoods implies that a researcher should place essentially no weight on the VAR model.

While the view from the computed marginal likelihoods are quite favorable for the DSGE model, we think an evaluation of "real-time" forecast accuracy may provide a more balanced view of the suitability of the DSGE model for forecast exercises. For example, the marginal likelihoods are quite decisive in favor of the DSGE/Edo model, and it is not immediately obvious how this result depends upon specifications of the priors over parameters for each model or sample period (e.g., Sims [2003a]). Such decisiveness also seems unlikely to hold given the large literature that suggests that averaging across models provides superior forecasts to reliance on single models (e.g., Clark and McCracken [2006]). In addition, the use of an ex ante criterion may provide a more reliable safeguard against overfitting and imposition of priors that actually reflect sample information.

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 after the color of the cover in which it is wrapped, 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 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 are most readily available on the web-site of the Federal Reserve Bank of Philadelphia, reflect the views of the staff and not the Committee members.

The maximum projection horizon for the Greenbook forecast vintages that we consider in this paper vary from six to ten quarters. 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. We use the forecasts produced for the FOMC meetings starting in September 1996 and ending in March 2001; this period includes the beginning of the period when the FRB/US model (discussed below) was employed. We choose March 2001 as the end point both because that month marked the beginning of a recession in the United States and, perhaps more importantly for our purposes, because Greenbook forecasts are only made public with a five-year lag, so forecasts through 2001 are the most recent vintage that is publicly available. An 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 June 2005.

4.2 The FRB/US Model Forecast

The Greenbook projection is a judgmental projection that reflects input from a large number of economists; it is not the output of any individual model. The second forecast that we compare with our DSGE model projection are those produced by the Federal Reserve’s FRB/US model, one of the tools used as input to the deliberations that lead to the Greenbook projection. These model forecasts are prepared at the same time as each Greenbook forecast is prepared. The FRB/US model forecast uses the projection for the federal funds rate used in the Greenbook projection, so all statistics related to the federal funds rate in our comparisions 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, its equations for most economic decisions are related to those based on explicit optimization like in Edo, but ad hoc elements are introduced to improve model fit in many cases. In addition, the specification of FRB/US has proceeded along an "equation-by-equation" route that some have criticized for a lack of attention to system properties and econometric rigor (e.g., Sims [2003b] and Sims [2006]). 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 twelve variables used in estimation of Edo. We choose a one-lag VAR both because of the large number of variables in the system and because this is the baseline used for comparions in Smets and Wouters [2007]. The second model is a one-lag Bayesian VAR that introduce onto the coeffients a modified version of the Litterman [1980] prior. [9]. We re-estimate these models for each forecast.

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 untouched since they were archived on the dates shown in the appendix describing the Greenbook.

[9] The values of the hyperparameters are the following: the overall tightness is 0.2; the cross-equation tightness is 0.5; and, finally, the harmonic lag decay is set to 0.5. The prior distributions of the "constant" parameters are normal densities centered at zero with a standard deviation of one. Since the Litterman prior was originally proposed for variables in levels while the Bayesian VAR model's variables are specified either in first-differences or stationary are in levels, we replace the unit root prior on the dynamic coefficient of a variable on its own first lag with a lower value of 0.7. The posterior distribution is sampled using Gibbs-Sampling methods as explained in Kadiyala and Karlsson [1997]

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 the past Greenbook databases we are assuming the same information set with which the Greenbook forecast was actually made. 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.

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 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 the NIPA takes place. This contrasts with the approach followed for the VAR forecasts, where re-estimation is performed for each forecast. Parameter estimates for Edo 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. Note that it is only the data used to estimate the model that remains 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 for these series contained in the most recent vintage of data (the October 2007 FOMC meeting).

5 Comparison with Time-Series Methods

We focus on two distinct sets of variables. The first are the “top-line” macroeconomic aggregate – hours per capital, the percent change in real GDP, GDP price inflation, 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 business 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 mean absolute error (MAE); results are little different for the root mean-squared error. Bias is measured as the fraction of forecasts at various horizons with positive errors.

5.1 The Main Macroeconomic Aggregates

The main macroeconomic aggregates examined are hours per capita, real GDP growth, GDP price inflation, 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 to 2005 period.

5.1.1 Hours per capita

As noted earlier, the state of the labor market in Edo is summarized by hours worked per capita. And the state of the labor market is one side of the Federal Reserve's dual mandate of full employment and price stability, so the ability of the Edo model to forecast hours per capita, relative to the ability of other models, is an important metric for model evaluation.

We first focus on the results presented in table 4. The statistics related to hours per capita are reported in the upper portion of the table. The line labeled VAR(1) reports the MAE at various forecast horizons (e.g., one through four quarters out and eight quarters out), in percentage points, for the forecast of hours per capita generated by the VAR model. These errors are large, between 2.5 and 3.5 percentage points at the reported horizons. The large size of these errors reflects the real-time nature of the exercise: the 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 MAE for the

BVAR and Edo models relative to the MAE for the VAR model; values below 1 indicate that the model performs better than VAR model. Several results are apparent. First, both the methods perform better than the VAR at all reported horizons. Second, the Edo model performs better than the BVAR model at all horizons.

Figure 2 presents a measure of the bias of the DSGE and BVAR forecasts for hours per capita. The bias measure is computed as the fraction of observations for which the forecast error is positive; an unbiased forecast should have a value of this bias indicator near 0.5, whereas a forecast that systematically forecast a valus that is too high should have a value substantially exceeding 0.5. The upper right panel presents the results for hours per capita. It is very clear that the forecasts for hours per capita were systematically too high. This suggests that, as we mentioned earlier, the period from 1996 to 2005 was one of unexpectedly strong growth in labor productivity. 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.

5.1.2 Real GDP Growth

The mid-1990s to early 2001 were a period of rapid growth in real GDP in the United States. Moreover, the pace of growth over this period was widely unanticipated; for example, Tulip [2005] finds that the staff at the Federal Reserve tended to systematically underpredict the pace of real GDP growth over this period.

The second set of results presented in table 4 focus on the forecasts for real GDP growth. The results for the VAR indicate that the MAE at the horizons reported was in the neighborhood of 0.5 percent. Turning to the comparison to the other models, the BVAR model performs better according to the MAE criterion at horizons less than one year. The Edo DSGE model performs better than the VAR models; the difference is economically meaningful, on the order of 20 percent.

Given the large bias reported earlier for hours per capita, the results on bias for GDP growth are quite interesting. As shown in the upper right panel of figure 2, the bias measures for the Edo and BVAR forecasts are in the neighborhood of 0.5. The relatively modest biases in forecasts for GDP growth suggest that the strong biases in forecasts for hours per capita reflect, in part, unexpectedly strong labor productivity for all the forecast methods examined; in other words, each method experienced relatively little bias for GDP growth but a noticeble positive bias for the level of hours per capita, indicating higher projections of labor input and lower projections for labor productivity over the period from 1996 to 2005 than was realized ex post.

5.1.3 GDP price inflation

The next set of results reported in table 4 focus on the forecasts for GDP prices. The line for the VAR indicates that the MAE for GDP price inflation is about 0.25 percentage point at each horizon. The comparison across forecast methods is again informative. Edo and the BVAR perform better than the VAR at all horizons. Once again the Edo DSGE model performs better than the VAR methods by an economically meaningful margin, especially out one year or more.

The bias statistic from Edo for GDP price inflation are near 0.5, suggesting little systematic bias.

5.1.4 The Federal Funds Rate

The final set of results reported in table 4 focus on the forecasts for the federal funds rate. The line for the VAR indicates that the MAE for the federal funds rate is about 0.1 percentage point at the one-quarter horizon and rises to 0.4 percentage point at the eight-quarter horizon. The comparison across between the BVAR and Edo is again informative. The VAR methods dominate Edo at short horizons; but Edo dominates at longer horizons, perhaps suggesting that the greater forecast accuracy for inflation and GDP growth at long horizons for Edo, combined with the policy rule linking the federal funds rate to these variables, helps forecast accuracy out a year or more.

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 (e.g., 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 business investment, and real residential investment in table 5. The structure of the reported statistics is the same as in table 4. We take away two summary points. First, the forecast of the Edo DSGE model, as summarized by the MAE, is more accurate than that of the VAR and BVAR for the components of consumption and business investment – sometimes by large margins. Second, Edo has more difficulty forecasting residential investment than the VAR 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, or superior to, those from time-series models or the staff of the Federal Reserve Board for a broad range of variables. The relative success of our DSGE model at forecasting provides support to the use of such models in a policy context. These findings are similar to those of Smets and Wouters [2007]. As emphasized in Edge *et al.* [2007b], we have included a further disaggregation of macroeconomic activity relative to Smets and Wouters [2007] for policy-relevant questions independent of forecasting, so the continued success in forecast performance is a positive finding for our detailed DSGE structure. The next section pushes harder on the policy-relevance question by examining Federal Reserve staff forecasts.

6 Comparison to Federal Reserve Staff Forecasts

We now examine the forecast performance of the Edo DSGE model relative to 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 VAR forecasts, in part because Federal Reserve forecasts have not placed much weight on VAR projections. Second, we attempt to identify what features of our model or the data contribute to the successes and failures recorded by the Edo model along the forecast dimension from 1996 to 2005, 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 and the onset of the recession in

March 2001 have led us to focus on comparisons of forecasts using data for FOMC meetings from September 1996 to March 2001.

6.1 Forecast Performance Prior to the 2001 Recession

Table 6 and 7 present statistics on forecast accuracy for the projections generated using the data from the September 1996 FOMC meeting to the March 2001 FOMC meeting – the period before the start of the recession in 2001.

With regard to the labor market, it is apparent that the staff projections in the Greenbook and from the FRB/US model for hours per capita are worse than the VAR forecasts, according to the MAE criteria, at all horizons except for the FRB/US forecast eight quarters out. The forecast performance of the Edo model dominates that of the Greenbook and FRB/US model at all horizons; this may be especially surprising at short horizons, where the Federal Reserve staff devote significant resources to assessing near-term developments (e.g., Romer and Romer [2000], Sims [2003b]). We think this is a significant finding: As we have emphasized in previous work (Edge *et al.* [2007b]), 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.

The results are also very similar for most other measures of economic activity, where again the forecast accuracy in a MAE sense of Edo is better than a VAR or the Greenbook projections. For example, table 6 reports that the MAE for GDP growth from Edo is 30 percent or more smaller than that of the Greenbook or FRB/US model at most horizons. Moreover, table 7 shows that the MAE of the Edo forecast for all four components of private expenditure is lower than that of the Greenbook at every horizon except for growth of residential investment eight quarters out.

Returning to table 6, the results for GDP price inflation continue to suggest that the Edo model is competitive with best practices. In particular, the Edo model dominates, in an MAE sense, the Greenbook and FRB/US at some horizons but is dominated by one of these Federal Reserve methods at other horizons. Sims [2003b] reported that the near-term inflation forecasts in the Greenbook were very good, so the competitive performance 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 Edo provides quite accurate forecasts for the federal funds rate in the period prior to March 2001.

6.2 Recent Performance and Implications

The availability of Greenbook projections to the public has limited our analysis to the period prior to March 2001. We report the forecast statistics for the period from May 2001 to June 2005 for the VAR, BVAR, and Edo in table 8 and table 9; these statistics help complete the comparison between Edo and the time-series methods, as the earlier tables reported results from September 1996 to June 2005 and for September 1996 to March 2001. Overall, the results are very similar: Edo is competitive or superior to other methods along several dimensions. We highlight two points. First, the MAE for the Edo projections of the federal funds rate are poor from May 2001 to June 2005. Second, the MAE for the Edo projections of the growth of residential investment rate are very poor from May 2001 to June 2005.

We interpret the entire set of results in two ways. First, the performance of Edo 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 to 2005 period. As a result, we view efforts to model the labor market in a more nuanced way, including allowance for an intensive and extensive margin, as a high priority. The most notable other aspect of the results for economic activity is the difference between the forecast performance for residential investment before and after 2001. Edo was quite accurate for residential investment prior to the period of fast growth early this decade. This suggests that factors that are not accounted for in Edo may have played a role in recent experience. Financial innovations, such as greater availability of mortgage finance, may have been one factor; behavioral factors, such as speculative investment, may have been another. We view structural investigations of these issues in general equilibrium models as an interesting topic for research.

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 very 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 complex DSGE model is competitive with reasonable forecast alternatives provides support for the use of such models in forecasting and other policy-relevant work. 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 very good forecasts. Some research has been very critical of New-Keynesian models of the Phillips curve (e.g., 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.

Our discussion also highlighted two areas where the results suggest 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 or other factors in the rise, and subsequent fall, of residential investment following 2001.

Finally, we would like to end with a caveat to our "real-time" evaluation. As we highlighted in our discussion of our approach in section 4, we took great care to base our forecasts using Edo and vector autoregressions on data and information available in "real-time" to place these forecasts on equal footing with the Greenbook and FRB/US model forecasts. However, the forecasts from the vector autoregressions and Edo are not truly real time. We have had several luxuries, especially the time to check that all of our codes are correct and that our data is correct. Perhaps as or more importantly, we have benefited, at least indirectly, from our previous research and that of others on what types of models are likely to explain the data. It is impossible to purge our analysis of these influences. As a result, we are cautious in our final verdict. Specifically, we view our analysis as clearly indicating that DSGE models like Edo are valuable forecasting tools and are likely to prove competitive with best practices at institutions like the Federal Reserve. We have some confidence in this view because our findsing are fairly systematic and do not result from excessive search (as, for example, we employ the same model previously employed in Edge *et al.* [2007b], Edge *et al.* [2007a], and Edge *et al.* [2007c]. However, we think it is reasonable to expect that the relative forecast performance of models like Edo in true real-time will be less successful than reported herein. We have been generating and archiving such true real-time forecasts since the May 2007 FOMC meeting.

References

Adolfson, M., Lind, J., and Villani, M. (2005). Forecasting performance of an open economy dynamic stochastic general equilibrium model. available at http://ideas.repec.org/p/hhs/rbnkwp/0190.html.

Altig, D., Christiano, L., Eichenbaum, M., and Linde, J. (2004). An estimated model of us business cycles. Manuscript.

Christoffel, K., Coenen, G., and Warne, A. (2007). The new area-wide model of the euro area: Specification, estimation results and properties. Manuscript, European Central Bank.

Clark, T. and McCracken, M. (2006). Averaging forecasts from vars with uncertain instabilities. *Journal of Applied Econometrics, forthcoming*.

Edge, R., Kiley, M., and Laforte, J. (2007a). Documentation of the research and statistics division's estimated dsge model of the u.s. economy. Federal Reserve Board FEDS Working paper 2007-53.

Edge, R., Kiley, M., and Laforte, J. (2007b). An estimated dsge model of the us economy with an application to natural rate measures. *Journal of Economic Dynamics and Control, forthcoming*.

Edge, R., Kiley, M., and Laforte, J. (2007c). The sources of fluctuations in residential investment. Manuscript.

Erceg, C., Guerrieri, L., and Gust, C. (2006). Sigma: A new open economy model for policy analysis. *International Journal of Central Banking*, **2**.

Faust, J. and Wright, J. (2007). Comparing greenbook and reduced form forecasts using a large realtime dataset. available at http://e105.org/faustj/download/gbts.pdf?d=n.

Gali, J. and Gertler, M. (2007). Macroeconomic modeling for monetary policy evaluation. *Journal of Economic Perspectives*, **21**(4), 25–45.

Gordon, R. (1998). Foundations of the goldilocks economy: Supply shocks and the time-varying nairu. *Brookings Papers on Economic Activity*, (2), p297 – 333.

IMF (2004). Gem; a new international macroeconomic model. IMF working paper, January.

Kadiyala, R. and Karlsson, S. (1997). Numerical methods for estimation and inferecne 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**.

Kohn, D. (2003). The strength in consumer durables and housing: Policy stabilization or 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, California, February 28.

Kydland, F. E. and Prescott, E. C. (1982). Time to build and aggregate fluctuations. *Econometrica*, **50**(6), p1345 – 1370.

Laforte, J. (2007). Pricing models: A bayesian dsge approach to the u.s. economy. *Journal of Money, Credit, and Banking*, **39**.

Langdon, D., Krantz, R., and Strople, M. (2004). Post-recession trends in nonfarm employment and related economic indicators. *Monthly Labor Review*, pages 49–56.

Lees, K., Matheson, T., and Smith, C. (2007). Open economy dsge-var forecasting and policy analysis - head to head with the rbnz published forecasts. available at http://ideas.repec.org/p/nzb/nzbdps/2007-01.html.

Litterman, R. (1980). A bayesian procedure for forecasting with vector autoregressions. Federal Reserve Bank of Minneapolis, manuscript.

Meyer, L. (1997). The role for structural macroeconomic models. Remarks at the American Economic Association panel on monetary and fiscal policy, New Orleans.

Negro, M. D. and Schorfheide, F. (2002). Priors from general equilibrium models for vars.

Reifschneider, D., Stockton, D., and Wilcox, D. (1997). Econometric models and the monetary policy process. *Carnegie-Rochester Conference Series on Public Policy*, **47**, 1–37.

Romer, C. and Romer, D. (2000). Federal reserve information and the behavior of interest rates. *American Economic Review*, **90**, 429–457.

Rudd, J. and Whelan, K. (2007). Modeling inflation dynamics: A critical survey of recent research. *Journal of Money, Credit, and Banking*, **39**.

Sims, C. (2003a). Probability models for monetary policy decisions. Manuscript.

Sims, C. (2003b). The role of models and probabilities in the monetary policy process. *Brookings Paper on Economic Activity*, **2002:2**, 1–63.

Sims, C. (2006). Improving monetary policy models. *Journal of Economic Dynamics and Control, forthcoming*.

Smets, F. and Wouters, R. (2007). Shocks and frictions in the us business cycles: A bayesian dsge approach. *American Economic Review*.

Tulip, P. (2005). Has output become more predictable? changes in greenbook forecast accuracy. *FEDS Working Paper No. 2005-31*.

β	α	ψ	δ^{nr}	δ^{cd}	δ^{r}	$\Theta_*^{x,cbi}$	$\Theta_*^{x,kb}$	Θ_*^{l}
0.990	0.260	5	0.030	0.055	0.004	7.000	7.000	7.000

Table 1: Calibrated Parameters

Param.	Prior Type	Prior Mean	Prior S.D.
h^{cnn}	Beta	0.500	0.015
h^{cd}	Beta	0.500	0.015
h^{r}	Beta	0.500	0.015
ν	Gamma	2.000	1.000
χ^{p}	Gamma	2.000	1.000
η^{p}	Normal	0.000	0.250
χ^{w}	Gamma	2.000	1.000
η^{w}	Normal	0.000	0.250
χ^{nr}	Gamma	2.000	1.000
χ^{cd}	Gamma	2.000	1.000
χ^{r}	Gamma	6.000	1.000
χ^{l}	Gamma	2.000	1.000
η^{l}	Normal	0.000	0.250
r^{π}	Normal	1.500	0.250
$r^{\triangle\pi}$	Normal	0.000	0.250
$r^{h,gdp}$	Normal	0.500	0.250
$r_{\triangle h,gdp}$	Normal	0.000	0.250
ρ^{r}	Beta	0.750	0.013
$\rho^{a,nr}$	Beta	0.500	0.023
$\rho^{a,cd}$	Beta	0.750	0.013
$\rho^{a,r}$	Beta	0.500	0.013
$\rho^{\xi,cnn}$	Beta	0.750	0.013
$\rho^{\xi,cd}$	Beta	0.750	0.013
$\rho^{\xi,r}$	Beta	0.750	0.013
$\rho^{\xi,l}$	Beta	0.750	0.013
$\rho^{\gamma,m}$	Beta	0.500	0.023
$\rho^{\gamma,kb}$	Beta	0.750	0.013
$\rho^{x,gf}$	Beta	0.750	0.013

Table 2: Prior Distributions for Parameters

Param.	Prior Type	Prior Mean	Prior S.D.
$\sigma_{xi,cnn}$	Inverted Gamma	3.000	2.000
$\sigma_{xi,cd}$	Inverted Gamma	3.000	2.000
$\sigma_{xi,r}$	Inverted Gamma	3.000	2.000
$\sigma_{xi,l}$	Inverted Gamma	3.000	2.000
$\sigma_{a,cd}$	Inverted Gamma	2.000	2.000
$\sigma_{a,r}$	Inverted Gamma	4.000	2.000
$\sigma_{a,kb}$	Inverted Gamma	4.000	2.000
$\sigma_{\gamma,m}$	Inverted Gamma	0.500	2.000
$\sigma_{\gamma,kb}$	Inverted Gamma	0.500	2.000
$\sigma_{\theta,m}$	Inverted Gamma	0.500	2.000
$\sigma_{\theta,kb}$	Inverted Gamma	0.500	2.000
σ_{r}	Inverted Gamma	0.200	2.000
ME_1	Inverted Gamma	0.500	2.000
ME_2	Inverted Gamma	0.500	2.000
ME_3	Inverted Gamma	0.500	2.000
ME_4	Inverted Gamma	0.500	2.000
ME_5	Inverted Gamma	0.500	2.000
ME_6	Inverted Gamma	0.500	2.000
ME_7	Inverted Gamma	0.500	2.000
ME_8	Inverted Gamma	0.500	2.000
ME_9	Inverted Gamma	0.500	2.000
ME_{10}	Inverted Gamma	0.500	2.000
ME_{11}	Inverted Gamma	0.500	2.000
ME_{12}	Inverted Gamma	0.500	2.000

Table 3: Prior Distributions for Standard Deviations. ME_j refers to the standard deviation of the measurement error associated with observable variable j.

Model	1Q	2Q	3Q	4Q	8Q
	Hours per capita				
VAR(1)	2.692	2.992	3.261	3.520	2.857
Relative MAE					
DSGE/Edo	0.926	0.904	0.861	0.844	0.764
BVAR(1)	0.984	0.955	0.928	0.905	0.874
	Real GDP Growth				
VAR(1)	0.433	0.446	0.472	0.488	0.349
Relative MAE					
DSGE/Edo	0.876	0.845	0.784	0.815	0.839
BVAR(1)	0.954	0.965	0.989	1.007	1.070
	GDP Inflation				
VAR(1)	0.249	0.273	0.278	0.298	0.211
Relative MAE					
DSGE/Edo	0.868	0.819	0.824	0.734	0.674
BVAR(1)	0.964	0.944	0.919	0.919	0.930
	Federal Funds Rate				
VAR(1)	0.113	0.195	0.272	0.364	0.386
Relative MAE					
DSGE/Edo	1.294	1.159	1.051	0.929	0.710
BVAR(1)	1.004	1.017	1.030	1.012	1.064

Table 4: Mean Absolute Errors of Models: Sep. 1996-June 2005

Model	1Q	2Q	3Q	4Q	8Q
	Real Consumption Growth Nondurables				
VAR(1)	0.287	0.348	0.305	0.301	0.183
Relative MAE					
DSGE/Edo	0.819	0.778	0.778	0.852	0.842
BVAR(1)	0.952	0.954	0.962	1.015	1.013
	Real Consumption Growth Durables				
VAR(1)	2.041	1.954	2.142	2.008	1.211
Relative MAE					
DSGE/Edo	0.909	0.892	0.843	0.924	1.004
BVAR(1)	1.015	0.904	0.931	0.974	0.968
	Real Investment Growth, Business				
VAR(1)	3.351	2.885	2.793	2.910	1.754
Relative MAE					
DSGE/Edo	0.824	0.871	0.926	0.875	0.901
BVAR(1)	0.956	0.987	1.015	1.041	1.105
	Real Investment Growth, Residential				
VAR(1)	1.491	1.413	1.781	2.099	1.827
Relative MAE					
DSGE/Edo	1.141	1.435	1.379	1.312	1.306
BVAR(1)	0.925	0.913	0.972	0.898	0.968

Table 5: Mean Absolute Errors of Models, Disaggregated Variables: Sep. 1996-June 2005

Model	1Q	2Q	3Q	4Q	8Q
			Hours per capita		
VAR(1)	2.229	2.321	2.478	2.696	2.496
Relative MAE					
DSGE/Edo	0.898	0.897	0.873	0.838	0.840
BVAR(1)	1.001	1.005	1.005	1.013	1.042
Greenbook	1.017	1.074	1.081	1.086	1.047
FRB/US	1.006	1.037	1.021	1.003	0.893
			Real GDP Growth		
VAR(1)	0.480	0.529	0.608	0.581	0.442
Relative MAE					
DSGE/Edo	0.781	0.788	0.715	0.777	0.770
BVAR(1)	1.044	1.001	1.047	1.040	1.096
Greenbook	1.120	1.111	1.105	1.092	0.957
FRB/US	1.033	1.229	1.113	1.054	0.891
			GDP Inflation		
VAR(1)	0.258	0.239	0.240	0.246	0.145
Relative MAE					
DSGE/Edo	0.795	0.832	0.846	0.847	0.864
BVAR(1)	0.964	0.970	0.969	0.972	0.965
Greenbook	0.798	0.871	0.609	0.602	0.926
FRB/US	0.659	0.900	0.888	0.793	0.884
			Federal Funds Rate		
VAR(1)	0.091	0.179	0.274	0.378	0.404
Relative MAE					
DSGE/Edo	1.043	0.851	0.756	0.693	0.755
BVAR(1)	1.020	1.019	1.020	1.020	0.993
Greenbook	0.747	0.682	0.760	0.816	0.926
FRB/US	0.747	0.682	0.760	0.816	0.926

Table 6: Mean Absolute Errors of Models: Sep. 1996-March 2001

Model	1Q	2Q	3Q	4Q	8Q
	Real Consumption Growth Nondurables and Services				
VAR(1)	0.341	0.433	0.362	0.349	0.267
Relative MAE					
DSGE/Edo	0.878	0.744	0.789	0.762	0.657
BVAR(1)	0.978	0.959	1.007	1.030	1.029
Greenbook	0.932	0.671	0.868	0.941	0.831
FRB/US	0.871	0.609	0.894	0.966	0.891
	Real Consumption Growth Durables				
VAR(1)	2.174	1.970	2.295	2.440	1.498
Relative MAE					
DSGE/Edo	0.787	0.883	0.862	0.844	0.983
BVAR(1)	1.033	0.925	0.965	0.975	1.042
Greenbook	0.967	0.980	0.895	0.906	0.997
FRB/US	1.050	0.974	0.830	0.918	1.035
	Real Investment Growth, Business				
VAR(1)	3.250	3.337	3.491	3.980	2.471
Relative MAE					
DSGE/Edo	0.895	0.845	0.881	0.867	0.842
BVAR(1)	0.971	1.038	1.033	1.028	1.052
Greenbook	1.177	1.070	1.197	1.118	0.961
FRB/US	1.061	1.244	1.137	0.978	1.046
	Real Investment Growth, Residential				
VAR(1)	1.436	1.844	2.411	2.637	1.857
Relative MAE					
DSGE/Edo	0.820	0.789	0.666	0.633	0.845
BVAR(1)	0.787	0.834	0.852	0.855	0.878
Greenbook	1.001	0.945	0.790	0.779	0.579
FRB/US	0.911	0.673	0.522	0.604	0.594

Table 7: Mean Absolute Errors of Models, Disaggregated Variables: Sep. 1996-March 2001

Model	1Q	2Q	3Q	4Q	8Q
			Hours per capita		
VAR(1)	1.515	1.767	1.954	2.097	1.539
Relative MAE					
DSGE/Edo	0.947	0.909	0.853	0.848	0.699
BVAR(1)	0.970	0.921	0.876	0.831	0.729
			Real GDP Growth		
VAR(1)	0.179	0.167	0.151	0.181	0.116
Relative MAE					
DSGE/Edo	1.011	0.940	0.929	0.881	0.979
BVAR(1)	0.826	0.904	0.864	0.951	1.018
			GDP Inflation		
VAR(1)	0.113	0.147	0.151	0.169	0.134
Relative MAE					
DSGE/Edo	0.956	0.807	0.806	0.647	0.565
BVAR(1)	0.965	0.922	0.877	0.878	0.909
			Federal Funds Rate		
VAR(1)	0.065	0.100	0.128	0.165	0.173
Relative MAE					
DSGE/Edo	1.479	1.449	1.384	1.213	0.654
BVAR(1)	0.993	1.015	1.040	1.003	1.151

Table 8: Mean Absolute Errors of Models: May 2001-June 2005

Model	1Q	2Q	3Q	4Q	8Q
	Real Consumption Growth Nondurables and Services				
VAR(1)	0.107	0.119	0.113	0.117	0.041
Relative MAE					
DSGE/Edo	0.719	0.843	0.757	0.994	1.471
BVAR(1)	0.908	0.945	0.887	0.990	0.957
	Real Consumption Growth Durables				
VAR(1)	0.894	0.914	0.931	0.720	0.420
Relative MAE					
DSGE/Edo	1.066	0.902	0.819	1.066	1.044
BVAR(1)	0.993	0.880	0.886	0.972	0.830
	Real Investment Growth, Business				
VAR(1)	1.636	1.124	0.950	0.810	0.450
Relative MAE					
DSGE/Edo	0.749	0.911	1.013	0.896	1.075
BVAR(1)	0.939	0.907	0.979	1.075	1.258
	Real Investment Growth, Residential				
VAR(1)	0.733	0.440	0.509	0.707	0.847
Relative MAE					
DSGE/Edo	1.473	2.861	3.162	2.648	1.840
BVAR(1)	1.068	1.087	1.271	0.983	1.072

Table 9: Mean Absolute Errors of Models, Disaggregated Variables: May 2001-June 2005

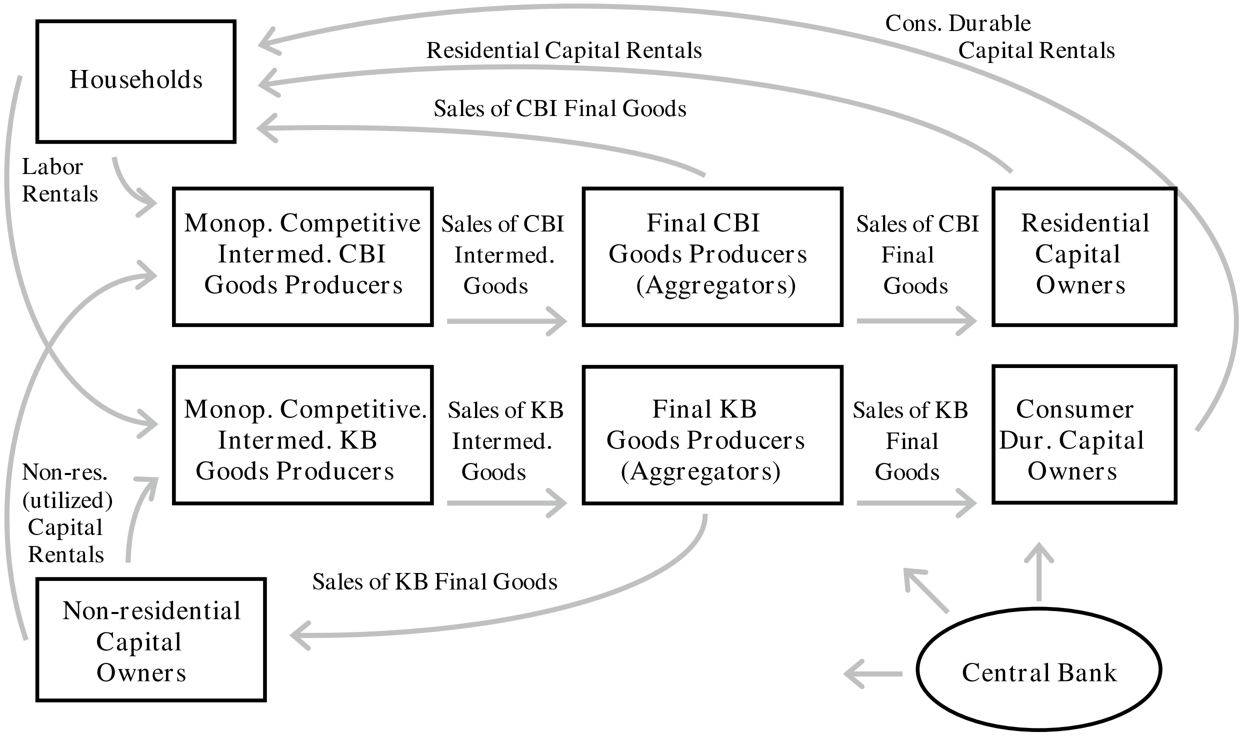

Figure 1: Model Overview

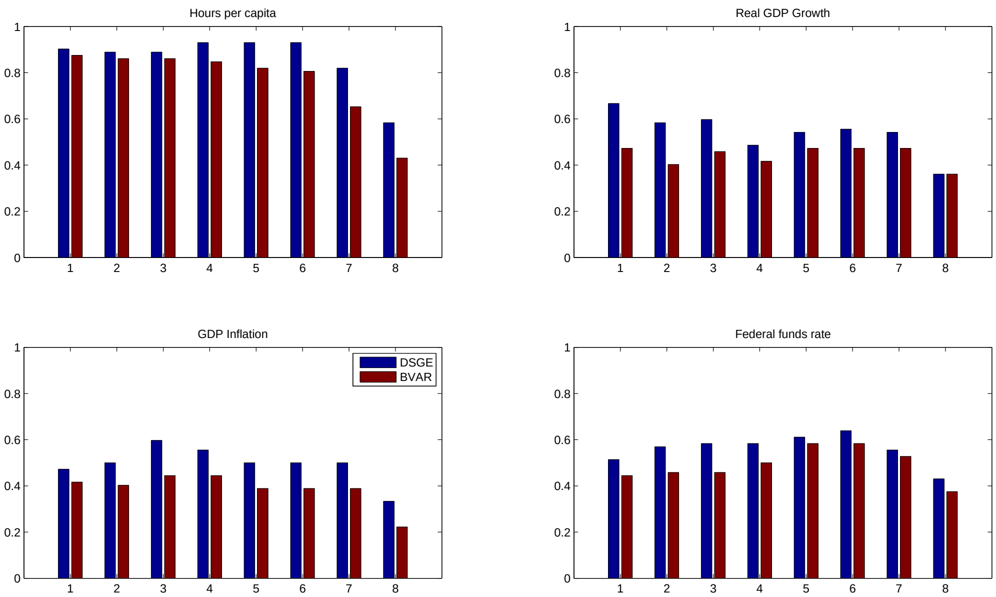

Figure 2: Bias in Forecasts

Appendix: Information on Greenbook Forecasts and Real-time Data

GB Name	Date GB Closed	Date(s) of FOMC Meeting	Estim. Period, DSGE Model	Last Qtr. of History	GB Forecast Horizon
Sep. 96	Sep. 18, 96	Sep. 24, 96	85:Q1-96:Q2	96:Q2	96:Q3-98:Q4
Interim NIPA releases: 96:Q2 Final (9/27/96), 96:Q3 Advan. (10/30/96)					
Nov. 96	Nov. 6, 96	Nov. 13, 96	85:Q1-96:Q2	96:Q3	96:Q4-98:Q4
Interim NIPA releases: 96:Q3 Prelim. (11/27/96)					
Dec. 96	Dec. 12, 96	Dec. 17, 96	85:Q1-96:Q2	96:Q3	96:Q4-98:Q4
Interim NIPA releases: 96:Q3 Final (12/20/96)					
Jan. 97	Jan. 29, 97	Feb. 4 & 5, 97	85:Q1-96:Q2	96:Q4	97:Q1-98:Q4
Interim NIPA releases: 96:Q4 Advan. (1/31/97), 96:Q4 Prelim. (2/28/97)					
Mar. 97	Mar. 19, 97	Mar. 25, 97	85:Q1-96:Q2	96:Q4	97:Q1-98:Q4
Interim NIPA releases: 96:Q4 Final (3/28/97), 97:Q1 Advan. (4/30/97)					
May 97	May 15, 97	May 20, 97	85:Q1-96:Q2	97:Q1	97:Q2-98:Q4
Interim NIPA releases: 97:Q1 Prelim. (5/30/97)					
Jun. 97	Jun. 25, 97	Jul. 1 & 2, 97	85:Q1-96:Q2	97:Q1	97:Q2-98:Q4
Interim NIPA releases: 97:Q1 Final (6/27/97), 97:Q2 Advan. & 94-96 Annual Revision (7/31/97)					
Aug. 97	Aug. 14, 97	Aug. 19, 97	85:Q1-97:Q2	97:Q2	97:Q3-98:Q4
Interim NIPA releases: 97:Q2 Prelim. (8/28/97)					
Sep. 97	Sep. 24, 97	Sep. 30, 97	85:Q1-97:Q2	97:Q2	97:Q3-99:Q4
Interim NIPA releases: 97:Q2 Final (9/26/97), 97:Q3 Advan. (10/31/97)					
Nov. 97	Nov. 6, 97	Nov. 12, 97	85:Q1-97:Q2	97:Q3	97:Q4-99:Q4
Interim NIPA releases: 97:Q3 Prelim. (11/26/97)					
Dec. 97	Dec. 11, 97	Dec. 16, 97	85:Q1-97:Q2	97:Q3	97:Q4-99:Q4
Interim NIPA releases: 97:Q3 Final (12/23/97)					
Jan. 98	Jan. 28, 98	Feb. 3 & 4, 98	85:Q1-97:Q2	97:Q4	98:Q1-99:Q4
Interim NIPA releases: 97:Q4 Advan. (1/30/98), 97:Q4 Prelim. (2/27/98)					
Mar. 98	Mar. 19, 98	Mar. 25, 98	85:Q1-97:Q2	97:Q4	98:Q1-99:Q4
Interim NIPA releases: 97:Q4 Final (3/26/98), 98:Q1 Advan. (4/30/98)					
May 98	May 14, 98	May 19, 98	85:Q1-97:Q2	98:Q1	98:Q2-99:Q4
Interim NIPA releases: 97:Q1 Prelim. (5/28/98)					

Table A.1: Greenbook and NIPA Release Dates (Sep. 96 to May 98).

GB Name	Date GB Closed	Date(s) of FOMC Meeting	Estim. Period, DSGE Models	Last Qtr. of History	GB Forecast Horizon
Jun. 98	Jun. 24, 98	Jun. 30 & Jul. 1, 98	85:Q1-97:Q2	98:Q1	98:Q2-99:Q4
Interim NIPA releases: 97:Q1 Final (6/25/98), 98:Q2 Advan. & 95-97 Annual Revision (7/31/98)					
Aug. 98	Aug. 13, 98	Aug. 18, 98	85:Q1-98:Q2	98:Q2	98:Q3-99:Q4
Interim NIPA releases: 98:Q2 Prelim. (8/27/98)					
Sep. 98	Sep. 23, 98	Sep. 29, 98	85:Q1-98:Q2	98:Q2	98:Q3-00:Q4
Interim NIPA releases: 98:Q2 Final (9/24/98), 98:Q3 Advan. (10/30/98)					
Nov. 98	Nov. 13, 98	Nov. 17, 98	85:Q1-98:Q2	98:Q3	98:Q4-00:Q4
Interim NIPA releases: 98:Q3 Prelim. (11/24/98)					
Dec. 98	Dec. 16, 98	Dec. 22, 98	85:Q1-98:Q2	96:Q3	98:Q4-00:Q4
Interim NIPA releases: 98:Q3 Final (12/23/98)					
Jan. 99	Jan. 28, 99	Feb. 2 & 3, 99	85:Q1-98:Q2	98:Q4	99:Q1-00:Q4
Interim NIPA releases: 98:Q4 Advan. (1/29/99), 98:Q4 Prelim. (2/26/99)					
Mar. 99	Mar. 24, 99	Mar. 30, 99	85:Q1-98:Q2	98:Q4	99:Q1-00:Q4
Interim NIPA releases: 98:Q4 Final (3/31/99), 98:Q1 Advan. (4/30/99)					
May 99	May 13, 99	May 18, 99	85:Q1-98:Q2	99:Q1	99:Q2-00:Q4
Interim NIPA releases: 99:Q1 Prelim. (5/27/99)					
Jun. 99	Jun. 23, 99	Jun. 29 & 30, 99	85:Q1-98:Q2	99:Q1	99:Q2-00:Q4
Interim NIPA releases: 99:Q1 Final (6/25/99), 99:Q2 Advan. (7/29/99)					
Aug. 99	Aug. 18, 99	Aug. 24, 99	85:Q1-99:Q2	99:Q2	99:Q3-00:Q4
Interim NIPA releases: 99:Q2 Prelim. (8/26/99)					
Sep. 99	Sep. 29, 99	Oct. 5, 99	85:Q1-99:Q2	99:Q2	99:Q3-01:Q4
Interim NIPA releases: 99:Q2 Final (9/30/99), 99:Q3 Advan. & Comprehensive Revision (10/28/99)					
Nov. 99	Nov. 10, 99	Nov. 16, 99	85:Q1-99:Q3	96:Q3	99:Q4-01:Q4
Interim NIPA releases: 99:Q3 Prelim. (11/24/99)					
Dec. 99	Dec. 15, 99	Dec. 21, 99	85:Q1-99:Q3	99:Q3	99:Q4-01:Q4
Interim NIPA releases: 99:Q3 Final (12/22/99)					
Jan. 00	Jan. 27, 00	Feb. 1 & 2, 00	85:Q1-99:Q3	99:Q4	00:Q1-01:Q4
Interim NIPA releases: 99:Q4 Advan. (1/28/00), 99:Q4 Prelim. (2/25/00)					

Table A.2: Greenbook and NIPA Release Dates (Jun. 98 to Jan. 00).

GB Name	Date GB Closed	Date(s) of FOMC Meeting	Estim. Period, DSGE Models	Last Qtr. of History	GB Forecast Horizon
Mar. 00	Mar. 15, 00	Mar. 21, 00	85:Q1-99:Q3	99:Q4	00:Q1-01:Q4
Interim NIPA releases: 99:Q4 Final (3/30/00), 99:Q1 Advan. (4/27/00)					
May 00	May 11, 00	May 16, 00	85:Q1-99:Q3	00:Q1	00:Q2-01:Q4
Interim NIPA releases: 00:Q1 Prelim. (5/25/00)					
Jun. 00	Jun. 21, 00	Jun. 27 & 28, 00	85:Q1-99:Q3	00:Q1	00:Q2-01:Q4
Interim NIPA releases: 00:Q1 Final (6/27/00), 00:Q2 Advan. & 97-99 Annual Revision (7/28/00)					
Aug. 00	Aug. 16, 00	Aug. 22, 00	85:Q1-00:Q2	00:Q2	00:Q3-01:Q4
Interim NIPA releases: 00:Q2 Prelim. (8/25/00)					
Sep. 00	Sep. 27, 00	Oct. 3, 00	85:Q1-00:Q2	00:Q2	00:Q3-02:Q4
Interim NIPA releases: 00:Q2 Final (9/28/00), 00:Q3 Advan. (10/27/00)					
Nov. 00	Nov. 8, 00	Nov. 15, 00	85:Q1-00:Q2	00:Q3	00:Q4-02:Q4
Interim NIPA releases: 00:Q3 Prelim. (11/29/00)					
Dec. 00	Dec. 13, 00	Dec. 19, 00	85:Q1-00:Q2	00:Q3	00:Q4-02:Q4
Interim NIPA releases: 00:Q3 Final (12/21/00)					
Jan. 01	Jan. 25, 01	Jan. 30 & 31, 01	85:Q1-00:Q2	00:Q4	01:Q1-02:Q4
Interim NIPA releases: 00:Q4 Advan. (1/31/01), 00:Q4 Prelim. (2/28/01)					
Mar. 01	Mar. 14, 01	Mar. 20, 01	85:Q1-00:Q2	00:Q4	01:Q1-02:Q4
Interim NIPA releases: 00:Q4 Final. (3/20/01), 01:Q1 Advan. (4/27/01)					
May. 01	May 9, 01	May. 15, 01	85:Q1-00:Q2	01:Q1	01:Q2-02:Q4
Interim NIPA releases: 01:Q1 Prelim (5/18/01)					
Jun. 01	Jun. 20, 01	Jun. 26 & 27, 01	85:Q1-00:Q2	01:Q1	01:Q2-02:Q4
Interim NIPA releases: 01:Q1 Final. (6/28/01), 01:Q2 Advan. (7/27/01)					
Aug. 01	Aug. 15, 01	Aug. 21, 01	85:Q1-01:Q2	01:Q2	01:Q3-02:Q4
Interim NIPA releases: 01:Q2 Prelim. (8/17/01), 01:Q2 Final.(9/24/01)					
Sept. 01	Sep. 26, 01	Oct. 2, 01	85:Q1-01:Q2	01:Q2	01:Q3-03:Q4
Interim NIPA releases: 01:Q3 Advan. (10/31/01)					
Nov. 01	Oct. 31, 01	Nov. 6, 01	85:Q1-01:Q2	01:Q3	01:Q4-03:Q4
Interim NIPA releases: 01:Q3 Prelim. (11/20/01)					

Table A.3: Greenbook and NIPA Release Dates (Mar. 00 to Nov. 01).

GB Name	Date GB Closed	Date(s) of FOMC Meeting	Estim. Period, DSGE Models	Last Qtr. of History	GB Forecast Horizon
Dec 01	Dec. 5, 01	Dec 11, 01	85:Q1-01:Q2	01:Q3	01:Q4-03:Q4
Interim NIPA releases: 01:Q3 Final. (12/19/01)					
Jan. 02	Jan. 23, 02	Jan. 29 & 30, 02	85:Q1-01:Q2	01:Q4	02:Q1-03:Q4
Interim NIPA releases: 01:Q4 Advan. (1/30/02), 01:Q4 Prelim. (2/28/02)					
Mar. 02	Mar. 13, 02	Mar. 19, 02	85:Q1-01:Q2	01:Q4	02:Q1-03:Q4
Interim NIPA releases: 01:Q4 Final. (3/19/02), 02:Q1 Advan. (4/26/02)					
May. 02	May. 1, 02	May. 7, 02	85:Q1-01:Q2	02:Q1	02:Q2-03Q4
Interim NIPA releases: 02:Q1 Prelim. (5/17/02)					
Jun. 02	Jun. 19, 02	Jun. 25 & 26, 02	85:Q1-01:Q2	02:Q1	02:Q2-03:Q4
Interim NIPA releases: 02:Q1 Final. (6/20/02), 02:Q2 Advan. (7/31/02)					
Aug. 02	Aug. 7, 02	Aug. 13, 02	85:Q1-02:Q2	02:Q2	02:Q3-03:Q4
Interim NIPA releases: 02:Q2 Prelim. (8/20/02)					
Sep. 02	Sep. 18, 02	Sep. 24, 02	85:Q1-02:Q2	02:Q2	02:Q3-04:Q4
Interim NIPA releases: 02:Q2 Final. (9/23/02), 02:Q3 Advan. (10/31/02)					
Nov. 02	Oct. 30, 02	Nov. 6, 02	85:Q1-02:Q2	02:Q3	02:Q4-04:Q4
Interim NIPA releases: 02:Q3 Prelim. (11/19/02)					
Dec. 02	Dec. 4, 02	Dec. 10, 02	85:Q1-02:Q2	02:Q3	02:Q4-04:Q4
Interim NIPA releases: 02:Q3 Final. (12/18/02)					
Jan. 03	Jan. 22, 03	Jan. 28 & 29, 03	85:Q1-02:Q2	02:Q4	03:Q1-04:Q4
Interim NIPA releases: 02:Q4 Advan. (1/30/03), 02:Q4 Prelim. (2/28/03)					
Mar. 03	Mar. 12, 03	Mar. 18, 03	85:Q1-02:Q2	02:Q4	03:Q1-04:Q4
Interim NIPA releases: 02:Q4 Final. (3/14/03), 03:Q1 Advan. (4/25/03)					
May. 03	Apr. 30, 03	May. 6, 03	85:Q1-02:Q2	03:Q1	03:Q2-04:Q4
Interim NIPA releases: 03:Q1 Prelim. (5/22/03)					
Jun. 03	Jun. 18, 03	Jun. 24 & 25, 03	85:Q1-02:Q2	03:Q1	03:Q2-04:Q4
Interim NIPA releases: 03:Q1 Final. (6/19/03), 03:Q2 Advan. (7/31/03)					
Aug 03	Aug 6, 03	Aug 12, 03	85:Q1-03:Q2	03:Q2	03:Q3-04:Q4
Interim NIPA releases: 03:Q2 Prelim. (8/14/03)					

Table A.4: Greenbook and NIPA Release Dates (Dec. 01 to Aug. 03).

GB Name	Date GB Closed	Date(s) of FOMC Meeting	Estim. Period, DSGE Models	Last Qtr. of History	GB Forecast Horizon
Sep. 03	Sep. 10, 03	Sep. 16, 03	85:Q1-03:Q2	03:Q2	03:Q3-05:Q4
Interim NIPA releases: 03:Q2 Final. (9/15/03)					
Oct. 03	Oct. 22, 03	Oct. 28, 03	85:Q1-03:Q2	03:Q3	03:Q4-05:Q4
Interim NIPA releases: 03:Q3 Advan. (10/30/03), 03:Q3 Prelim. (11/13/03)					
Dec. 03	Dec. 3, 03	Dec. 9, 03	85:Q1-03:Q2	03:Q3	03:Q4-05:Q4
Interim NIPA releases: 03:Q3 Final. (12/16/03)					
Jan. 04	Jan. 21, 04	Jan. 27 & 28, 04	85:Q1-03:Q2	03:Q4	04:Q1-05:Q4
Interim NIPA releases: 03:Q4 Advan. (1/30/04), 03:Q4 Prelim. (2/27/04)					
Mar. 04	Mar. 10, 04	Mar. 16, 04	85:Q1-03:Q2	03:Q4	04:Q1-05:Q4
Interim NIPA releases: 03:Q4 Final. (3/25/03), 04:Q1 Advan. (4/29/04)					
May. 04	Apr. 28, 04	May. 4, 04	85:Q1-03:Q2	04:Q1	04:Q2-05:Q4
Interim NIPA releases:04:Q1 Prelim. (5/27/04)					
Jun. 04	Jun. 23, 04	Jun. 29 & 30, 04	85:Q1-03:Q2	04:Q1	04:Q2-05:Q4
Interim NIPA releases: 04:Q1 Final. (6/25/04), 04:Q2 Advan. (7/30/04)					
Aug. 04	Aug. 4, 04	Aug. 10, 04	85:Q1-04:Q2	04:Q2	04:Q3-05:Q4
Interim NIPA releases: 04:Q2 Prelim. (8/27/04)					
Sep. 04	Sep. 15, 04	Sep. 21, 04	85:Q1-04:Q2	04:Q2	04:Q3-06:Q4
Interim NIPA releases: 04:Q2 Final. (9/29/04), 04:Q3 Advan. (10/29/04)					
Nov. 04	Nov. 3, 04	Nov. 10, 04	85:Q1-04:Q2	04:Q3	04:Q4-06:Q4
Interim NIPA releases: 04:Q3 Prelim. (11/30/04)					
Dec 04	Dec 8, 04	Dec 14, 04	85:Q1-04:Q2	04:Q3	04:Q4-06:Q4
Interim NIPA releases: 04:Q3 Final. (12/22/04)					
Feb. 05	Jan. 26, 05	Feb. 1 & 2, 05	85:Q1-04:Q2	04:Q4	05:Q1-06:Q4
Interim NIPA releases: 04:Q4 Advan. (1/28/05), 04:Q4 Prelim. (2/25/05)					
Mar. 05	Mar. 16, 05	Mar. 22, 05	85:Q1-04:Q2	04:Q4	05:Q1-06:Q4
Interim NIPA releases: 04:Q4 Final. (3/30/05), 05:Q1 Advan. (4/28/05)					
May. 05	Apr. 28, 05	May. 3, 05	85:Q1-04:Q2	05:Q1	05:Q2-06:Q4
Interim NIPA releases: 05:Q1 Prelim. (5/26/05)					
Jun. 05	Jun. 22, 05	Jun. 29 & 30, 05	85:Q1-04:Q2	05:Q1	05:Q2-06:Q4
Interim NIPA releases: 05:Q1 Final. (6/29/05), 05:Q2 Advan. (7/29/05)					

Table A.5: Greenbook and NIPA Release Dates (Sep. 03 to Jun. 05).