WORKING PAPER NO. 14-29
ANALYZING DATA REVISIONS WITH A DYNAMIC
STOCHASTIC GENERAL EQUILIBRIUM MODEL

Dean Croushore
Professor of Economics and Rigsby Fellow
University of Richmond
and
Visiting Scholar, Federal Reserve Bank of Philadelphia

Keith Sill
Federal Reserve Bank of Philadelphia

September 23, 2014

Analyzing Data Revisions with a Dynamic Stochastic General Equilibrium Model

Dean Croushore

University of Richmond

Keith Sill[*]

*Federal Reserve Bank
of Philadelphia*

September 23, 2014

Abstract

We use a structural dynamic stochastic general equilibrium model to investigate how initial data releases of key macroeconomic aggregates are related to final revised versions and how identified aggregate shocks influence data revisions. The analysis sheds light on how well preliminary data approximate final data and on how policy makers might condition their view of the preliminary data when formulating policy actions. The results suggest that monetary policy shocks and multifactor productivity shocks lead to predictable revisions to the initial release data on output growth and inflation.

1 Introduction

Much of the data used by policymakers to assess economic conditions and evaluate alternative policy actions is subject to ongoing revision. Data are often still being revised months or years after policy decisions are made. What is the relationship between the early release data, which are unrevised or lightly revised, and the final release data that ultimately reveal a clearer picture of the economy? Do initial release and final release data respond in systematically different ways to identified economic shocks? We investigate this issue using a structural DSGE model framework that links early release data to state variables derived from fully revised data.

Consider the case of the quarterly GDP series released by the Bureau of Economic Analysis. There are three monthly releases of a given quarter's output, followed by three annual revisions and a sequence of benchmark revisions that occur approximately every five years thereafter. Some series, such as the unemployment rate and CPI inflation are largely unrevised, but this is more the exception than the norm. We can view the early estimates of economic variables like GDP as forecasts of a true measure of GDP that will be revealed as more source data are finalized. For the most part, though, the benchmark structural models that are used to assess the stance of monetary policy and alternative policy actions do not take explicit account of the data revision process. Rather, the models are largely constructed and estimated under the assumption that data are not revised and that the observations in hand are good measures of the underlying economic conditions, even at the tail end of the sample.

Kishor and Koenig (2012) highlight an asymmetric treatment of the data in the use of models estimated using data that contain early observations that are heavily revised and late observations that are lightly revised (or not revised at all). That estimation typically ignores the fact that data toward the end of the sample have undergone little or no revision while data at the beginning of the sample are heavily revised. This raises a concern that the data-generating process for the earlier parts of the sample may differ from that at the end of the sample. For example, when generating forecasts or undertaking current policy analysis, the key end-of-sample data are lightly revised while the model parameters and filtered history of latent variables are largely based on heavily revised data that may have very different statistical properties.

We investigate the relationship between initial release and final release data using a structural DSGE model that offers a clear channel from identified economic shocks to the

dynamics of observed macroeconomic variables. We use a DSGE model to investigate data revisions but for tractability we treat initial release data as a nonmodeled auxiliary variable that is a function of the DSGE model's latent state variables. The DSGE model itself is viewed as being informative about an economic structure that characterizes final release, fully revised data, i.e., the "truth." The model framework allows us to identify and recover the economic shocks that are most important in accounting for the variance of real output growth and inflation in the revised data. We investigate how initial release data are related to these shocks and how the dynamic responses of initial release data to economic shocks differ from final release responses.

This paper is based on the empirical method of Schorfheide et al. (2010). That paper uses a simple two-step estimation approach, based on an empirical model that consists of a medium-scale DSGE model for a set of core macroeconomic variables and a set of measurement equations or auxiliary regressions that link the state variables of the DSGE model to noncore variables. The first step in that approach is to estimate the DSGE model using the core variables as measurements. Based on the estimates of the DSGE model parameters, the Kalman filter is used to get estimates of the latent state variables. The filtered state variables are then used as regressors to estimate simple linear measurement equations with serially correlated idiosyncratic errors.

The literature on real-time data analysis has largely focused on examining the size of data revisions and their impact on forecasts and monetary policy. Evidence from the real-time literature, described in Croushore (2011), suggests that such revisions may be crucial for forecasting and policy analysis. In forecasting, data revisions may change the estimated parameters of forecasting models, the specification of the model, and the jumping-off point for forecasts. Policy may be formulated in error if it relies too much on data that are measured with error and that are subject to revision. The implementation of sound monetary policy would seem to require an examination of the significance of data revisions and the role they play when combining structural models with latest-vintage data for current analysis.

We do not provide an explicit structural model of how initial release data are generated. Instead, we think of a government data agency as following a protocol for releasing data in which it first makes initial releases of the data based on a small sample. Over time, the sample grows in size, and the agency provides more precise releases of the data. Initially, the agency is forecasting many components of the aggregate data, so its early releases are similar to forecasts. We also compare the dynamic responses of initial release data with

forecasts made by professional forecasters and find that they are quite similar. As the data agency gets additional source data, the responses of its data releases to identified shocks conform more to those implied by the estimated structural DSGE model.

The paper proceeds as follows. In section 2, we describe the medium-scale DSGE model that we use. Section 3 describes the methods we use to estimate the model. Section 4 discusses the results of the model's estimation and shows the relationship between data revisions and estimated structural shocks. Section 5 interprets the results and discusses their importance.

2   A Medium-Scale DSGE Model

We begin with a brief description of the medium-scale New Keynesian model that underlies our empirics. The model incorporates many of the main elements that are standard in the New Keynesian DSGE literature, including habit formation, costs of adjusting capital investment, wage and price rigidities, and variable capital utilization. The baseline model is similar to Smets and Wouters (2003) and Christiano et al. (2005), and the specific log-linearized implementation is described in more detail in Schorfheide et al. (2010).

2.1   Final Goods Producers

There is a final good Y_t that is produced as a composite of a continuum of intermediate goods $Y_t(i)$ using the technology:

$$Y_t = \left[\int_0^1 Y_t(i)^{\frac{1}{1+\lambda_{f,t}}} \, di \right]^{1+\lambda_{f,t}} \tag{1}$$

with $\lambda_{f,t} \in (0, \infty)$ following the exogenous process:

$$\ln \lambda_{f,t} = (1 - \rho_{\lambda_f}) \ln \lambda_f + \rho_{\lambda_f} \ln \lambda_{f,t-1} + \sigma_{\lambda_f} \epsilon_{\lambda,t}. \tag{2}$$

The variable $\lambda_{f,t}$ is the desired markup over marginal cost that intermediate goods producers would like to charge. From the first-order conditions for profit maximization and the zero-profit condition (final goods producers are perfectly competitive firms) the demand for intermediate goods is given by:

$$Y_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\frac{1+\lambda_{f,t}}{\lambda_{f,t}}} Y_t \tag{3}$$

with the composite good price given by:

$$P_t = \left[\int_0^1 P_t(i)^{-\frac{1}{\lambda_{f,t}}} \, di \right]^{-\lambda_{f,t}} \tag{4}$$

2.2   Intermediate Goods Producers

There is a continuum of intermediate goods indexed by i. Intermediate goods are produced using the technology:

$$Y_t(i) = Z_t^{1-\alpha} K_t(i)^\alpha L_t(i)^{1-\alpha}, \tag{5}$$

where Z_t is an exogenous technological progress that is assumed to be nonstationary. We define $z_t = \ln(Z_t/Z_{t-1})$ and assume that it follows the process:

$$(z_t - \gamma) = \rho_z(z_{t-1} - \gamma) + \varepsilon_{z,t}.$$

Prices are assumed to be sticky and adjust following Calvo (1983). Each firm can readjust prices optimally with probability $1 - \zeta_p$ in each period. Firms that are unable to reoptimize their prices $P_t(i)$ adjust prices mechanically according to:

$$P_t(i) = (\pi_{t-1})^{\iota_p} (\pi_*)^{1-\iota_p} \tag{6}$$

where $\pi_t = P_t/P_{t-1}$ and π_* is the steady state inflation rate of the final good. Those firms that reoptimize price choose a price level $\tilde{P}_t(i)$ that maximizes the expected present discounted value profits in all states of nature in which the firm maintains that price in the future:

$$\begin{aligned} max_{\tilde{P}_t(i)} \quad & \Xi_r^p \left(\tilde{P}_t(i) - MC_t \right) Y_t(i) + \\ E_t \sum_{s=1}^\infty & \zeta_p^s \beta^s \Xi_{t+s}^p \left(\tilde{P}_t(i)(\Pi_{l=1}^s \pi_{t+l-1}^{\iota_p} \pi_*^{1-\iota_p}) - MC_{t+s} \right) Y_{t+s} \end{aligned} \tag{7}$$

subject to

$$Y_{t+s}(i) = \left(\frac{\tilde{P}_t(i) \left(\Pi_{l=1}^s \pi_{t+l-1}^{\iota_p} \pi_*^{1-\iota_p} \right)}{P_{t+s}} \right)^{-\frac{1+\lambda_{f,t}}{\lambda_{f,t}}} Y_{t+s},$$

where $\pi_t \equiv P_t/P_{t-1}$, $\beta^s \Xi_{t+s}^p$ is the household's discount factor, and MC_t is the firm's marginal cost. Markets are assumed to be complete so all households face the same discount factor. All firms that can readjust price face an identical problem. We will consider only a symmetric equilibrium in which all adjusting firms choose the same price, which means that we can drop the i index. It then follows that the aggregate price level can be expressed as follows:

$$P_t = \left[(1 - \zeta_p) \tilde{P}_t^{-\frac{1}{\lambda_f}} + \zeta_p \left(\pi_{t-1}^{\iota_p} \pi_*^{1-\iota_p} P_{t-1} \right)^{-\frac{1}{\lambda_f}} \right]^{-\lambda_f}.$$

In the estimation, we shut down inflation indexation by setting $\iota_p = 0$.

2.3 Households

The objective function for household j is given by:

$$E_t \sum_{s=0}^{\infty} b_{t+s} \left[\ln(C_{t+s}(j) - hC_{t+s-1}(j)) - \frac{\varphi_{t+s}}{1 + \nu_l} L_{t+s}(j)^{1+\nu_l} + \frac{\chi_{t+s}}{1 - \nu_m} \left(\frac{M_{t+s}(j)}{Z_{t+s} P_{t+s}} \right)^{1-\nu_m} \right],$$

where $C_t(j)$ is consumption, $L_t(j)$ is labor supply, and $M_t(j)$ is money holdings. Household preferences are subject to three shocks: an intertemporal shifter b_t, a labor supply shock φ_t, and a money demand shock χ_t. Real balances are deflated by the stochastic trend growth to make real money demand stationary. All preference shocks are assumed to follow an AR(1) process in logs. The household budget constraint, written in nominal terms, is given by:

$$P_{t+s} C_{t+s}(j) + P_{t+s} I_{t+s}(j) + B_{t+s}(j) \leq R_{t+s} B_{t+s-1}(j) + M_{t+s-1}(j) +$$
$$\Pi_{t+s} + W_{t+s}(j) L_{t+s}(j) + R_{t+s}^k u_{t+s}(j) \hat{K}_{t+s-1}(j) - P_{t+s} a(u_{t+s}(j)) \hat{K}_{t+s-1}(j),$$

where $I_t(j)$ is investment, $\hat{K}_t(j)$ is capital holdings, $u_t(j)$ is the rate of capital utilization, and $B_t(j)$ is holdings of government bonds. The gross nominal interest rate paid on government bonds is R_t, and Π_t is the per-capita profit the household gets from owning firms. Household labor is paid wage $W_t(j)$, and households rent an "effective" amount of capital to firms $K_t(j) = u_t(j)\hat{K}_{t-1}(j)$. In return, they receive $R_t^k u_t(j)\hat{K}_{t-1}(j)$. Households pay a consumption cost associated with capital utilization given by $a(u_t(j))\hat{K}_{t-1}(j)$. Capital accumulation is governed by:

$$\hat{K}_t(j) = (1 - \delta)\hat{K}_{t-1}(j) + \mu_t \left(1 - S \left(\frac{I_t(j)}{I_{t-1}(j)} \right) \right) I_t(j),$$

where δ is the rate of depreciation, $S(\cdot)$ is the cost of adjusting investment ($S' > 0$, $S'' > 0$), and μ_t is a stochastic shock to the price of investment relative to consumption, assumed to follow an AR(1) process in logs.

2.4 The Labor Market

The labor market has labor packers that buy labor from households, combine it, and resell it to the intermediate goods producing firms. Labor used by the intermediate goods producers is a composite:

$$L_t = \left[\int_0^1 L_t(j)^{\frac{1}{1+\lambda_{w,t}}} \, dj \right]^{1+\lambda_{w,t}}$$

The labor packers maximize profits in a perfectly competitive environment, which leads to the labor demand:

$$L_t(j) = \left(\frac{W_t(j)}{W_t} \right)^{-\frac{1+\lambda_{w,t}}{\lambda_{w,t}}}.$$

5

Combining labor demand with the zero-profit condition leads to the aggregate wage expression:

$$W_t = \left[\int_0^1 W_t(j)^{\frac{1}{\lambda_{w,t}}} dj \right]^{\lambda_{w,t}}.$$

In the estimation, we fix $\lambda_{w,t} = \lambda_w \in (0, \infty)$. Households have market power, but wage adjustment is subject to a rigidity as in Calvo (1983). Each period, a fraction $1 - \zeta_w$ of households reoptimize their wage. For those that are unable to reoptimize, $W_t(j)$ adjusts as a geometric average of the steady state rate increase in wages and last period's productivity times last period's inflation. For those households that can reoptimize, the problem is to choose a wage $\tilde{W}_t(j)$ that maximizes utility in all states of nature in which the household wage is to be held at its chosen value:

$$max_{\tilde{W}_t(j)} E_t \sum_{s=0}^{\infty} (\zeta_w \beta)^s b_{t+s} \left[-\frac{\varphi_{t+s}}{1 + \nu_l} L_{t+s}(j)^{1+\nu_l} + \ldots \right]$$

subject to

$$W_{t+s}(j) = \left(\Pi_{l=1}^s (\pi_*)^{1-\iota_w} (\pi_{t+l-1} e^{z^*_{t+l-1}})^{\iota_w} \right) \tilde{W}_t(j)$$

for $s = 1, \ldots, \infty$ as well as to the household budget constraint and the labor demand condition. In the estimation, we shut down nominal wage indexation by setting $\iota_w = 0$.

2.5   Government Policies

The government consists of a fiscal authority and a monetary authority. The monetary authority sets the nominal interest rate according to the feedback rule:

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R} \right)^{\rho_R} \left[\left(\frac{\pi_t}{\pi_*} \right)^{\psi_R} \left(\frac{Y_t}{Y} \right)^{\psi_Y} \right]^{1-\rho_R} \epsilon_{R,t}.$$

The fiscal authority balances its budget by issuing short-term bonds. Government spending is exogenous and given by:

$$G_t = (1 - 1/g_t) Y_t,$$

where the government spending shock g_t is assumed to follow an AR(1) process.

2.6   Exogenous Processes

There are seven exogenous shocks in the model. These follow the processes:

- Technology process. Let $z_t = \ln(Z_t/Z_{t-1})$:

$$(z_t - \gamma) = \rho_z(z_{t-1} - \gamma) + \sigma_z \epsilon_{z,t}$$

- Preference for leisure:

$$\ln \phi_t = (1 - \rho_\phi) \ln \phi + \rho_\phi \ln \phi_{t-1} + \sigma_\phi \epsilon_{\phi,t}$$

- Money demand:

$$\ln \chi_t = (1 - \rho_\chi) \ln \chi + \rho_\chi \ln \chi_{t-1} + \sigma_\chi \epsilon_{\chi,t}$$

- Price-markup shock:

$$\ln \lambda_{f,t} = (1 - \rho_{\lambda_f}) \ln \lambda + \rho_\lambda \ln \lambda_{f,t-1} + \sigma_\lambda \epsilon_{\lambda_f,t}$$

- Capital adjustment cost (marginal efficiency of investment):

$$\ln \mu_t = (1 - \rho_\mu) \ln \mu + \rho_\mu \ln \mu_{t-1} + \sigma_\mu \epsilon_{\mu,t}$$

- Intertemporal preference shifter:

$$\ln b_t = \rho_b \ln b_{t-1} + \sigma_b \epsilon_{b,t}$$

- Government spending shock:

$$\ln g_t = (1 - \rho_g) \ln g + \rho_g \ln g_{t-1} + \sigma_g \epsilon_{g,t}$$

- Monetary policy shock:

$$\epsilon_{R,t}$$

2.7 Model Solution

The model's real variables, output Y_t, consumption C_t, investment I_t, capital K_t, effective capital $\hat{K}_t$, and real wage W_t/P_t all grow at the same rate as Z_t. These variables are detrended while the nominal interest rate, inflation, and hours worked are stationary. A steady state is constructed for the stationary representation of the model. The method of Sims (2002) is used to construct a log-linear approximation of the model around its steady state.

3 Estimation

3.1 DSGE Model Estimation

Following the discussion and methodology described in Schorfheide et al. (2010) we log-linearize the rational expectations model, which can then be expressed as a vector autoregressive law of motion for a vector of fundamental state variables. Calling these variables s_t, the law of motion is given by:

$$s_t = \Phi_1(\theta)s_{t-1} + \Phi_\epsilon(\theta)\epsilon_t. \tag{8}$$

The coefficients of the matrices Φ_1, and Φ_ϵ are functions of the DSGE model parameters θ and the vector s_t is given by

$$s_t = [c_t, i_t, \bar{k}_t, R_t, w_t, z_t, \phi_t, \mu_t, b_t, g_t, \lambda_{f,t}]'.$$

The variables c_t, i_t, $\bar{k}_t$, R_t, and w_t are endogenous state variables and the remaining elements of s_t are exogenous state variables.

To estimate the DSGE model based on a sequence of observations $Y^T = [y_t, \ldots, y_T]$, we construct a state-space model that links the observable variables y_t to the state variables via a system of measurement equations.

The vector of observables y_t in the measurement equation comprises quarter-to-quarter growth rates (measured in percentages) of real GDP, consumption, investment, and nominal wages, as well as a measure of hours worked, GDP deflator inflation, and the federal funds rate. The use of growth rates for some variables requires the set of model states to be augmented by lagged values of output, consumption, investment, and real wages. Since lagged consumption, investment, and real wages are elements of the vector s_{t-1}, and lagged output, y_{t-1}, can be expressed as a linear function of the elements of s_{t-1}, we can write

$$[y_{t-1}, c_{t-1}, i_{t-1}, w_{t-1}]' = M_s(\theta)s_{t-1}$$

for a suitably chosen matrix $M_s(\theta)$ and define

$$\varsigma_t = [s'_t, s'_{t-1}M'_s(\theta)]'. \tag{9}$$

This allows us to express the set of measurement equations as

$$y_t = A_0(\theta) + A_1(\theta)\varsigma_t. \tag{10}$$

8

The state-space representation of the DSGE model comprises (8), (9), and (10).

Assuming that the innovations ϵ_t are normally distributed, the likelihood function for the structural model, denoted by $p(Y^T|\theta)$, can be evaluated with the Kalman filter, which can also be used to generate estimates of the state vector:

$$\varsigma_{t|t}(\theta) = I\!E[\varsigma_t|\theta, Y^t]. \tag{11}$$

Bayesian estimation of the model combines a prior $p(\theta)$ with the likelihood function $p(Y^T|\theta)$ to obtain a joint probability density function for data and parameters. We use Markov-Chain-Monte-Carlo (MCMC) methods as described in An and Schorfheide (2007) to implement the Bayesian inference.

We view the structural DSGE model as the data generating process for the final release version of the data, which has undergone a thorough set of revisions and so represents the "truth." Operationally, we estimate the model using data that have been substantially revised. In particular, we estimate the DSGE model parameters using the data vintage available as of 2013Q3, but the effective sample period for the estimation runs from 1984Q1 through 2010Q4. Thus, for our sample, the 2010Q4 end point has undergone two annual revisions and one comprehensive benchmark revision. If we were to add more recent data to the sample in the form of extending the sample period beyond 2010Q4, we begin to add observations that are less thoroughly revised and so less approximated by our notion of final data. Note, though, that our data sample will continue to be revised with the BEA's comprehensive benchmark revisions. We take the view though that further revisions are likely to have only minor effects on our estimated parameters and filtered states. We chose 1984Q1 as the starting point for the estimation to avoid the influence of potential regime shifts in monetary policy surrounding the Volcker disinflation episode.

The structural model parameters are estimated using a vector of core variables y_t that comprises seven series: the growth rates of output, consumption, investment, and nominal wages, in addition to the levels of hours worked, inflation, and the nominal interest rate. Where appropriate, the variables enter in per-capita terms. The data series are obtained from Haver Analytics (Haver mnemonics are in italics). Real output is computed as nominal GDP (*GDP*) divided by the population that is 16 years and older (*LN16N*) and the chained-price GDP deflator (*JGDP*). Consumption is defined as nominal personal consumption expenditures (*C*) less consumption of durables (*CD*). The nominal series is divided by *LN16N* and then deflated by *JGDP*. Investment is defined as *CD* plus nominal gross private domestic investment (*I*). As with the other real series, it is deflated by population and the

output price deflator. Quarter-to-quarter growth rates are computed as the log difference of real per-capita variables and are multiplied by 100 to give percentages.

The hours worked series is calculated using nonfarm business sector hours of all persons (*LXNFH*), dividing it by *LN16N*, and then scaling it to get mean quarterly average hours to about 257. The series is then logged and multiplied by 100 so that all figures can be interpreted as percentage deviations from the mean. Nominal wages are the total compensation of employees (*YCOMP*) divided by the product of *LN16N* and the measure of average hours. Inflation is computed as the log difference of the GDP deflator converted into percentages. The nominal interest rate is defined as the average effective federal funds rate (*FFED*) over the quarter and is annualized.

The model is solved using Sims' Gensys method and estimated using the code from Schorfheide et al. (2010), available on Frank Schorfheide's website.[1]

3.2 DSGE Model Estimates

Table 1 reports information on the prior of the DSGE model parameters and the posterior of the parameter estimates. The choice of prior is the same as in Schorfheide et al. (2010) and follows the "standard" prior in Del Negro and Schorfheide (2008). For details, the reader is referred to Schorfheide et al. (2010).

Table 1 also gives means and 90 percent confidence intervals for the posterior distribution of parameter estimates. The estimate of the average technology growth rate implies that long-run per-capita output, consumption, and investment grow about 1.5 percent per year. The estimates of β and π_* imply that the long-run short-term nominal interest rate is 3.4 percent. Steady state nominal wage growth is estimated at about 4 percent per year.

The estimated monetary policy rule shows a strong reaction to inflation with $\hat{\psi}_1 = 2.85$ and a weak reaction to the output gap (measured as the deviation of output from its long-run growth path) at $\hat{\psi}_2 = 0.04$. Estimated price stickiness is fairly high with $\hat{\zeta}_p = 0.85$, implying an average time between price changes of about 6.7 quarters. On the other hand, estimated wage stickiness is low with $\hat{\zeta}_w = 0.28$, implying an average duration of 1.4 quarters between wage changes.

The estimates for the shock processes are shown in Table 1 part 2. The estimated technology growth process shows little persistence ($\rho_z = 0.2$), while most remaining shocks show high to very high persistence. For the most part, these estimates are similar to

---

[1]http://www.ssc.upenn.edu/ schorf/programs/ijf-ssk_code.zip

those in Schorfheide et al. (2010) (which estimated the model over the period 1984Q1 to 2007Q3). Figure 1 plots the estimates of the latent variables driving the model, generated from the Kalman filter and evaluated at the posterior mean. There is significant low-frequency movement in the labor shock, government spending shock, and more recently the discount factor shock. The Great Recession episode shows up clearly in a dramatic increase in ϕ (which leads to a reduction in hours worked) and dramatic decreases in the marginal efficiency of investment shock μ and discount factor shock b.

The in-sample fit and forecasting properties of NKDSGE models similar to this one are explored in several papers (e.g., Schorfheide et al. (2010), Del Negro and Schorfheide (2013), and Del Negro et al. (2007)), so we do not undertake further evaluation of those properties here.

3.3   Auxiliary Model Estimation

We now turn to the analysis of the relationship between early release data on output growth and inflation and final release data on those variables. Loosely speaking, our strategy is to use the estimated DSGE model and final release data to uncover a set of state variables (or factors) that are then used as a set of regressors in a linear projection of early release data on the factors. This is the same methodology used in Schorfheide et al. (2010) where those authors investigated how nonmodeled variables can be forecast using a DSGE model. For us, the nonmodeled variables are initial release and first annual revision data on output growth and inflation. The first annual revision is interesting because it incorporates information from annual tax returns, Social security information, and annual manufacturing survey data, as well as other new source data. Since the DSGE model provides shock identification, we can then investigate how shocks to monetary policy, TFP, and labor supply influence early release data and revisions of that data to final release data. To make this paper self-contained, we begin by reviewing the estimation strategy laid out in Schorfheide et al. (2010).

Let x_t denote a variable that is not formally included in the DSGE model but that is of interest in that we wish to use the DSGE model to forecast or to understand the dynamics of x_t. In our case, x_t will denote an observation of early release data on either output growth or inflation, which we know will subsequently be revised to its observed final version. We view x_t as not formally modeled in the DSGE framework since the structural model is estimated using final release data, not early release data. Conceptually, though,

11

the structural model does not distinguish final release data from early release data, except through the measurement equation. Since we do not include measurement error in the measurement equations, the model could, in principle, fully explain the initial release data if its statistical properties are close enough to those of the final release data. The variable x_t will be modeled as a linear function of the state variables of the DSGE model ζ_t. The Kalman filter can be used in conjunction with the DSGE model and final release data to deliver a sequence $\zeta_{t|t}(\theta)$ that is obtained by using the posterior mean estimate $\hat{\theta}_T$ as a replacement for θ.

Let $\hat{s}_t$ denote a subset of the DSGE model state variables that are fundamental in the sense that all other state variables in the DSGE model can be derived as a linear combination of these fundamental states. We model the auxiliary variable x_t as:

$$x_t = \alpha_0 + \hat{s}'_{t|t}\alpha_1 + \xi_t, \quad \xi_t = \rho\xi_{t-1} + \eta_t, \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2). \tag{12}$$

The error term ξ_t is idiosyncratic for each auxiliary variable that is modeled. The structure of the auxiliary equation is much like a factor model with the factors derived from the core variables that are used in the measurement equation of the DSGE model. Bayesian methods are used to estimate the auxiliary regression (12).

Rewrite (12) in quasi-differenced form:

$$
\begin{aligned}
x_1 &= \alpha_0 + \hat{s}'_{1|1}\alpha_1 + \xi_1 \\
x_t &= \rho x_{t-1} + \alpha_0(1-\rho) + [\hat{s}'_{t|t} - \hat{s}'_{t-1|t-1}\rho]\alpha_1 + \eta_t, \quad t = 2,\ldots,T.
\end{aligned}
\tag{13}
$$

As in Schorfheide et al. (2010), we assume $\xi_1 \sim \mathcal{N}(0, \tau^2)$, where τ is interpreted as the prior standard deviation of the idiosyncratic error. We set τ to be in the range of 15 percent to 20 percent of the sample variance of x_t.

Since our variables x_t (early release data on output growth and inflation) are presumably closely related to their DSGE modeled counterparts (final release data on output growth and inflation), we can use information from the structural model to set the priors for the auxiliary regression parameters. As in Schorfheide et al. (2010), the prior takes the form:

$$\alpha \sim \mathcal{N}(\mu_{\alpha,0}, V_{\alpha,0}), \quad \rho \sim \mathcal{U}(-1,1), \quad \sigma_\eta \sim \mathcal{IG}(\nu, \tau), \tag{14}$$

where $\mathcal{N}(\mu, V)$ denotes a normal distribution with mean μ and covariance matrix V, $\mathcal{U}(a,b)$ is a uniform distribution on the interval (a,b), and $\mathcal{IG}(\nu, s)$ denotes the Inverse Gamma distribution with density $p_{\mathcal{IG}}(\sigma|\nu, s) \propto \sigma^{-(\nu+1)}e^{-\nu s^2/2\sigma^2}$. We will use the same τ to characterize the standard deviation of ξ_1 and the prior for σ_η.

The prior mean $\mu_{\alpha,0}$ is chosen based on the DSGE model-implied factor loadings for its counterpart. For example, take the case of x_t being initial release real GDP growth. The DSGE model estimated using final release real GDP gives a set of α's such that observed final release real GDP growth is an exact linear combination of the fundamental states (recall that the measurement equation does not include measurement error). We use the model-implied loadings on the fundamental states as our prior mean for the projection of initial release real GDP growth on the fundamental states. We undertake a similar exercise for initial release inflation. For the covariance matrix our prior is given by a diagonal matrix with elements:

$$diag(V_{\alpha,0}) = \left[\lambda_0, \frac{\lambda_1}{\omega_1}, \ldots, \frac{\lambda_1}{\omega_J} \right]. \tag{15}$$

The parameters λ_0 and λ_1 are hyperparameters that determine the degree of shrinkage for the intercept α_0 and the loadings α_1 of the state variables. The diagonal elements of $V_{\alpha,0}$ are scaled by ω_j^{-1}, $j = 1, \ldots, J$, where ω_j denotes the DSGE model's implied variance of the j'th element of $\hat{s}_{t|t}$ (evaluated at the posterior mean of θ). Draws from the posterior distribution are obtained using the Gibbs sampler described in Schorfheide et al. (2010). We set the degrees of freedom parameter ν of the inverted gamma prior for σ_n equal to 2, restrict $\lambda_0 = \lambda_1 = \lambda$ and consider three values: 1.00, 0.10, and 1e-5. The prior mean and posterior estimate essentially coincide when $\lambda =$ 1e-5, and as λ is increased, the factor loadings α are allowed to differ to an increasing extent from the prior mean.

Note that the strategy we are taking here is very different from using initial release data in the state-space representation measurement equation and adding a measurement error to the system. That strategy would not use final release data to help pin down the DSGE model parameters. Rather, the "truth" is a latent state implied by the model parameters estimated using early release data. Our strategy is instead to use both early release and final release data in the estimation and then to use the model to link the series.

3.3.1 Auxiliary Equation Estimates

To begin, Figure 2 shows plots of initial release data on inflation and output growth and compares them with their final release counterparts. For inflation, final release data appear to be much less volatile than the initial release data, especially up until 2005. For real GDP growth, the high frequency differences between the series are less evident. Indeed, the sample standard deviation for initial release real GDP growth (per capita) is about 0.52 compared with 0.62 for the final release data (deflated by the same population series). For

output growth, up until about the year 2000, there is a tendency for growth rate revisions to be positive, i.e., for growth rates computed using final release data growth rates to exceed growth rates computed using initial release data. Since 2000, it is the opposite: There is a tendency for final release growth rates to be somewhat less than initial release growth rates.

The auxiliary equation estimates for initial release GDP deflator inflation and real GDP growth are presented in Tables 2 and 3. The fundamental state variables on which the early release data are projected are consumption, investment, effective capital, nominal interest rate, real wage, TFP growth, labor shock, marginal efficiency of investment shock, discount factor shock, government shock, and price-markup shock. These states are sufficient to fully explain final release output and inflation. The second column in the table is the prior mean. Under the prior mean, the auxiliary model regressions return the observed values for final release output growth and inflation to approximation error. The third through fifth columns in the tables show the posterior mean estimates and 90 percent probability coverage intervals for the estimated auxiliary models under three assumption on λ, which governs the tightness of the prior. When λ is small, the prior is tight – as evidenced by the column five estimates, which are very close to the prior mean estimates. As λ is moved further away from zero, the role of the prior on the estimates is relaxed, and the estimated factor loading can begin to differ from the prior means. Generally though, we found that the difference in estimates and their implications between $\lambda = 1$ and $\lambda = 0.1$ were rather small. Consequently, in the impulse response analysis in what follows we will only present results for the case of $\lambda = 0.1$.

Consider first Table 2, which shows the auxiliary regression estimates for initial release GDP deflator inflation under three different settings of λ. These estimates imply that inflation revisions are predictable in response to certain shocks to the economy. Often, statements about the predictability of revisions are made unconditionally, and the finding tends to be a lack of predictability. We find that once we identify shocks, there is some evidence for the predictability of revisions conditional on our knowledge of the shocks.

In Table 2, estimated loadings that differ significantly from the prior means indicate state variables that are important in accounting for the deviation of initial release inflation from its final release version. The loadings on three factors (the nominal one-period interest rate R_t, TFP growth z_t, and the marginal efficiency of investment μ_t) are all outside of or at least very close to the bounds of the 90 percent coverage intervals. This suggests that movements in interest rates, multifactor productivity growth, and investment productivity can account, to some extent, for the deviation of initial release inflation from final release

14

inflation. In other words, these factors help to explain revisions to the inflation data. This is not the case for the other factors/shocks that are identified by the model.

Figure 3 plots the predicted values for initial release inflation against actual values for initial release inflation for each of the models $\lambda = 1$, $\lambda = 0.1$, and $\lambda = 1\text{E-}5$. There is little difference in the predicted series for $\lambda = 1$ compared with $\lambda = 0.1$, and a greater degree of difference for $\lambda = 1\text{e-}5$. The predicted series for $\lambda = 1$ and $\lambda = 0.1$ are notably smoother than the actual series but still appear to pick up a fair amount of the movement in actual initial release inflation. For the case of $\lambda = 0.1$, the estimated linear combination of state variables accounts for about 70 percent of the variance of initial release inflation.

Table 3 shows the estimation results for the auxiliary regressions of initial release real GDP on the DSGE model factors. Since output enters the model in levels, we add a lag of output to the vector of fundamental states and regress initial release real GDP growth on the DSGE state vector. The estimates indicate again that we can find some predictability of real GDP growth revisions in response to identified shocks. In particular, movements in TFP growth z_t, the preference for leisure ϕ_t, government spending g_t, and price markups $\lambda_{f,t}$ lead to significant differences between initial release data and final release data.

Figure 4 plots the predicted values of initial release real GDP growth for the three different values of λ as well as the actual initial release values of real GDP growth. Again, the models do a fairly good job of picking up the low frequency movement in initial release real output growth. For $\lambda = 1$ and $\lambda = 0.1$, the predicted series is about as volatile as the actual series, while for $\lambda = 1\text{e-}5$, which essentially replicates final release real GDP growth, the predicted series shows less volatility than the initial release series. For the case $\lambda = 0.1$, the estimated linear combination of state variables accounts for about 56 percent of the variance of initial release real GDP growth.

4   Impulse Response Analysis

With the DSGE model's state variables evidently accounting in part for differences between initial release and final release data on output growth and inflation, the structural model's identified shocks will have differential impacts on initial release and final release data. That is, structural shocks may help explain dynamics of data revisions. We investigate this further using impulse response analysis to investigate how initial release and final release data respond to the model's exogenous shocks.

The auxiliary regression models are easily amenable to impulse response analysis. We perturb shocks in (8) to trace out a sequence of states s_t. Given these state variable responses, we can use (12) to calculate not only how initial release data respond to changes in the structural errors but also how first annual revision data respond to the structural errors.

To calculate the impulse responses, we compute the evolution of the model's state vector, evaluated at the posterior mean for the structural model parameter estimates, in response to each of the model's structural shocks. We then draw from the posterior of the auxiliary equation parameter distribution. The impulse response functions from this exercise are shown in Figures 5-8. Each figure shows two impulse responses: The solid dark line is the response as calculated from the DSGE model i.e., the response in the final data. The dashed lines are the responses of initial release data in Figures 5 and 7, and annual release data in Figures 6 and 8, together with 90 percent probability coverage intervals.

Consider first the impulse responses for inflation that are shown in Figure 5 (for initial release inflation) and Figure 6 (for first annual revision data on inflation). Figure 5 shows that there are significant differences in the response of initial and final release inflation to marginal efficiency of investment shocks, monetary policy shocks, and possibly TFP shocks. In all three cases, the impulse responses for the initial estimates are closer to zero than is the case for the final revision estimates, and of these three cases, only the monetary policy shock impulse response is significantly different from zero. More generally, though, the initial release impulse responses differ significantly from zero for government spending shocks, monetary policy shocks, labor shocks, discount factor shocks, and price-markup shocks. For these shocks (with the exception of the monetary policy shock), the impulse responses of initial release inflation data is largely the same as the response of the final release inflation data. It appears that shocks to TFP, the marginal efficiency of investment, and monetary policy partly explain revisions to the data between the initial and final releases.

Presumably, as the data get revised over time, the response of earlier release data to shocks will approach the response pattern that we see in the final release data. Confirming evidence for this conjecture is shown in Figure 6, which plots the impulse responses from an auxiliary model estimated using first annual revision data (rather than initial release data) along with the impulse responses from the DSGE model and final release data for comparison. We see now that with the possible exception of the monetary policy shock impulse, the responses of the final release data lie within the probability coverage intervals of the annual revision impulse responses. As was the case in Figure 5, the annual release data are largely less responsive to identified shocks than are the final release counterparts.

16

Figures 7 and 8 show the same exercise for initial release and first annual revision release data on output growth. For output, shocks to TFP, labor supply, and possibly price markups lead to significant revisions from initial to final release. This pattern is less pronounced, but is still evident in the first annual revision data on output growth. As was the case for the inflation series, the initial release data are less responsive to economic shocks than are the final release data. For the other shocks in the model, there is not much difference in the response of initial versus final release data or first annual revision versus final data. Looking across both the inflation and output series, TFP shocks lead to significant revisions in both inflation and output growth while the other shocks in the model tend to affect revisions to either inflation or output growth but not both. Discount factor shocks appear not to affect revisions to either inflation or output growth, judging by the 90 percent coverage intervals.

4.1 Initial Data Releases as Forecasts

The impulse response functions indicate that initial release data on output growth and inflation are less responsive to identified macroeconomic shocks than the final release data. If the BEA is largely forecasting final release data with its initial release estimates, we might expect those estimates to be smoother than the final release data. Consider how the BEA calculates the initial release of quarterly GDP and inflation. According to Grimm and Weadock (2006), about 25 percent of the source data that are used to calculate the first quarterly estimate of GDP are trend-based data, which means that the estimate is typically calculated from previous data using moving averages, regressions, and BEA judgment. Another 30 percent of the source data are in the form of "monthly data and trend-based data," which typically include two months of source data but limited or no data for the third month, which must then be calculated. The remainder of the source data are in the form of "monthly or quarterly data," which includes either monthly data for all three months of a quarter or complete quarter data. This breakdown suggests that a substantial portion of the first estimate of GDP in a quarter is tantamount to a forecast of missing components that are then aggregated into the GDP estimates.

To characterize the BEA's forecast approach in constructing its initial release estimates, we ask how the estimates compare with private sector forecasts in response to identified economic shocks. Of course, the BEA has access to confidential data when it prepares its initial release data, so we might expect some differences in the responses. But to the extent that the BEA initial release incorporates trends from past observable data, the responses should be similar to professional forecasts. We investigate this issue using data from the

17

Survey of Professional Forecasters (SPF). In particular, we use the SPF median nowcast of the current quarter, taken before the release of the first GDP estimate, in the auxiliary model. We estimate the models over the same sample period and present the impulse responses of the SPF nowcasts to structural shocks in Figures 9 and 10.

Figure 9 shows the response of SPF nowcasts of real output growth (the dashed line) compared with the responses in the initial release data (the solid line), as well as the 90 percent probability coverage interval for the nowcast impulse response. The SPF responses do, in general, look quite similar to the responses from the BEA's initial release data on output growth. The SPF nowcast is more responsive to TFP shocks, less responsive to labor supply shocks, and less responsive to price markup shocks. For the most part though, the impulse responses from the SPF data and the initial release data line up fairly closely.

When we look at the inflation data, the impulse responses show a bit more divergence. Monetary policy shocks and price-markup shocks lead to quite different responses from the forecasters compared with the BEA. The SPF nowcasts for inflation show little response to price markups, while the response in the BEA data is dramatic. The SPF nowcast shows a positive response to the monetary policy shocks while the BEA data show a negative response. There is also a tendency for the SPF inflation nowcast to under-respond to labor supply shocks compared with the BEA initial release data. On balance, the SPF inflation nowcast is less responsive to the model's identified economic shocks than is the BEA initial release data. Indeed, for most of the shocks, the 90 percent probability coverage interval for the impulse responses contains zero. Overall, this exercise suggests that the information set used by the BEA to construct the price index is quite different from that of the private forecasters, and that it is not very well approximated by the historical data that forecasters condition on. However, this seems to be much more the case for inflation than it does for real output growth.

5 Discussion

The model estimation and impulse response analysis lead to four principal findings: *(i)* initial release data on inflation and output growth are explained in significant part by the model state variables derived from final release data, *(ii)* initial release data generally under-respond to identified shocks relative to final release data, *(iii)* identified shocks help explain data revisions from initial to final release, but there is not a general pattern across the series and shocks, and *(iv)* the response of professional forecaster nowcasts of current

output growth to shocks is similar to the responses of the BEA's initial release data – this is less so for the nowcasts of inflation.

These findings are broadly consistent with the view that the BEA's initial release data are forecasts of the final release data. The initial release data are much less responsive to TFP shocks than the final release data for both inflation and output growth. This suggests that it is particularly difficult to recognize productivity shocks in real time and to incorporate them into initial estimates of output and inflation. For the other shocks in the model the story is more nuanced. The initial release estimates of inflation show little response to investment shocks, while on the output side, the initial and final release estimates track closely. Similarly, the initial release estimate of inflation under-responds significantly to monetary policy shocks (relative to the final release response) while on the output side the responses of initial release and final release data are again quite similar. For price-markup shocks, the initial release data over-respond on the output side, while on the inflation side initial release and final release track closely.

Consider again Figure 7, which shows the impulse responses of initial release real GDP growth compared with final release. Recall that the structural model state variables account for about 60 percent of the sample variance of initial output growth. With the exception of TFP shocks and possibly price-markup shocks, the response of initial release real GDP growth to the identified shocks is quite close to the response in the final release data. More generally, though, how important are the various identified shocks in accounting for the movements in initial release data? We get a sense of this by calculating the variance contributions of the shocks using the estimated model. Table 4 shows the contribution of each shock to the variance of initial and final release output and inflation. The table also reports the 90 percent probability coverage interval for the estimates (which incorporates both parameter and shock uncertainty). For the initial release data, the most important shock is the labor supply shock, which accounts for about 40 percent of the variance of initial release output growth. TFP shocks, investment shocks, and government spending shocks each account for about 15 percent of output growth variance. Price-markup shocks account for only about 3 percent of the variance. When contrasting these results with the decomposition of final release output growth, the shock contributions are similar, with the exception of TFP growth and labor shocks. For final release output growth, TFP shocks are most important, accounting for about 37 percent of the variance of output, followed by labor supply shocks, which account for about 20 percent of the variance. This suggests that it is difficult to untangle labor supply shocks and TFP shocks in the early release data. The

remaining shock contributions match up fairly closely between the initial release and final release data.

For inflation, recall that the model state variables account for about 70 percent of the sample variance of the initial release data. With the exception of TFP shocks, investment shocks, and monetary policy shocks, the response of initial release data and final release data are close to each other. The model's variance decomposition for inflation indicates that labor supply shocks and price-markup shocks are the most important contributors, at about 35 percent each. The results are similar in the final release data with labor supply shocks and price-markup shocks contributing about 30 percent each to the variance of inflation. The initial release and the final release contributions are quite similar across the columns, suggesting that underlying shocks to the economy are reasonably well accounted for in the initial release data. Unlike the case of output growth, the contributions from TFP shocks and labor supply shocks are not confounded in the intial release and final release data.

These results suggest that a cautious approach should be taken when interpreting the initial release data on inflation and output growth in terms of how they relate to the fundamental structure of the economy. While the economic shocks that account for the final release data are influencing, to a substantial extent, the dynamics of the initial release data, there is also a substantial influence from shocks that ultimately gives rise to revisions in the data. In this paper, we have taken some first steps in quantifying these influences.

References

An, S., Schorfheide, F., 2007. Bayesian Analysis of DSGE Models. Econometric Reviews 26, 113–172.

Calvo, G., 1983. Staggered prices in a utility-maximizing framework. Journal of Monetary Economics 12, 383–398.

Christiano, L. J., Eichenbaum, M., Evans, C., 2005. Nominal rigidities and the dynamic effects of a shock to monetary policy. Journal of Political Economy 113:1, 1–45.

Croushore, D., 2011. Frontiers of real-time data analysis. Journal of Economic Literature 49, 72–100.

Del Negro, M., Schorfheide, F., 2008. Forming priors for dsge models (and how it affects the assessment of nominal rigidities). Journal of Monetary Economics 55 (7), 1191–1208.

Del Negro, M., Schorfheide, F., 2013. Chapter 2 - {DSGE} model-based forecasting. In: Elliott, G., Timmermann, A. (Eds.), Handbook of Economic Forecasting. Vol. 2, Part A of Handbook of Economic Forecasting. Elsevier, pp. 57 – 140.
URL http://www.sciencedirect.com/science/article/pii/B9780444536839000025

Del Negro, M., Schorfheide, F., Smets, F., Wouters, R., 2007. On the fit of new Keynesian models. Journal of Business and Economic Statistics 25, 123–162.

Grimm, B. T., Weadock, T. L., February 2006. Gross Domestic Product: Revisions and Source Data. Survey of Current Business, 11–15.

Kishor, N. K., Koenig, E. F., 2012. VAR estimation and forecasting when data are subject to revision. Journal of Business and Economic Statistics 30 (2), 181–190.

Schorfheide, F., Sill, K., Kryshko, M., 2010. DSGE Model-based forecasting of non-modelled variables. International Journal of Forecasting 26, 348–373.

Sims, C., 2002. Solving rational expectations models. Computational Economics 20, 1–20.

Smets, F., Wouters, R., 2003. An estimated dynamic stochastic general equilibrium of the euro area. Journal of the European Economic Association 1:5, 1123–1175.

Table 1: Prior and Posterior of DSGE Model Parameters (Part 1)

		Prior		Posterior	
Name	Density	Para (1)	Para (2)	Mean	90% Intv.
Household					
h	$\mathcal{B}$	0.700	0.050	0.728	[0.679 , 0.779]
a''	$\mathcal{G}$	0.200	0.100	0.307	[0.134 , 0.476]
ν_l	$\mathcal{G}$	2.000	0.750	1.738	[0.976 , 2.489]
ζ_w	$\mathcal{B}$	0.600	0.200	0.284	[0.216 , 0.351]
$400(1/\beta - 1)$	$\mathcal{G}$	2.000	1.000	0.870	[0.318 , 1.394]
Firms					
α	$\mathcal{B}$	0.330	0.100	0.165	[0.128 , 0.199]
ζ_p	$\mathcal{B}$	0.600	0.200	0.850	[0.812 , 0.888]
S''	$\mathcal{G}$	4.000	1.500	4.884	[2.765 , 6.890]
λ_f	$\mathcal{G}$	0.150	0.100	0.155	[0.014 , 0.291]
Monetary Policy					
$400\pi_*$	$\mathcal{N}$	3.000	1.500	2.501	[1.808 , 3.229]
ψ_1	$\mathcal{G}$	1.500	0.400	2.854	[2.335 , 3.339]
ψ_2	$\mathcal{G}$	0.200	0.100	0.041	[0.022 , 0.061]
ρ_R	$\mathcal{B}$	0.500	0.200	0.853	[0.823 , 0.883]

Table 1: Prior and Posterior of DSGE Model Parameters (Part 2)

		Prior		Posterior	
Name	Density	Para (1)	Para (2)	Mean	90% Intv.
			Shocks		
400γ	$\mathcal{G}$	2.000	1.000	1.737	[1.285 , 2.145]
g_*	$\mathcal{G}$	0.300	0.100	0.292	[0.130 , 0.438]
ρ_z	$\mathcal{B}$	0.200	0.100	0.204	[0.096 , 0.310]
ρ_μ	$\mathcal{B}$	0.800	0.050	0.817	[0.761 , 0.868]
ρ_{λ_f}	$\mathcal{B}$	0.600	0.200	0.325	[0.208 , 0.440]
ρ_g	$\mathcal{B}$	0.800	0.050	0.959	[0.944 , 0.975]
ρ_b	$\mathcal{B}$	0.600	0.200	0.832	[0.754 , 0.910]
ρ_ϕ	$\mathcal{B}$	0.600	0.200	0.978	[0.963 , 0.990]
σ_z	$\mathcal{IG}$	0.750	2.000	0.623	[0.549 , 0.694]
σ_μ	$\mathcal{IG}$	0.750	2.000	0.428	[0.349 , 0.499]
σ_{λ_f}	$\mathcal{IG}$	0.750	2.000	0.175	[0.148 , 0.203]
σ_g	$\mathcal{IG}$	0.750	2.000	0.357	[0.317 , 0.400]
σ_b	$\mathcal{IG}$	0.750	2.000	0.381	[0.296 , 0.467]
σ_ϕ	$\mathcal{IG}$	4.000	2.000	2.761	[2.019 , 3.526]
σ_R	$\mathcal{IG}$	0.200	2.000	0.140	[0.122 , 0.156]

Notes: Para (1) and Para (2) list the means and the standard deviations for the Beta ($\mathcal{B}$), Gamma ($\mathcal{G}$), and Normal ($\mathcal{N}$) distributions; s and ν for the Inverse Gamma ($\mathcal{IG}$) distribution, where $p_{\mathcal{IG}}(\sigma|\nu,s) \propto \sigma^{-(\nu+1)}e^{-\nu s^2/2\sigma^2}$. The joint prior distribution is obtained as a product of the marginal distributions tabulated in the table and truncating this product at the boundary of the determinacy region. Posterior summary statistics are computed based on the output of the posterior sampler. The following parameters are fixed: $\delta = 0.025$, $\lambda_w = 0.3$; estimation sample: 1984Q1 to 2010Q4.

Table 2: PRIOR AND POSTERIORS OF AUXILIARY INFLATION REGRESSIONS

Factor	Prior Mean	Posterior		
		$\lambda = 1$	$\lambda = 0.1$	$\lambda = 1e^{-5}$
α_0	0.648	0.669	0.668	0.648
		[0.608, 0.738]	[0.605,0.728]	[0.642,0.653]
c	0.029	0.048	0.043	0.029
		[0.001, 0.093]	[0.009,0.081]	[0.028,0.030]
i	0.010	0.011	0.011	0.010
		[0.002, 0.021]	[0.003,0.019]	[0.010,0.010]
$\bar{K}$	-0.023	-0.024	-0.023	-0.023
		[-0.050, 0.002]	[-0.045,-0.000]	[-0.024,-0.023]
R	-0.290	-0.134	-0.126	-0.287
		[-0.297, 0.031]	[-0.270,0.025]	[-0.297,-0.277]
w	0.052	0.030	0.033	0.052
		[-0.016, 0.075]	[-0.007,0.073]	[0.050,0.054]
z	-0.009	0.044	0.044	-0.008
		[-0.012, 0.102]	[-0.012,0.098]	[-0.016,0.001]
ϕ	0.029	0.035	0.033	0.029
		[0.013, 0.058]	[0.014,0.050]	[0.028,0.029]
μ	0.107	0.003	0.003	0.107
		[-0.103, 0.106]	[-0.098,0.101]	[0.100,0.113]
b	0.202	0.188	0.187	0.204
		[0.060, 0.309]	[0.065,0.300]	[0.196,0.211]
g	0.039	0.069	0.064	0.040
		[0.016, 0.127]	[0.016,0.114]	[0.036,0.044]
λ_f	1.600	1.437	1.435	1.597
		[1.222, 1.659]	[1.209,1.636]	[1.568,1.624]
ρ		-0.164	-0.160	0.001
		[-0.344, 0.011]	[-0.330,0.023]	[-0.167,0.167]
σ_η	0.114	0.029	0.029	0.033
		[0.022, 0.036]	[0.022,0.036]	[0.026,0.041]

Note: The 90 percent probability coverage intervals are in brackets.

Table 3: PRIOR AND POSTERIORS OF AUXILIARY OUTPUT REGRESSIONS

Factor	Prior Mean	Posterior		
		$\lambda = 1$	$\lambda = 0.1$	$\lambda = 1e^{-5}$
α_0	0.441	0.205	0.193	0.441
		[0.090, 0.323]	[0.093,0.301]	[0.436,0.446]
c	0.842	0.759	0.849	0.842
		[0.605, 0.905]	[0.791,0.909]	[0.841,0.843]
i	0.179	0.151	0.170	0.179
		[0.121, 0.181]	[0.157,0.184]	[0.179,0.180]
$\bar{K}$	-0.021	-0.017	-0.014	-0.021
		[-0.063, 0.030]	[-0.047,0.019]	[-0.022,-0.021]
R	0.000	-0.044	0.093	-0.002
		[-0.353, 0.255]	[-0.148,0.343]	[-0.012,0.008]
w	0.017	-0.024	0.010	0.017
		[-0.105, 0.054]	[-0.052,0.072]	[0.016,0.019]
z	0.000	-0.423	-0.328	-0.002
		[-0.564, -0.291]	[-0.429,-0.230]	[-0.009,0.007]
ϕ	0.000	0.032	0.025	0.000
		[-0.006, 0.067]	[0.001,0.048]	[-0.000,0.000]
μ	0.006	0.096	-0.003	0.005
		[-0.108, 0.297]	[-0.176,0.165]	[-0.002,0.012]
b	-0.000	0.078	0.013	-0.001
		[-0.147, 0.310]	[-0.189,0.215]	[-0.008,0.006]
g	1.021	0.771	0.874	1.020
		[0.606, 0.944]	[0.784,0.959]	[1.016,1.024]
λ_f	0.000	-0.418	-0.416	0.001
		[-0.801, -0.026]	[-0.801,-0.054]	[-0.027,0.029]
y_{lag}	-1.000	-0.848	-0.968	-1.000
		[-0.991, -0.705]	[-1.026,-0.911]	[-1.001,-0.999]
ρ		-0.086	-0.100	-0.004
		[-0.270, 0.090]	[-0.281,0.070]	[-0.167,0.154]
σ_η	0.303	0.086	0.088	0.157
		[0.065, 0.107]	[0.067,0.109]	[0.122,0.191]

Note: The 90 percent probability coverage intervals are in brackets.

Table 4: VARIANCE CONTRIBUTION OF SHOCKS

Shocks	Output Growth		Inflation	
	Initial	Final	Initial	Final
z	0.372	0.159	0.010	0.012
	[0.295, 0.453]	[0.104,0.216]	[0.001,0.020]	[0.003,0.020]
ϕ	0.192	0.391	0.288	0.320
	[0.134, 0.249]	[0.316,0.466]	[0.050,0.511]	[0.149,0.485]
μ	0.114	0.151	0.056	0.083
	[0.058, 0.169]	[0.093,0.204]	[0.003,0.112]	[0.031,0.134]
b	0.075	0.105	0.171	0.161
	[0.042, 0.108]	[0.062,0.145]	[0.032,0.300]	[0.087,0.232]
g	0.114	0.145	0.017	0.003
	[0.085, 0.143]	[0.107,0.181]	[0.000,0.041]	[0.000,0.005]
λ_f	0.037	0.029	0.281	0.372
	[0.023, 0.052]	[0.014,0.043]	[0.135,0.424]	[0.254,0.486]
ϵ_R	0.011	0.021	0.017	0.049
	[0.005, 0.017]	[0.010,0.032]	[0.003,0.031]	[0.021,0.076]

Note: The 90 percent probability coverage intervals are in brackets and are computed on a shock-by-shock basis. Hence, column entries need not sum to 100 percent.

Figure 1: ESTIMATED LATENT VARIABLES
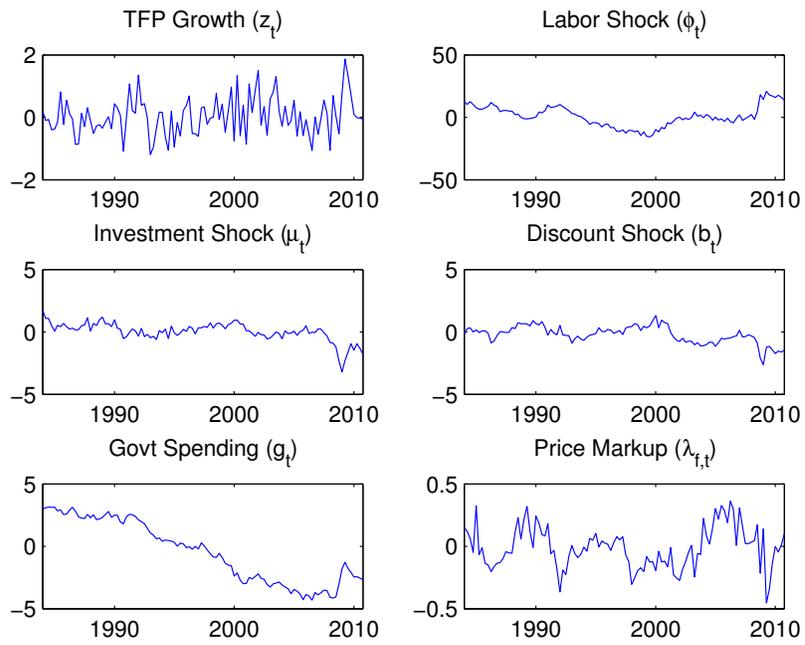
Figure 2: FINAL AND INTIAL RELEASE DATA

Final and Initial Release Inflation

Final and Initial Release Real GDP Growth
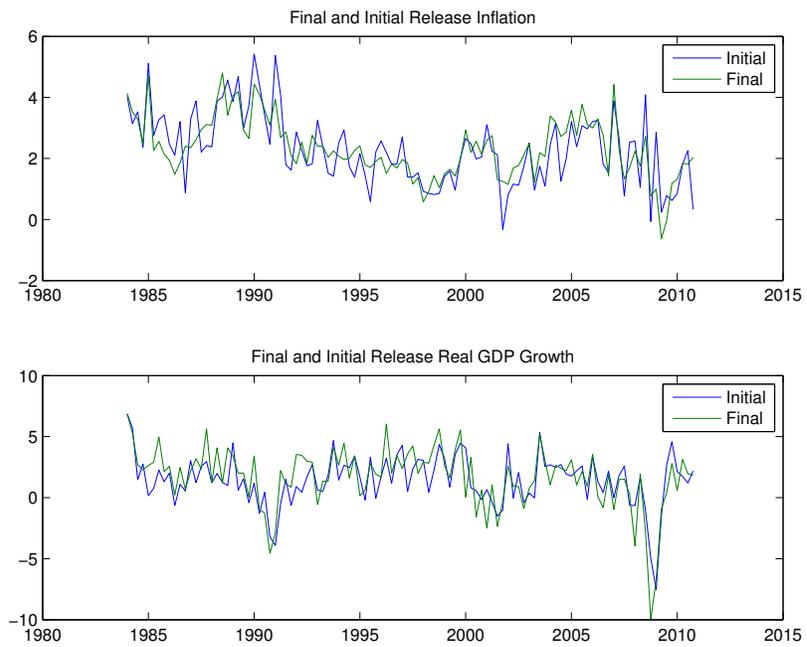
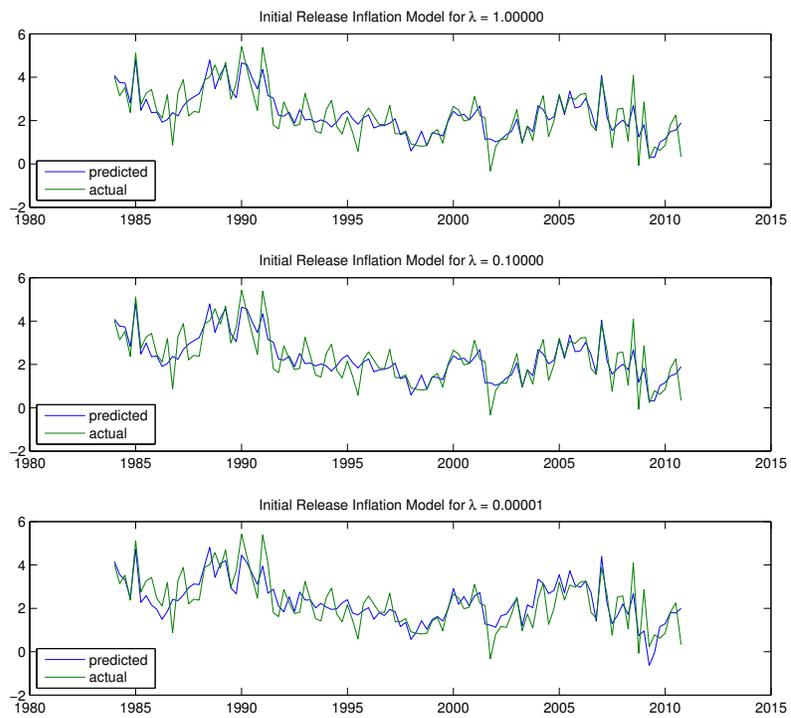Figure 3: PREDICTED VS. ACTUAL FOR INITIAL RELEASE INFLATION

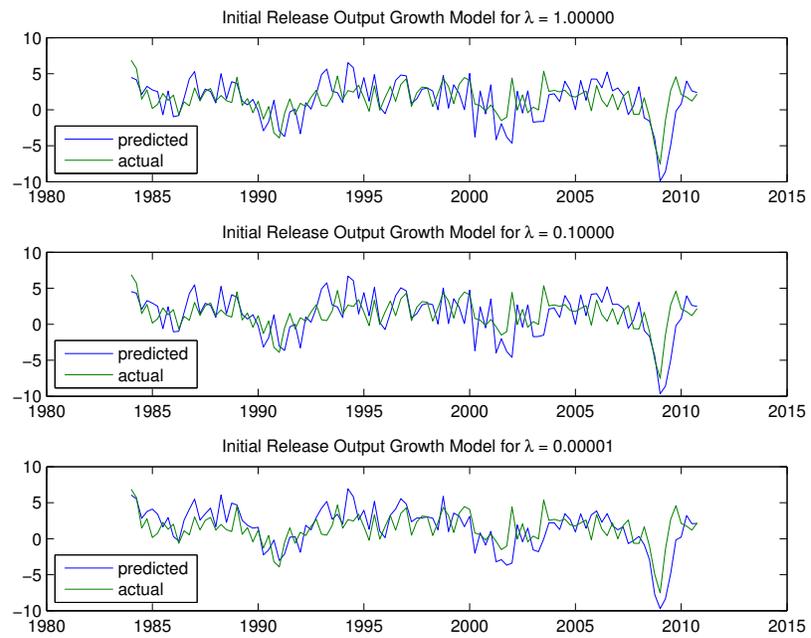Figure 4: PREDICTED VS. ACTUAL FOR INITIAL RELEASE OUTPUT GROWTH

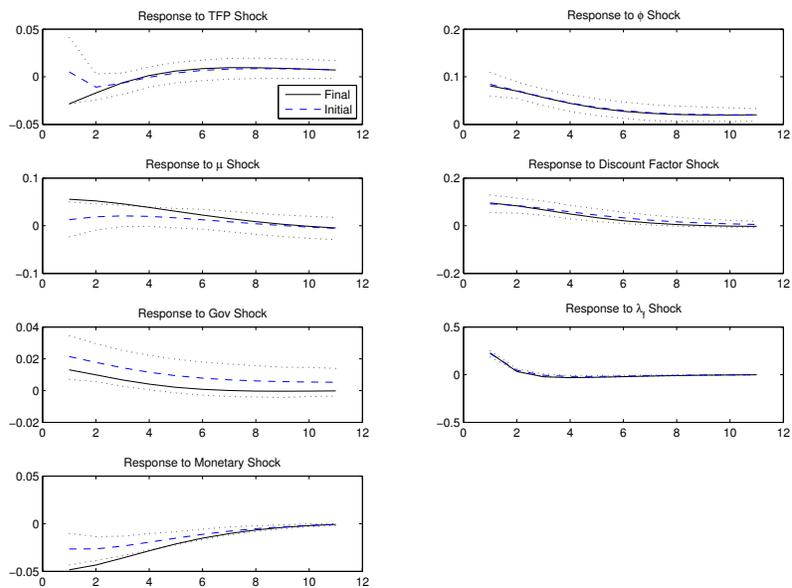Figure 5: INFLATION IMPULSE RESPONSES FOR INITIAL RELEASE VS. FINAL

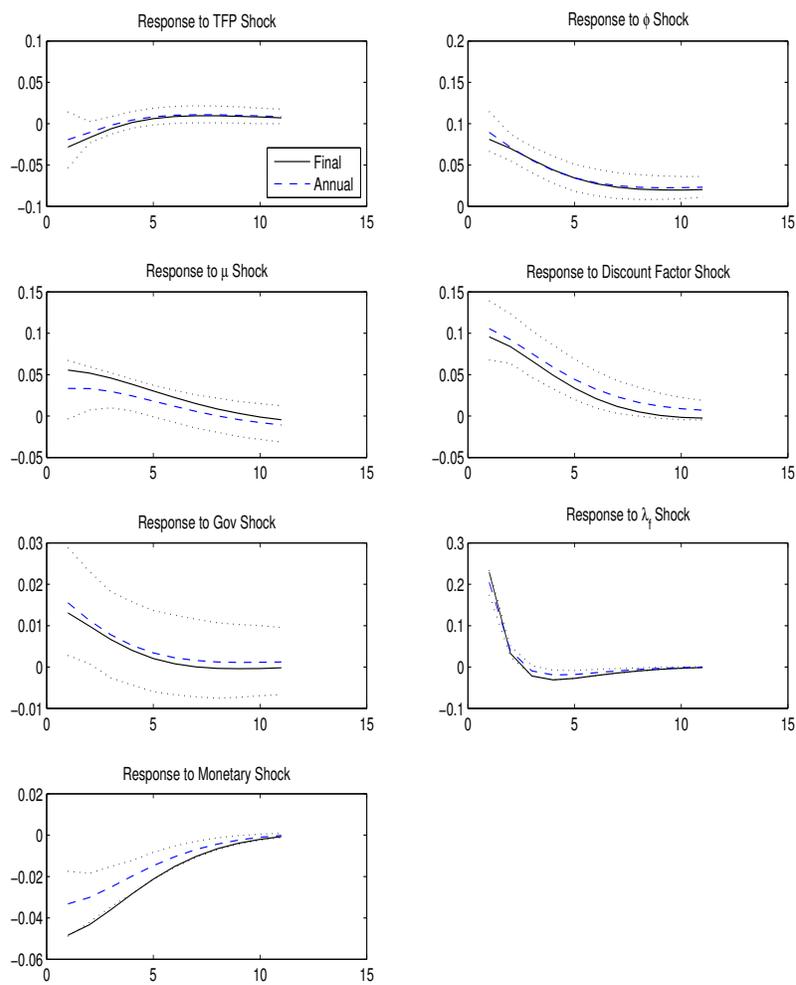Figure 6: Inflation Impulse Responses for Annual Revision vs. Final
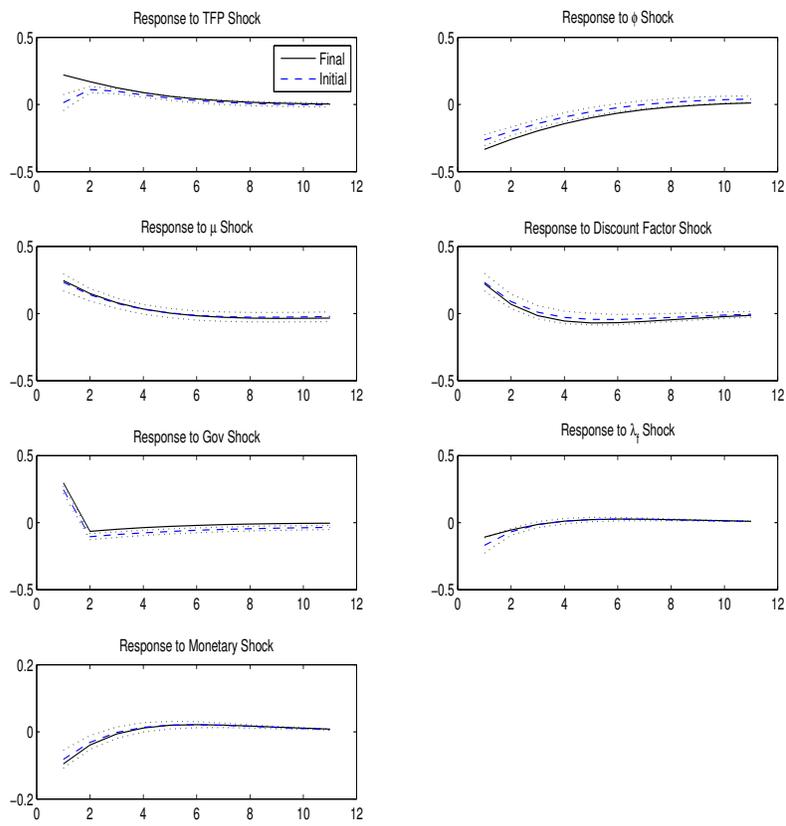
Figure 7: OUTPUT GROWTH IMPULSE RESPONSES FOR INITIAL RELEASE VS. FINAL

Figure 8: OUTPUT GROWTH IMPULSE RESPONSES FOR ANNUAL REVISION VS. FINAL
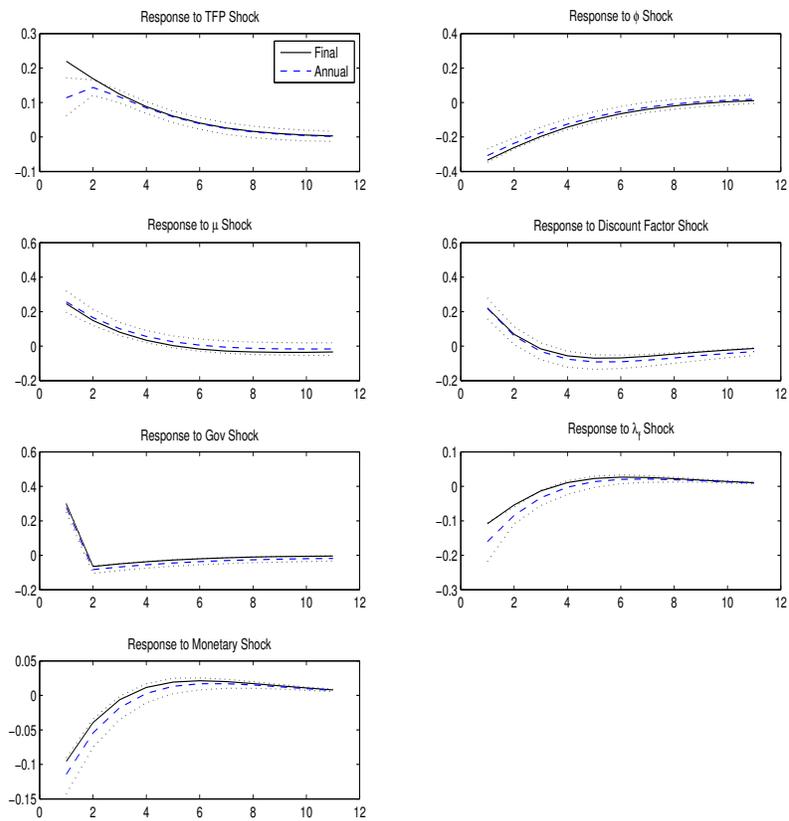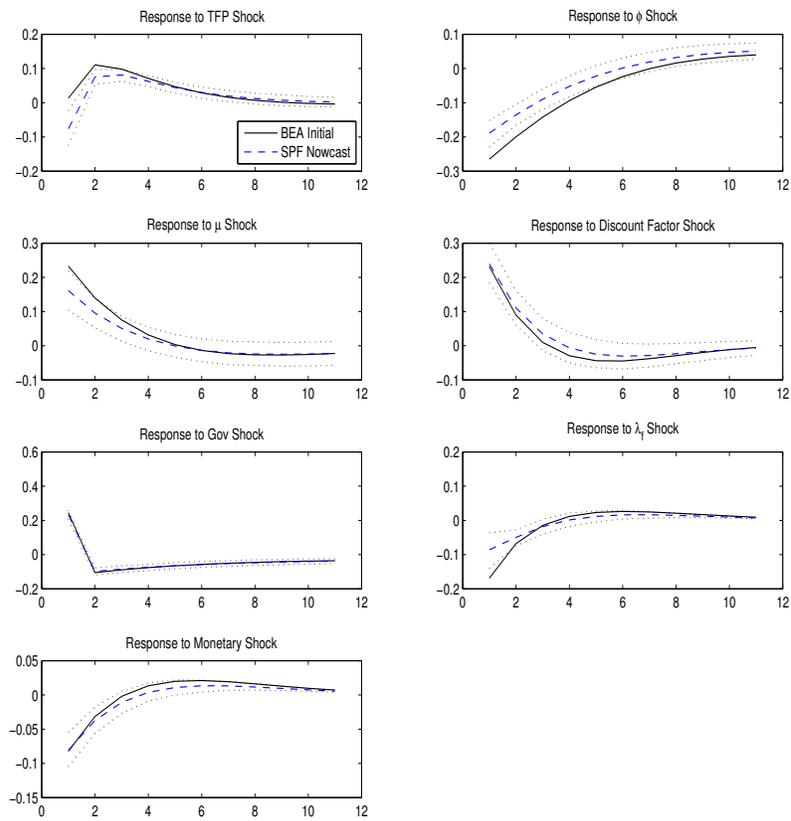
Figure 9: SPF Output Growth Impulse Responses vs. Initial

Figure 10: SPF Inflation Impulse Responses vs. Initial