ESSAYS IN ESTIMATION OF DYNAMIC STOCHASTIC GENERAL EQUILIBRIUM MODELS

MAXYM KRYSHKO

A DISSERTATION

in

ECONOMICS

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2010

Frank Schorfheide, Professor of Economics
Supervisor of Dissertation

Dirk Krueger, Professor of Economics
Graduate Group Chairperson

Dissertation Committee:

Frank Schorfheide, Professor of Economics

Francis X. Diebold, Professor of Economics

Jesús Fernández-Villaverde, Professor of Economics

ESSAYS IN ESTIMATION OF DYNAMIC STOCHASTIC GENERAL EQUILIBRIUM MODELS

Acknowledgements

The author is deeply grateful to his main advisor Frank Schorfheide, and the thesis committee members Frank Diebold and Jesús Fernández-Villaverde for the continued support, strong encouragement and wise guidance throughout the process of writing this dissertation. The author would also like to thank Cristina Fuentes-Albero, Yuriy Gorodnichenko, Ed Herbst, Dirk Krueger, Leonardo Melosi, Emanuel Moench, Andriy Norets, Keith Sill, Kevin Song, Sergiy Stetsenko and other participants of the Penn Econometrics Seminar, Penn Macro lunch and Penn Econometrics lunch for valuable discussions and many useful comments and suggestions.

ABSTRACT

ESSAYS IN ESTIMATION OF DYNAMIC STOCHASTIC GENERAL EQUILIBRIUM MODELS

Maxym Kryshko

Frank Schorfheide

Dynamic factor models (DFM) and dynamic stochastic general equilibrium (DSGE) models are widely used for empirical research in macroeconomics. The empirical factor literature argues that the co-movement of large panels of macroeconomic and financial data can be captured by relatively few common unobserved factors. Similarly, the dynamics in DSGE models are often governed by a handful of state variables and exogenous processes such as latent preference and/or technology shocks. A general topic of this dissertation is the estimation of DSGE models on a rich panel of macroeconomic and financial data by combining a DSGE with a dynamic factor model. By incorporating richer information, this combination allows to obtain DSGE model predictions and to do more reliable policy analysis with a broader range of data series of interest than before. Moreover, the combination of a DSGE and a dynamic factor model can be used as a tool for evaluating a DSGE model. This dissertation consists of three essays summarized below.

Chapter 1 *"Bayesian Dynamic Factor Analysis of a Simple Monetary DSGE Model"*: We take a standard New Keynesian business cycle model to a richer data set. When estimating DSGE models, the number of observable economic variables is usually kept small, and for convenience it is assumed that the model variables are perfectly measured by a single – often quite arbitrarily selected – data series. We relax these two assumptions and estimate a fairly simple monetary DSGE model on a richer data set. Building upon Boivin and Giannoni (2006), the framework can be seen as a combination of a DSGE model and a dynamic factor model in which factors are economic state variables and the factor dynamics are governed by a DSGE model solution. Using post-1983 U.S. data on real output, inflation, nominal interest rates, measures of inverse money velocity, and a large panel of informational series, we compare the data-rich DSGE model with a regular – few observables, perfect measurement – DSGE model in terms of deep parameter estimates, propagation of monetary policy and technology shocks and sources of business cycle fluctuations. We document that the data-rich DSGE model generates a higher implied duration of Calvo price contracts and a lower slope of the New Keynesian Phillips curve. Because of the data set's high panel dimension, the likelihood-based estimation of the data-rich DSGE model is computationally very

challenging. To reduce the costs, we employed a novel speedup as in Jungbacker and Koopman (2008) and achieved the computational time savings of 60 percent.

Chapter 2 *"Data-Rich DSGE and Dynamic Factor Models"*: In addition to a data-rich DSGE model with a standard New Keynesian core, we consider an unrestricted dynamic factor model and estimate both on a rich panel of U.S. macroeconomic and financial data compiled by Stock and Watson (2008). We find that the spaces spanned by the common empirical factors and by the data-rich DSGE model states are very close. First, this implies that a DSGE model indeed captures the essential sources of co-movement in the data and that the differences in fit between a data-rich DSGE model and a DFM are potentially due to restricted factor loadings in the former. Second, this also implies a greater degree of comfort about propagation of structural shocks to a wide array of macro and financial series. Third, the proximity of factor spaces facilitates economic interpretation of a dynamic factor model, as the empirical factors are now isomorphic to the DSGE model state variables with clear economic meaning. Finally, the proximity of factor spaces allows us to propagate monetary policy and technology innovations in an otherwise completely non-structural dynamic factor model to obtain predictions for many more series than just a handful of traditional macro variables including measures of real activity, price indices, labor market indicators, interest rate spreads, money and credit stocks, and exchange rates. We can therefore provide a more complete and comprehensive picture of the effects of monetary policy and technology shocks.

Chapter 3 *"DSGE Model Based Forecasting of Non-Modeled Variables"* (joint work with Frank Schorfheide and Keith Sill): We develop and illustrate a simple method to generate a DSGE model-based forecast for variables that do not explicitly appear in the model (non-core variables). Estimation is performed in two steps. First, we estimate the regular DSGE model on core observables. Second, we obtain filtered DSGE model state variables and use them as regressors in auxiliary linear regressions – resembling DFM measurement equations – for the non-core variables. Predictions for the non-core variables are then obtained by applying their estimated measurement equations to DSGE model-generated forecasts of the state variables.

This estimation approach can be viewed as a simplified version of a data-rich DSGE model estimation in which we essentially decouple the analysis of the non-core measurement equations and the estimation of a DSGE model on the core observables. The proposed shortcut is practically appealing: we considerably reduce the associated computational costs and we can incorporate and forecast an additional non-core variable without having to re-estimate the whole DSGE model, a feature useful in real-time applications. We apply our approach to generate and evaluate recursive forecasts for personal consumption expenditure (PCE) inflation, core PCE inflation, the unemployment rate, and housing starts.

Table of Contents

ACKNOWLEDGEMENTS .... **III**

LIST OF TABLES .... **VIII**

LIST OF FIGURES .... **IX**

CHAPTER 1. BAYESIAN DYNAMIC FACTOR ANALYSIS OF A SIMPLE MONETARY DSGE MODEL **1**

1 INTRODUCTION 1
2 DATA-RICH DSGE MODEL 3
2.1 Regular vs. Data-Rich DSGE Models *5*
2.2 Environment *7*
2.2.1 Households 8
2.2.2 Final Good Firms 11
2.2.3 Intermediate Goods Firms 12
2.2.4 Monetary and Fiscal Policy 16
2.2.5 Aggregation 17
3 ECONOMETRIC METHODOLOGY 19
3.1 Estimation of the Data-Rich DSGE Model *19*
3.2 Speed-Up: Jungbacker and Koopman 2008 .... *26*
4 DATA AND TRANSFORMATIONS 29
5 EMPIRICAL RESULTS 32
5.1 Priors *32*
5.2 Posteriors: Regular vs. Data-Rich DSGE Model *37*
5.3 Estimated States: Regular vs. Data-Rich DSGE Model *39*
5.4 Sources of Business Cycle Fluctuations *41*
5.5 Impulse Response Analysis *44*
6 CONCLUSIONS 49
APPENDIX A. DSGE MODEL 52
Appendix A1. First-Order Conditions of Household *52*
Appendix A2. First-Order Conditions of Intermediate Goods Firm *54*
Appendix A3. Evolution of Price Dispersion *58*
Appendix A4. Equilibrium Conditions and Aggregate Disturbances *58*
Appendix A5. Steady State and Log-Linearized Equilibrium Conditions *61*
APPENDIX B. DETAILS OF MARKOV CHAIN MONTE CARLO ALGORITHM 65
APPENDIX C. DATA: DESCRIPTION AND TRANSFORMATIONS 72
APPENDIX D. TABLES AND FIGURES 74

CHAPTER 2. DATA-RICH DSGE AND DYNAMIC FACTOR MODELS **83**

1 INTRODUCTION 83
2 TWO MODELS 88
2.1 Dynamic Factor Model *88*
2.2 Data-Rich DSGE Model *89*
3 ECONOMETRIC METHODOLOGY 91
3.1 Estimation of the Data-Rich DSGE Model *91*
3.2 Estimation of the Dynamic Factor Model *91*
4 DATA 98
5 EMPIRICAL ANALYSIS 99
5.1 Priors and Posteriors *100*

5.2 *Empirical Factors and Estimated DSGE Model States* ... *102*
5.3 *How Well Factors Trace Data* ... *103*
5.4 *Comparing Factor Spaces* ... *105*
5.5 *Propagation of Monetary Policy and Technology Innovations* ... *106*
6 CONCLUSIONS ... 115
APPENDIX E. DFM: GIBBS SAMPLER: DRAWING TRANSITION EQUATION MATRIX ... 116
APPENDIX F. TABLES AND FIGURES ... 118

CHAPTER 3. DSGE MODEL BASED FORECASTING OF NON-MODELED VARIABLES ... 132

1 INTRODUCTION ... 132
2 THE DSGE MODEL ... 136
3 ECONOMETRIC METHODOLOGY ... 142
3.1 *DSGE Model Estimation* ... *143*
3.2 *Linking Model States to Non-Core Variables* ... *145*
3.3 *Forecasting* ... *149*
4 EMPIRICAL APPLICATION ... 150
4.1 *Data and Priors* ... *151*
4.2 *DSGE Model Estimaton and Forecasting of Core Variables* ... *155*
4.3 *Forecasting Non-Core Variables with Auxiliary Regressions* ... *161*
4.4 *Multivariate Considerations* ... *171*
5 CONCLUSION ... 182
APPENDIX. MCMC IMPLEMENTATION ... 184

BIBLIOGRAPHY ... 189

List of Tables

Table D1. Data-Rich DSGE Model: Parameters Fixed During Estimation - Calibration and Normalization 74
Table D2. Data-Rich DSGE Model: Prior Distributions 75
Table D3. Data-Rich DSGE Model: Posterior Estimates 76
Table D4. Data-Rich DSGE Model: Summary of the Unconditional Variance Decomposition 77
Table D5. Data-Rich DSGE vs. Regular DSGE Model: Unconditional Variance Decomposition 78
Table F1. DFM: Principal Components Analysis 119
Table F2. Pure DFM: Fraction of Unconditional Variance Captured by Factors 120
Table F3. Data-Rich DSGE Model: Fraction of Unconditional Variance Captured by DSGE Model States 120
Table F4. Pure DFM: Unconditional Variance Captured by Factors 121
Table F5. Data-Rich DSGE Model: Fraction of Unconditional Variance Captured by DSGE Model States 123
Table F6. Regressing Data-Rich DSGE Model States on DFM Factors 125
Table F7. Regressing DFM Factors on Data-Rich DSGE Model States 125
Table 1. Prior and Posterior of DSGE Model Parameters: Part 1 153
Table 2. Prior and Posterior of DSGE Model Parameters: Part 2 154
Table 3. RMSE Comparison: DSGE Model versus AR(1) 159
Table 4. One-Step-Ahead Forecast Performance of DSGE Models 160
Table 5. Non-Modelled and Related DSGE Model Variables 162
Table 6. Auxiliary Regression Estimates 165
Table 7. Root Mean Squared Errors for Auxiliary Regressions 169
Table 8. Posterior Predictive Check: Cross-Correlations 172
Table 9. RMSE Ratios: Conditional (on Interest Rates) vs. Unconditional Forecasts ... 181
Table 10. RMSE Ratios: Conditional (on GDP Deflator Inflation) vs. Unconditional Forecasts 181

List of Figures

Figure D1. Data-Rich DSGE Model (iid errors): Estimated Model States....................... 79
Figure D2. Impulse Responses to Structural Shocks: Primary Observables 80
Figure D3. Impact of Monetary Policy Innovation on Core Macro Series: Regular vs. Data-Rich DSGE Model.. 81
Figure D4. Impact of Technology Innovation on Core Macro Series: Regular vs. Data-Rich DSGE Model 82
Figure F1. DFM: Principal Components Analysis 118
Figure F2. Pure DFM (iid errors): Estimated Factors.. 126
Figure F3. Do Empirical Factors and DSGE Model State Variables Span the Same Space? 127
Figure F4. Impact of Monetary Policy Innovation on Core Macro Series 128
Figure F5. Impact of Monetary Policy Innovation on Non-Core Macro Series 129
Figure F6. Impact of Technology Innovation on Core Macro Series............................ 130
Figure F7. Impact of Technology Innovation on Non-Core Macro Series.................... 131
Figure 1. Latent State Variables of the DSGE Model 157
Figure 2. Non-Core Variables and Related DSGE Model Variables............................. 163
Figure 3. Non-Core Variables and Factors 167
Figure 4. Impulse Response to a Contractionary Monetary Policy Shock 174
Figure 5. Bivariate One-Step-Ahead Predictive Distributions 179

CHAPTER 1. BAYESIAN DYNAMIC FACTOR ANALYSIS OF A SIMPLE MONETARY DSGE MODEL

1 Introduction

When estimating dynamic stochastic general equilibrium (DSGE) models, the number of observable economic variables is usually kept small, and for convenience it is assumed that the model variables are perfectly measured by a single – often quite arbitrarily selected – data series. In this chapter, we relax these two assumptions and estimate a version of the monetary DSGE model with a standard New Keynesian core on a richer data set. Building upon Boivin and Giannoni (2006), this so called *data-rich DSGE model* can be seen as a combination of a regular DSGE model and a dynamic factor model in which factors are the economic state variables of the DSGE model and the transition of factors is governed by a DSGE model solution.

We use the post-1983 U.S. data on real output, inflation, nominal interest rates, measures of inverse money velocity and a large panel of the other informational macroeconomic and financial series compiled by Stock and Watson (2008) to estimate and compare the new data-rich DSGE model with a regular – few observables, perfect measurement – DSGE model, both sharing the same theoretical core. The estimation involves Bayesian Markov Chain Monte Carlo (MCMC) methods. Because of the data set's high panel dimension, the likelihood-based estimation of the data-rich DSGE model is computationally very challenging. To reduce the costs, we employed a novel speed-up as in Jungbacker and Koopman (2008) and achieved the computational time savings of 60 percent.

We document that the data-rich DSGE model generates a higher duration of the Calvo price contracts and a lower implied slope of the New Keynesian Phillips curve measuring the elasticity of current inflation to real marginal costs. As we move from the regular to the data-rich DSGE model, we find that: (i) the role of technology innovations in generating fluctuations in real output, inflation and the interest rates is noticeably reduced; and that (ii) the contribution of monetary policy shocks to cyclical fluctuations of the interest rates increases from 4 to 14-17 percent. Regarding dynamic propagation, we establish that (i) despite some slight on-impact differences, the responses of all primary observables (real GDP, GDP deflator inflation, fed funds rate and real M2) to the monetary policy innovation remain theoretically plausible and quantitatively close in the regular and in the data-rich DSGE models; and that (ii) the regular DSGE model tends to overestimate all effects of TFP shocks, though on impact they might not have been too different. Finally, we find some puzzling results for the responses of the industrial production, the PCE deflator inflation and the CPI inflation to monetary tightening, which may indicate the potential misspecification of our theoretical DSGE model.

The chapter is organized as follows. In Section 2, we present a data-rich DSGE model with a New Keynesian core to be used in the subsequent empirical analysis. Our econometric methodology to estimate the data-rich DSGE model and also the Jungbacker-Koopman computational speed-up are discussed in Section 3. Section 4 describes our data set and transformations. In Section 5 we proceed by conducting the empirical analysis of the regular and the data-rich DSGE models. We begin by discussing the choice of the prior distributions of model parameters and then describe the posterior estimates of deep structural parameters in both models. Second, we compare the estimated DSGE state variables from our data-rich and from the regular DSGE model. Finally, we explore the differences that the regular and the data-rich DSGE models imply about the sources of business cycle fluctuations and about the propagation of structural innovations, notably the monetary policy and technology shocks, to the real output, inflation, interest rates and the real money balances. Section 6 concludes.

2 Data-Rich DSGE Model

In this section, we begin by defining what we refer to as the data-rich DSGE model and contrast it with the regular DSGE model. Then, we present a fairly standard New Keynesian business cycle core that will be shared by both types of models.

In any DSGE model, economic agents solve intertemporal optimization problems built from explicit preferences and technology assumptions. Moreover, decision rules of these agents depend upon a number of exogenous stochastic disturbances that characterize uncertainty in the economic environment. The equilibrium dynamics of a DSGE model are captured by a system of non-linear expectational difference equations. The standard approach in the literature is to derive a log-linear approximation to this non-linear system around its deterministic steady state and then to solve numerically the resulting linear rational expectations system by one of the available methods.[1]

This numerical solution delivers a vector autoregressive process for S_t, the vector collecting all non-redundant state variables of the DSGE model, and a linear relationship between the remaining DSGE model variables z_t and the current state S_t:

$$z_t = \mathbf{D}(\boldsymbol{\theta})S_t \tag{1}$$

$$S_t = \mathbf{G}(\boldsymbol{\theta})S_{t-1} + \mathbf{H}(\boldsymbol{\theta})\varepsilon_t, \qquad \text{where } \varepsilon_t \sim iid\ \ N(0,\mathbf{Q}(\boldsymbol{\theta})). \tag{2}$$

The matrices in (1) and (2) are the functions of structural parameters $\boldsymbol{\theta}$ characterizing preferences and technology in a DSGE model. For convenience, we assume that the exogenous shocks ε_t are mean-zero normal random variables with diagonal covariance matrix $\mathbf{Q}(\boldsymbol{\theta})$. In what follows we will refer to S_t as the *DSGE model states* or the DSGE model state variables. We will also refer to the elements of $\bar{S}_t = [z_t', S_t']'$, the vector collecting all variables in a given DSGE model, as the DSGE model concepts or simply *model concepts*. The typical examples of model concepts could be inflation, output, technology shock, capital stock and so on. By definition of $\bar{S}_t$:

$$\bar{S}_t = \begin{bmatrix} \mathbf{D}(\boldsymbol{\theta}) \\ \mathbf{I} \end{bmatrix} S_t \tag{3}$$

In order to estimate our DSGE model on a set of observables $X^T = [X_1,...,X_T]'$, a state-space representation of the model is constructed by augmenting (1)-(2) with a

[1] Please see Sims (2002), Blanchard and Kahn (1980), Klein (2000), Uhlig (1999), and King and Watson (2002).

number of measurement equations that connect model concepts in $\bar{S}_t$ to data indicators in vector X_t.

2.1 Regular vs. Data-Rich DSGE Models

Depending on the number of data indicators and on how we connect them to the model concepts, we will distinguish regular and data-rich DSGE models. In *regular* DSGE models, the number of observables contained in X_t is usually kept small (most often equal to the number of structural shocks) and model concepts are often assumed to be perfectly measured by a single data indicator.[2] For example, Lubik and Schorfheide (2004), in a DSGE model with three structural shocks, specify the following measurement equations for real output $\tilde{x}_t$, inflation $\tilde{\pi}_t$, and the nominal interest rate $\tilde{R}_t$ (we omit the intercept for simplicity):

$$\underbrace{\begin{bmatrix} \text{RealGDP}_t \\ \text{CPI\_Inflation}_t \\ \text{FedFundsRate}_t \end{bmatrix}}_{X_t} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 4 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 4 & 0 & \cdots & 0 \end{bmatrix}}_{\mathbf{\Lambda}} \cdot \underbrace{\begin{bmatrix} \tilde{x}_t \\ \tilde{\pi}_t \\ \tilde{R}_t \\ \vdots \end{bmatrix}}_{\bar{S}_t} \tag{4}$$

Similarly, Smets and Wouters (2007) estimate a DSGE model with seven structural shocks on seven key U.S. macro variables: again assuming one-to-one model concept-data indicator correspondence and perfect measurement.

[2] The underlying reason is to avoid the so-called stochastic singularity. The likelihood function for observables X_t with dimension exceeding the number of structural shocks will be degenerate, since according to DSGE model some X_t's can be perfectly (deterministically) predicted from others and this is obviously not true in the data. The solution is to add measurement errors (or theoretical gaps between the model concept and the data indicator) as e.g. in Altug (1989), Sargent (1989), and Ireland (2004), or to add more shocks, e.g., as in Leeper and Sims (1994), and Adolfson, Laseen, Linde, Villani (2008).

Following an important contribution of Boivin and Giannoni (2006), *data-rich* DSGE models relax these assumptions and allow for: (i) the presence of measurement errors or, alternatively, of terms capturing the theoretical gap between a particular data indicator and a model concept it is supposed to measure; (ii) multiple data indicators $X_{j,t}$ measuring the same model concept $\overline{S}_{i,t}$, and (iii) many informational data series in X_t with an unknown link to specific model concepts that load on all DSGE model states (and that may contain useful information about the state of the economy). We call the *core series* X_t^F the part of X_t in which each data indicator loads on a single model concept $\overline{S}_{i,t}$ only (although same $\overline{S}_{i,t}$ may have several data indicators measuring it):

$$X_t^F = \mathbf{\Lambda_F}\overline{S}_t + e_t^F , \tag{5}$$

where each row of $\mathbf{\Lambda_F}$ contains just one non-zero element. We call the *non-core series* X_t^S the remaining part of X_t that is not supposed to measure any model concept and therefore loads freely on all DSGE model states:

$$X_t^S = \mathbf{\Lambda_S} S_t + e_t^S \tag{6}$$

For example, in a simple closed-economy DSGE model of Lubik and Schorfheide (2004), the core series might have been various measures of real output (e.g., real GDP, industrial production), of inflation (e.g., CPI inflation, PCE deflator inflation) or of the nominal interest rate; the non-core series might include exchange rates, real exports and imports, stock returns and similar data indicators not related directly to any model concept. We partition $\mathbf{\Lambda_F} = \left[\mathbf{\Lambda_{F,1}} \mid \mathbf{\Lambda_{F,2}} \right]$ conformably and use definition (3) to obtain the measurement equation in the data-rich DSGE model for demeaned X_t:

$$\underbrace{\begin{bmatrix} X_t^F \\ \hline X_t^S \end{bmatrix}}_{X_t} = \underbrace{\begin{bmatrix} \mathbf{\Lambda_{F,1}} \mathbf{D}(\boldsymbol{\theta}) + \mathbf{\Lambda_{F,2}} \\ \hline \mathbf{\Lambda_S} \end{bmatrix}}_{\mathbf{\Lambda}(\boldsymbol{\theta})} S_t + \underbrace{\begin{bmatrix} e_t^F \\ \hline e_t^S \end{bmatrix}}_{e_t}, \qquad (7)$$

where the measurement errors e_t may be serially correlated, but uncorrelated across different data indicators ($\mathbf{\Psi}$, $\mathbf{R}$ are diagonal):

$$e_t = \mathbf{\Psi} e_{t-1} + v_t, \qquad v_t \sim iid\ N(\mathbf{0}, \mathbf{R}). \qquad (8)$$

So the state-space representation of the data-rich DSGE model consists of transition equation (2) and measurement equations (7)-(8).

2.2 Environment

In this chapter, we use a relatively standard New Keynesian business cycle core that will be shared by the data-rich and the regular DSGE models. It features capital as the factor of production, nominal rigidities in price setting, and investment adjustment costs. The real money stock enters households' utility in additively separable fashion as in Walsh (2003, Ch. 5), and Sidrauski (1967). In terms of a specific version of the model, we draw upon the work of Aruoba and Schorfheide (2009) and their money-in-the-utility specification.

The economy is populated by households, final and intermediate goods-producing firms and a central bank (monetary authority). A representative household works, consumes, saves, holds money balances and accumulates capital. It consumes the final output manufactured by perfectly competitive final good firms. The final good producers produce by combining a continuum of differentiated intermediate goods supplied by monopolistically competitive intermediate goods firms. To manufacture their output, intermediate goods producers hire labor and capital services from households. Also, when optimizing their prices, intermediate goods firms face the nominal price rigidity a la Calvo (1983), and those firms that are unable to re-optimize may index their price to lagged inflation. Monetary policy is conducted by the central bank setting the one-period nominal interest rate on public debt via a Taylor-type interest rate feedback rule. Given the interest rate, the central bank supplies enough nominal money balances to meet equilibrium demand from households.

Our DSGE model is more elaborate than the basic three-equation model used in Woodford (2003), but is "lighter" than the models in Smets and Wouters (2003, 2007) and Christiano, Eichenbaum and Evans (2005): it abstracts from wage rigidities, habit formation in consumption and variable capital utilization.

2.2.1 Households

In our environment, there is a continuum of households indexed by $j \in [0;1]$. Each household maximizes the following utility function:

$$E_0 \sum_{t=0}^{\infty} \beta^t \left\{ U(x_t(j)) - Ah_t(j) + \frac{\chi_t}{1-\nu_m} \left[\frac{A}{Z_*^{1/(1-\alpha)}} \frac{m_t(j)}{P_t} \right]^{(1-\nu_m)} \right\}, \tag{9}$$

which is additively separable in consumption $x_t(j)$, labor supply $h_t(j)$ and real money balances $m_t(j)/P_t$. Here β stands for the discount factor, A denotes disutility of labor,

ν_m controls the elasticity of money demand and χ_t is an aggregate preference shifter that affects households' marginal utility from holding real money balances.[3] The law of motion for χ_t is:

$$\ln \chi_t = (1-\rho_\chi)\ln \chi_* + \rho_\chi \ln \chi_{t-1} + \varepsilon_{\chi,t}, \qquad \text{where } \varepsilon_{\chi,t} \sim N(0,\sigma_\chi^2) \tag{10}$$

We assume that households are able to trade on a complete set of Arrow-Debreu (A-D) securities, which are contingent on all aggregate and idiosyncratic events $\omega \in \Omega$ in the economy. Let $a_{t+1}(j)(\omega)$ denote the quantity of A-D securities (that pay 1 unit of consumption in period $t+1$ in the event ω) acquired by household j at time t at real price $q_{t+1,t}(j)$. Then household j's budget constraint in nominal terms is given by:

$$\begin{aligned} &P_t x_t(j) + P_t i_t(j) + b_{t+1}(j) + m_{t+1}(j) + P_t \int_\Omega q_{t+1,t}(j) a_{t+1}(j)(\omega) d\omega = \\ &= P_t W_t h_t(j) + P_t R_t^k k_t(j) + \Pi_t + R_{t-1} b_t(j) + m_t(j) + P_t a_t(j) - T_t \end{aligned} \tag{11}$$

where P_t is the period t price of the final good, $i_t(j)$ is investment, $b_t(j)$ and $m_t(j)$ are government bond and money holdings, R_t is the gross nominal interest rate on government bonds, W_t and R_t^k are the real wage and real return on capital earned by households, Π_t stands for profits from owning the firms, and T_t is the nominal amount of lump-sum taxes paid. Households also accumulate capital $k_t(j)$ according to the following law of motion:

[3] As in Aruoba and Schorfheide (2009), scaling $m_t(j)/P_t$ by a factor $A/Z_*^{1/(1-\alpha)}$ can be viewed as re-parameterization of χ_t, in which the steady-state money velocity remains constant when we move around A and Z_*.

$$k_{t+1}(j)=(1-\delta)k_t(j)+\left[1-S\left(\frac{i_t(j)}{i_{t-1}(j)}\right)\right]i_t(j), \tag{12}$$

where δ is the depreciation rate and $S(\bullet)$ is an adjustment cost function satisfying $S(1)=0$, $S'(1)=0$ and $S''(1)>0$.

The problem of each household j is to maximize the utility function (9) subject to budget constraint (11) and capital accumulation equation (12) for all t. Associate Lagrange multipliers $\lambda_t(j)$ and $Q_t(j)$ with constraints (11) and (12), respectively. The first-order conditions are provided in Appendix A1. We do not take the first-order conditions with respect to A-D securities holdings $a_{t+1}(j)$ explicitly, because we make use of the result in Erceg, Henderson and Levin (2000). This result says that under the assumption of complete markets for A-D securities and under the additive separability of labor and money balances in households' utility, the equilibrium price of A-D securities will be such that optimal consumption will not depend on idiosyncratic shocks. Hence, all households will share the same marginal utility of consumption, and the Lagrange multiplier $\lambda_t(j)$ will also be the same across all households: $\lambda_t(j)=\lambda_t$, all *j* and *t*. This implies that in equilibrium all households will choose the same consumption, money and bond holdings, investment and capital. Note that we don't have wage rigidity in this model: therefore, the choice of optimal labor will also be same. Therefore we can safely drop index j from all household-related conditions and variables and proceed accordingly.

Let us define the stochastic discount factor $\Xi^{p}_{t+1|t}$ that the firms – whose behavior we are going to describe shortly – will use to value streams of future profits:

$$\Xi^{p}_{t+1|t} = \frac{\lambda_{t+1}}{\lambda_t} = \frac{U'(x_{t+1})}{U'(x_t)} \frac{1}{\pi_{t+1}}, \tag{13}$$

where $\pi_t = P_t / P_{t-1}$ denotes final good price inflation.

2.2.2 Final Good Firms

There is single final good Y_t in our economy manufactured by combining a continuum of intermediate goods $Y_t(i)$ indexed by $i \in [0;1]$ according to the following production function:

$$Y_t = \left(\int_0^1 Y_t(i)^{\frac{1}{1+\lambda}} di \right)^{(1+\lambda)}, \tag{14}$$

where the elasticity of substitution between any goods *i* and *j* is $\frac{1+\lambda}{\lambda}$.

The final good firms purchase intermediate goods in the market, package them into a composite final good, and sell the final good to households. These firms are perfectly competitive and maximize one-period profits subject to production function (14), taking as given intermediate goods prices $P_t(i)$ and own output price P_t:

$$\begin{aligned} &\max_{Y_t, Y_t(i)} \quad P_t Y_t - \int_0^1 P_t(i) Y_t(i) di \\ &\text{s.t.} \quad Y_t = \left(\int_0^1 Y_t(i)^{\frac{1}{1+\lambda}} di \right)^{(1+\lambda)} \end{aligned} \tag{15}$$

The first-order condition leads to the optimal demand for good *i*:

$$Y_t(i) = \left(\frac{P_t(i)}{P_t} \right)^{-\frac{(1+\lambda)}{\lambda}} Y_t. \qquad (16)$$

Since final good firms are perfectly competitive and there is free entry, they earn zero profits in equilibrium, which, together with optimal demand (16), yields the price of the final good:

$$P_t = \left[\int_0^1 P_t(i)^{-\frac{1}{\lambda}} di \right]^{-\lambda}. \qquad (17)$$

2.2.3 Intermediate Goods Firms

Our economy is populated by a continuum of intermediate goods firms. Each intermediate goods firm i uses the following technology to produce its output:

$$Y_t(i) = \max\left\{ Z_t K_t(i)^{\alpha} H_t(i)^{(1-\alpha)} - \tilde{F}, 0 \right\}, \qquad (18)$$

where $K_t(i)$ is the amount of capital that the firm i rents from households, $H_t(i)$ is the amount of labor input and Z_t is the level of neutral technology evolving according to the law of motion:

$$\ln Z_t = (1-\rho_Z)\ln Z_* + \rho_Z \ln Z_{t-1} + \varepsilon_{Z,t}, \text{ where } \varepsilon_{Z,t} \sim N(0, \sigma_Z^2). \qquad (19)$$

Parameter α stands for the capital share of production, while parameter $\tilde{F}$ controls the amount of fixed costs in production that guarantee that the firm's economic profits will be zero in the steady state. Unlike with the final good producers, we do not allow for free entry or exit on the part of the intermediate goods firms.

All intermediate goods producers are *monopolistically competitive*, in that they take all factor prices (W_t and R_t^k), as well as the prices of other firms, as given, but can optimally choose their own price $P_t(i)$ subject to optimal demand (16) for good i from final good firms. Intermediate firms solve a two-stage optimization problem.

In the first stage, the firms hire capital and labor from households to minimize total nominal costs:

$$\begin{aligned} \min_{K_t(i),H_t(i)} \quad & P_t W_t H_t(i) + P_t R_t^k K_t(i) \\ \text{s.t.} \quad & Y_t(i) = \max\left\{ Z_t K_t(i)^{\alpha} H_t(i)^{(1-\alpha)} - \tilde{F}, 0 \right\} \end{aligned} \tag{20}$$

Assuming interior solution, optimality conditions imply ($\eta_t(i)$ is the Lagrange multiplier attached to (18)):

$$P_t W_t = \eta_t(i) P_t(i)(1-\alpha) Z_t K_t(i)^{\alpha} H_t(i)^{-\alpha}$$

$$P_t R_t^k = \eta_t(i) P_t(i) \alpha Z_t K_t(i)^{\alpha-1} H_t(i)^{1-\alpha}$$

Take the ratio of two conditions to obtain:

$$\frac{K_t(i)}{H_t(i)} = \frac{\alpha}{1-\alpha} \frac{W_t}{R_t^k} \tag{21}$$

If we define aggregate capital stock $K_t = \int_0^1 K_t(i)di$ and aggregate labor $H_t = \int_0^1 H_t(i)di$, integrating both sides of (21) yields:

$$K_t = \frac{\alpha}{1-\alpha} \frac{W_t}{R_t^k} H_t \tag{22}$$

Now we can factorize total real variable cost $VC_t(i)$ into real marginal cost MC_t and the variable part of firm i's output $Y_t^{\text{var}}(i) = Z_t K_t(i)^{\alpha} H_t(i)^{(1-\alpha)}$:

$$VC_t(i) = \left(W_t + R_t^k \frac{K_t(i)}{H_t(i)} \right) H_t(i) = \left(W_t + R_t^k \frac{K_t(i)}{H_t(i)} \right) \frac{1}{Z_t} \left(\frac{K_t(i)}{H_t(i)} \right)^{-\alpha} Y_t^{\text{var}}(i) \quad (23)$$

Plugging in the optimal capital labor ratio (21), real marginal cost MC_t turns out to be the same across all intermediate goods firms:

$$MC_t \overset{def}{=} \left(W_t + R_t^k \frac{K_t(i)}{H_t(i)} \right) \frac{1}{Z_t} \left(\frac{K_t(i)}{H_t(i)} \right)^{-\alpha} = \left(\frac{1}{\alpha} \right)^{\alpha} \left(\frac{1}{1-\alpha} \right)^{(1-\alpha)} \frac{W_t^{1-\alpha} \left(R_t^k \right)^{\alpha}}{Z_t} \quad (24)$$

The intuition is that all firms face identical technology shocks and hire inputs at the same factor prices.

In the second stage, all intermediate goods firms have to choose their own price $P_t(i)$ that maximizes total discounted nominal profits subject to demand curve (16). Given optimal choices of inputs from the first stage, the one-period nominal profits of firm i are:

$$\begin{aligned} \Pi_t(i) &= P_t(i) Y_t(i) - P_t W_t \tilde{H}_t(i) - P_t R_t^k \tilde{K}_t(i) = P_t(i) Y_t(i) - P_t \left(MC_t Y_t^{\text{var}}(i) \right) = \\ &= \left(P_t(i) - P_t MC_t \right) Y_t(i) - P_t MC_t \tilde{F} \end{aligned} \quad (25)$$

Note that we can ignore the term $P_t MC_t \tilde{F}$ since it doesn't depend on a firm's choice.

We assume that intermediate goods firms face *nominal price rigidity* a la Calvo (1983). In each period, a fraction $(1-\zeta)$ of firms can optimize their prices. As in Aruoba and Schorfheide (2009), we modify Calvo's original set-up and assume that all other firms cannot adjust their prices and can only index $P_t(i)$ by a geometric weighted average of the fixed rate π_{**} and of the previous period's inflation π_{t-1}, with weights $(1-\iota)$ and ι respectively. The corresponding price adjustment factor is:

$$\pi^{adj}_{t+s|t} = \begin{cases} 1, & s=0 \\ \prod_{l=1}^{s}\left(\pi^{\iota}_{t+l-1}\pi^{(1-\iota)}_{**}\right), & s>0 \end{cases} \tag{26}$$

The firms allowed to re-optimize must choose the optimal price $P^o_t(i)$ that maximizes the discounted value of profits in all states of nature in which the firm faces that price in the future:

$$\begin{aligned} &\max_{P^o_t(i)} \quad \Xi^p_{t|t}(P^o_t(i) - P_t MC_t)Y_t(i) + E_t\left\{\sum_{s=1}^{\infty}(\zeta\beta)^s \Xi^p_{t+s|t}(P^o_t(i)\pi^{adj}_{t+s|t} - P_{t+s}MC_{t+s})Y_{t+s}(i)\right\} \\ &\text{s.t.} \quad Y_{t+s}(i) = \left[\frac{P^o_t(i)\pi^{adj}_{t+s|t}}{P_{t+s}}\right]^{-\frac{(1+\lambda)}{\lambda}} Y_{t+s}, \qquad s = 0,1,2,.... \end{aligned} \tag{27}$$

Notice that $\beta^s \Xi^p_{t+s|t}$ is the period *t* value of a future dollar for the consumer/household in period *t*+*s*.

Since we consider only a *symmetric equilibrium* in which all firms re-optimizing their prices will choose the same price $P^o_t(i) = P^o_t$, we can drop the indices i from firms' conditions and variables. Given (17) and Calvo pricing, the aggregate price index P_t should evolve as:

$$P_t = \left[(1-\zeta)\left(P^o_t\right)^{-\frac{1}{\lambda}} + \zeta\left(\pi^{\iota}_{t-1}\pi^{(1-\iota)}_{**}P_{t-1}\right)^{-\frac{1}{\lambda}}\right]^{-\lambda} \tag{28}$$

and, dividing by P_{t-1} and defining $p^o_t = P^o_t / P_t$, yields:

$$\pi_t = \left[(1-\zeta)\left(\pi_t p_t^o\right)^{-\frac{1}{\lambda}} + \zeta\left(\pi_{t-1}^{\iota}\pi_{**}^{(1-\iota)}\right)^{-\frac{1}{\lambda}} \right]^{-\lambda} \tag{29}$$

As is standard in the literature, the first-order conditions (Appendix A2) of intermediate firms' problem (27) connect the evolution of inflation to the dynamics of real marginal costs and output, and thus imply the New Keynesian Phillips curve.

2.2.4 Monetary and Fiscal Policy

The central bank sets the one-period nominal interest rate on public debt via a Taylor-type interest rate feedback rule responding to deviations of inflation and real output from their target levels:

$$\frac{R_t}{R_*} = \left(\frac{R_{t-1}}{R_*}\right)^{\rho_R} \left(\left(\frac{\pi_t}{\pi_*}\right)^{\psi_1} \left(\frac{Y_t}{Y_*}\right)^{\psi_2} \right)^{(1-\rho_R)} e^{\varepsilon_{R,t}}, \qquad \text{where } \varepsilon_{R,t} \sim N(0,\sigma_R^2) \tag{30}$$

where R_*, π_* and Y_* are the steady-state values of the gross nominal interest rate, final good inflation and real final output, respectively. Parameter ρ_R is introduced to control for the degree of interest rate smoothing that we observe in the postwar U.S. data. Also, the central bank supplies enough money balances M_t to meet demand from households, given the desired nominal interest rate.

Every period the government spends G_t in real terms to purchase goods in the final goods market, issues nominal bonds B_{t+1} that pay R_t in gross interest next period and collects nominal lump-sum taxes from households T_t. Each period, the combined government (central bank + Treasury) budget constraint is:

$$P_t G_t + R_{t-1} B_t + M_t = T_t + B_{t+1} + M_{t+1} \tag{31}$$

Real government spending is modeled as a stochastic fraction of total output (i.e., fiscal policy is passive):

$$G_t = \left(1 - \frac{1}{g_t}\right) Y_t, \tag{32}$$

where g_t is an exogenous process shifting G_t:

$$\ln g_t = (1-\rho_g)\ln g_* + \rho_g \ln g_{t-1} + \varepsilon_{g,t}, \text{ where } \varepsilon_{g,t} \sim N(0, \sigma_g^2). \tag{33}$$

2.2.5 Aggregation

We now derive the aggregate demand condition. To that end, we integrate budget constraints across all households and combine the result with the government budget constraint (31), introducing aggregate variables – consumption $X_t = \int_0^1 x_t(j)dj$ and investment $I_t = \int_0^1 i_t(j)dj$:

$$P_t X_t + P_t I_t + P_t G_t = P_t W_t H_t + P_t R_t^k K_t + \Pi_t . \tag{34}$$

We derive the expression for aggregate profits Π_t from intermediate firms' problems, combine it with (34) and divide the result by P_t to obtain the *aggregate demand condition*:

$$X_t + I_t + G_t = Y_t \tag{35}$$

From the supply side, the aggregate output of intermediate goods firms $\overline{Y}_t$ is given by:

$$\overline{Y}_t = \int_0^1 Z_t K_t(i)^{\alpha} H_t(i)^{(1-\alpha)} di - \tilde{F} = Z_t \int_0^1 \left(\frac{K_t(i)}{H_t(i)} \right)^{\alpha} H_t(i) di - \tilde{F} = Z_t K_t^{\ \alpha} H_t^{(1-\alpha)} - \tilde{F}, \quad (36)$$

where we have used the fact that the capital/labor ratio is constant across firms. However, from (16):

$$\overline{Y}_t = \int_0^1 Y_t(i) di = Y_t \int_0^1 \left(\frac{P_t(i)}{P_t} \right)^{-\frac{(1+\lambda)}{\lambda}} di \quad (37)$$

Hence, the *aggregate supply condition* becomes:

$$Y_t = \frac{1}{D_t} (Z_t K_t^{\alpha} H_t^{1-\alpha} - \tilde{F}), \quad (38)$$

with $D_t = \int_0^1 \left(\frac{P_t(i)}{P_t} \right)^{-\frac{(1+\lambda)}{\lambda}} di$ measuring the extent of aggregate loss of efficiency caused by price dispersion across intermediate goods firms. In Appendix A3, we show that aggregate price dispersion D_t evolves according to:

$$D_t = \zeta \left[\left(\frac{\pi_{t-1}}{\pi_t} \right)^{\iota} \left(\frac{\pi_{**}}{\pi_t} \right)^{(1-\iota)} \right]^{-\frac{(1+\lambda)}{\lambda}} D_{t-1} + (1-\zeta) \left[\frac{P_t^o}{P_t} \right]^{-\frac{(1+\lambda)}{\lambda}} \quad (39)$$

For convenience, we collect all DSGE model parameters in the vector $\mathbf{\theta}$ and stack all innovations in vector $\varepsilon_t = [\varepsilon_{Z,t}, \varepsilon_{\chi,t}, \varepsilon_{g,t}, \varepsilon_{R,t}]'$. We then derive a log-linear approximation to the system of equilibrium conditions (summarized in Appendix A4 and A5) around its deterministic steady state. The resulting linear rational expectations system is solved by the method described in Sims (2002).

3 Econometric Methodology

In this section, we first provide the details on a Markov Chain Monte Carlo (MCMC) algorithm to estimate the data-rich DSGE model, including the choice of the prior for factor loadings. Second, we present the novel speed-up suggested by Jungbacker and Koopman (2008), which enhances the speed of our Bayesian estimation procedure.

3.1 Estimation of the Data-Rich DSGE Model

As discussed in the previous section, the state-space representation of our data-rich DSGE model consists of a transition equation of model states S_t and a set of measurement equations relating the states[4] to data X_t:

$$\underbrace{S_t}_{N\times 1} = \underbrace{\mathbf{G}(\boldsymbol{\theta})}_{N\times N}\underbrace{S_{t-1}}_{N\times 1} + \underbrace{\mathbf{H}(\boldsymbol{\theta})}_{N\times N_\varepsilon}\underbrace{\varepsilon_t}_{N_\varepsilon\times 1} \tag{40}$$

$$\underbrace{X_t}_{J\times 1} = \underbrace{\boldsymbol{\Lambda}(\boldsymbol{\theta})}_{J\times N}\underbrace{S_t}_{N\times 1} + \underbrace{e_t}_{J\times 1} \tag{41}$$

$$e_t = \boldsymbol{\Psi} e_{t-1} + v_t, \tag{42}$$

where $\varepsilon_t \sim iid\, N(\mathbf{0}, \mathbf{Q}(\boldsymbol{\theta}))$, $v_t \sim iid\, N(\mathbf{0}, \mathbf{R})$ and where $\mathbf{Q}(\boldsymbol{\theta})$, $\mathbf{R}$ and $\boldsymbol{\Psi}$ are assumed diagonal. An essential feature of a data-rich framework is that the panel dimension of data set J is much higher than the number of DSGE model states N. For convenience, collect state-space matrices from the measurement equation into $\Gamma = \{\boldsymbol{\Lambda}(\boldsymbol{\theta}), \boldsymbol{\Psi}, \mathbf{R}\}$ and DSGE states-factors into $S^T = \{S_1, S_2, \ldots, S_T\}$. Because of the normality of structural

[4] In measurement equations (41) we keep only the non-redundant state variables of a DSGE model. Because some of the DSGE states are merely linear combinations of the other states, one can interpret this as minimum-state-variable approach in the spirit of McCallum (1983, 1999, 2003). Here, though, the main rationale is to avoid multicollinearity on the right hand side of (41). We always set the corresponding factor loadings in $\boldsymbol{\Lambda}$ equal to zero.

shocks ε_t and measurement error innovations v_t, system (40)-(42) is a linear Gaussian state-space model and the likelihood function of data $p(X^T \mid \boldsymbol{\theta}, \Gamma)$ can be evaluated using a Kalman filter.

Following Boivin and Giannoni (2006), we use Bayesian techniques to estimate the unknown model parameters $(\boldsymbol{\theta}, \Gamma)$. We combine prior $p(\boldsymbol{\theta}, \Gamma) = p(\Gamma \mid \boldsymbol{\theta}) p(\boldsymbol{\theta})$ with the likelihood function $p(X^T \mid \boldsymbol{\theta}, \Gamma)$ to obtain the posterior distribution of parameters given data:

$$p(\boldsymbol{\theta}, \Gamma \mid X^T) = \frac{p(X^T \mid \boldsymbol{\theta}, \Gamma) p(\boldsymbol{\theta}, \Gamma)}{\int p(X^T \mid \boldsymbol{\theta}, \Gamma) p(\boldsymbol{\theta}, \Gamma) d\boldsymbol{\theta} d\Gamma} \tag{43}$$

We use Markov Chain Monte Carlo (MCMC) method to estimate posterior density $p(\boldsymbol{\theta}, \Gamma \mid X^T)$ by constructing a Markov chain with the property that its limiting invariant distribution is our posterior distribution. Similarly to Boivin and Giannoni (2006), the Markov chain is constructed by the Gibbs sampling method with a Metropolis-within-Gibbs step to generate draws from the posterior distribution $p(\boldsymbol{\theta}, \Gamma \mid X^T)$ and to compute the approximations to posterior means and covariances of parameters of interest.

But before we turn to describing the Gibbs sampler, we must elaborate on *how we connect the DSGE model states to data indicators*. This is important, because, unlike in Boivin and Giannoni (2006), the link is primarily through the prior on factor loadings $\boldsymbol{\Lambda}(\boldsymbol{\theta})$. The priors for the rest of the parameters ($\boldsymbol{\theta}$, $\boldsymbol{\Psi}$ and $\mathbf{R}$) are discussed in detail in the section "Empirical Results: Priors" below. Recall that we have *core* data series that measure specific model concepts and *non-core* informational variables that are related to all states of the DSGE model. Consider the following hypothetical example:

$$\begin{matrix}\text{core}\\ \\ \\ \text{non-core}\end{matrix}\begin{cases}\\ \\ \\ \end{cases}\underbrace{\begin{bmatrix}\text{output \#1}\\ \text{output \#2}\\ \text{inflation \#1}\\ \text{inflation \#2}\\ \vdots\\ \hline \text{exchange rate}\\ X_t^{S,\,rest}\end{bmatrix}}_{X_t}=\underbrace{\begin{bmatrix}\lambda'_{Y1}\\ \lambda'_{Y2}\\ \lambda'_{\pi1}\\ \lambda'_{\pi2}\\ \vdots\\ \hline \lambda'_{ER}\\ \mathbf{\Lambda}_S\end{bmatrix}}_{\mathbf{\Lambda(\theta)}}\cdot\underbrace{\begin{bmatrix}\hat{Y}_t\\ \hat{\pi}_t\\ \vdots\end{bmatrix}}_{S_t}+\underbrace{\begin{bmatrix}e_t^F\\ \hline e_t^S\end{bmatrix}}_{e_t} \tag{44}$$

As a matter of general principle, for each of the core series we center the prior mean of λ's at regular-DSGE-model-implied factor loadings of a corresponding model concept. In the example above, this corresponds to the conditional prior for core loadings being:

$$\begin{aligned}p(\lambda_{Y1}\,|\,\boldsymbol{\theta}) &= p(\lambda_{Y2}\,|\,\boldsymbol{\theta}) = N\left([1,0,0,...,0]',\Omega(\boldsymbol{\theta})\right)\\ p(\lambda_{\pi1}\,|\,\boldsymbol{\theta}) &= p(\lambda_{\pi2}\,|\,\boldsymbol{\theta}) = N\left([0,4,0,...,0]',\Omega(\boldsymbol{\theta})\right).\end{aligned} \tag{45}$$

This means that in regular DSGE model, the output #1 in the data is equal to 1 times output $\hat{Y}_t$ in the model, and inflation #1 in the data is equal to 4 times inflation $\hat{\pi}_t$ in the model (conversion from quarterly to annual inflation). In the data-rich DSGE model, we do not impose $\lambda_{Y,0}=[1,0,0,...,0]'$ and $\lambda_{\pi,0}=[0,4,0,...,0]'$ on loadings λ_Y and λ_π, but instead use them to center the prior means for λ_Y and λ_π. This is different from Boivin and Giannoni (2006), who restrict core factor loadings λ_Y and λ_π to be either $\lambda_{Y,0}$ and $\lambda_{\pi,0}$ or proportional to these.

For non-core series, we center the prior mean of factor loadings at zero vector with an identity covariance matrix. In terms of example (44), the conditional prior is:

$$p(\lambda_{ER} \mid \boldsymbol{\theta}) = p(\boldsymbol{\Lambda}'_{S,k} \mid \boldsymbol{\theta}) = N\left([0,0,0,...,0]', \mathbf{I}_N\right), \qquad (46)$$

where sub-index k selects one row from matrix $\boldsymbol{\Lambda}_S$.

Note that prior means for core loadings may in general depend on DSGE model parameters $\boldsymbol{\theta}$. For instance, if core series contain a measure of inverse money velocity IVM_t, then the DSGE model counterpart $\hat{M}_t - \hat{Y}_t$ (real money balances minus real output in logs) depends on state S_t indirectly, say, via $\hat{M}_t - \hat{Y}_t = d_{IVM}(\boldsymbol{\theta})S_t$. As a result, the conditional prior for loadings in the IVM measurement equation would be $p(\lambda_{IVM_1} \mid \boldsymbol{\theta}) = N\left(d_{IVM}(\boldsymbol{\theta})', \Omega(\boldsymbol{\theta})\right)$.

Also note that to prevent the data-rich DSGE model from drifting too far away from parameter estimates of a regular DSGE model and to fix the scale of the estimated DSGE model state variables, we make the prior for one of the core series within each core subgroup *perfectly tight*. In example (44), we have two subgroups of core series – output and inflation. This implies, without loss of generality, the perfectly tight prior on loadings in the output #1 and inflation #1 equations. Therefore, we write $\boldsymbol{\Lambda}(\boldsymbol{\theta})$ to underscore that some loadings will explicitly depend on the DSGE model's structural parameters.

Now let us turn to the description of our Gibbs sampler. MCMC implementation for the linear Gaussian state-space model (40)-(42) is based on the following conditional posterior distributions:

$$p(\Gamma \mid \boldsymbol{\theta}; X^T) \quad p(S^T \mid \Gamma, \boldsymbol{\theta}; X^T) \quad p(\Gamma \mid S^T, \boldsymbol{\theta}; X^T) \quad p(\boldsymbol{\theta} \mid \Gamma; X^T) \tag{47}$$

Essentially, the Gibbs sampler iterates on conditional posterior densities $p(\Gamma \mid \boldsymbol{\theta}; X^T)$ and $p(\boldsymbol{\theta} \mid \Gamma; X^T)$ to generate draws from the joint posterior distribution $p(\boldsymbol{\theta}, \Gamma \mid X^T)$ of the state-space parameters Γ and the structural DSGE model parameters $\boldsymbol{\theta}$. It uses an intermediate step to draw DSGE states S^T, because this simplifies sampling the elements of Γ conditional on S^T and $\boldsymbol{\theta}$. The sampling of $\boldsymbol{\theta}$ relies on a Metropolis-within-Gibbs step, since the conditional posterior density $p(\boldsymbol{\theta} \mid \Gamma; X^T)$ is generally intractable.

The main steps of the *Gibbs sampler* are (we provide full details in Appendix B):

1. Specify initial values $\boldsymbol{\theta}^{(0)}$ and $\Gamma^{(0)}$.
2. Repeat for $g = 1, 2, ..., n_{sim}$

   2.1. Solve the DSGE model numerically at $\boldsymbol{\theta}^{(g-1)}$ and obtain matrices $\mathbf{G}(\boldsymbol{\theta}^{(g-1)})$, $\mathbf{H}(\boldsymbol{\theta}^{(g-1)})$ and $\mathbf{Q}(\boldsymbol{\theta}^{(g-1)})$

   2.2. Draw from $p(\Gamma \mid \boldsymbol{\theta}^{(g-1)}; X^T)$:

   a) Generate unobserved states $S^{T,(g)}$ from $p(S^T \mid \Gamma^{(g-1)}, \boldsymbol{\theta}^{(g-1)}; X^T)$ using the Carter-Kohn (1994) forward-backward algorithm;

   b) Generate state-space parameters $\Gamma^{(g)}$ from $p(\Gamma \mid S^{T,(g)}, \boldsymbol{\theta}^{(g-1)}; X^T)$ by drawing from a complete set of known conditional densities $[\mathbf{R} \mid \boldsymbol{\Lambda}, \boldsymbol{\Psi}; \Xi]$, $[\boldsymbol{\Lambda} \mid \mathbf{R}, \boldsymbol{\Psi}; \Xi]$ and $[\boldsymbol{\Psi} \mid \boldsymbol{\Lambda}, \mathbf{R}; \Xi]$, where $\Xi = \left\{ S^{T,(g)}, \boldsymbol{\theta}^{(g-1)}, X^T \right\}$.

   2.3. Draw DSGE parameters $\boldsymbol{\theta}^{(g)}$ from $p(\boldsymbol{\theta} \mid \Gamma^{(g)}; X^T)$ using Metropolis step:

   a) Propose

$$\boldsymbol{\theta}^* \sim q(\boldsymbol{\theta} \mid \boldsymbol{\theta}^{(g-1)}; \Gamma^{(g)}) \tag{48}$$

b) Draw $u \sim Uniform(0,1)$ and set

$$\boldsymbol{\theta}^{(g)} = \begin{cases} \boldsymbol{\theta}^* & \text{if } u \le \alpha(\boldsymbol{\theta}^* \mid \Gamma^{(g)}, \boldsymbol{\theta}^{(g-1)}) \\ \boldsymbol{\theta}^{(g-1)} & \text{otherwise} \end{cases} \tag{49}$$

where acceptance probability $\alpha(\bullet) = \min\left\{1, r(\boldsymbol{\theta}^{(g-1)}, \boldsymbol{\theta}^*, \Gamma^{(g)})\right\}$ and

$$r(\boldsymbol{\theta}^{(g-1)}, \boldsymbol{\theta}^*, \Gamma^{(g)}) = \frac{p(\boldsymbol{\theta}^*, \Gamma^{(g)} \mid X^T)}{p(\boldsymbol{\theta}^{(g-1)}, \Gamma^{(g)} \mid X^T)} = \frac{p(X^T \mid \boldsymbol{\theta}^*, \Gamma^{(g)})\, p(\Gamma^{(g)} \mid \boldsymbol{\theta}^*)\, p(\boldsymbol{\theta}^*)}{p(X^T \mid \boldsymbol{\theta}^{(g-1)}, \Gamma^{(g)})\, p(\Gamma^{(g)} \mid \boldsymbol{\theta}^{(g-1)})\, p(\boldsymbol{\theta}^{(g-1)})}. \tag{50}$$

3. Return $\left\{\boldsymbol{\theta}^{(g)}, \Gamma^{(g)}\right\}_{g=1}^{n_{sim}}$

The Carter-Kohn (1994) algorithm in step 2.2.(a) proceeds as follows. First, it applies a Kalman filter to the state-space system (40)-(42) to generate filtered DSGE states $\hat{S}_{t|t}$, $t = 1..T$. Then, starting from $\hat{S}_{T|T}$, it rolls back in time along Kalman smoother recursions to draw elements of $S^{T,(g)}$ from a sequence of conditional Gaussian distributions.

The intermediate step to generate DSGE model states $S^{T,(g)}$ is used to facilitate sampling state-space matrices $\Gamma^{(g)}$ in 2.2.(b). Conditional on $S^{T,(g)}$, the elements of matrices $\Gamma^{(g)} = \left\{\boldsymbol{\Lambda}^{(g)}, \boldsymbol{\Psi}^{(g)}, \mathbf{R}^{(g)}\right\}$ are the parameters of simple linear regressions (41)-(42) and we can draw them equation by equation using the approach of Chib and Greenberg (1994). It is a straightforward procedure, since we assume conjugate priors for Γ and conditional posterior densities are all of known functional forms.

To generate DSGE model parameters $\boldsymbol{\theta}^{(g)}$, we introduce Metropolis step 2.3. It is required because density $p(\boldsymbol{\theta} \mid \Gamma; X^T)$ is generally intractable and cannot be easily factorized into known conditionals. We choose to use the *random-walk version of Metropolis step* (e.g., An and Schorfheide, 2007) in which the proposal density $q(\boldsymbol{\theta}' \mid \boldsymbol{\theta})$ is a multivariate Student-t with mean equal to the previous draw $\boldsymbol{\theta}^{(g-1)}$ and a covariance matrix proportional to the inverse Hessian from the *regular* DSGE model[5] evaluated at the posterior mode.

To *initialize* our Gibbs sampler, we first run a regular DSGE model estimation (see footnote 5), compute the posterior mean of DSGE model parameters and generate smoothed model states $S^{T,reg}$. Then we take the rich panel of macro and financial series X^T and run equation-by-equation OLS regressions of X_k^T on smoothed DSGE states $S^{T,reg}$ to back out initial values for $\boldsymbol{\Lambda}$, $\boldsymbol{\Psi}$ and $\mathbf{R}$.

Under regularity conditions satisfied here for the linear Gaussian state-space model, the Markov chain $\{\boldsymbol{\theta}^{(g)}, \Gamma^{(g)}\}$ constructed by the Gibbs sampler above converges to its invariant distribution and, starting from some $g > \bar{g}$, contains draws from the posterior distribution of interest $p(\boldsymbol{\theta}, \Gamma \mid X^T)$. Sample averages of these draws (or their appropriate transformations) converge almost surely to respective population moments under our posterior density (Tierney 1994, Chib 2001, Geweke 2005).

---

[5] Running a bit ahead, in our empirical analysis this regular DSGE estimation features the same underlying theoretical DSGE model as in the data-rich version, but only four (equal to the number of shocks) core observables assumed to have been measured without errors. These core observables are (appropriately transformed) real GDP, GDP deflator inflation, the federal funds rate and the inverse velocity of money based on M2S. See details in the Data and Transformations section. Also see the notes to Table D3.

3.2 Speed-Up: Jungbacker and Koopman 2008

The data-rich DSGE model (40)-(42) is potentially a high-dimensional object (the panel dimension J could be as high as 100+), and therefore, the MCMC algorithm outlined above spends a lot of time evaluating the likelihood function with the Kalman filter and sampling the DSGE states S_t at every iteration. To reduce the computational costs associated with a likelihood-based analysis of dynamic factor models (of which our data-rich DSGE model is a special case), Jungbacker and Koopman (2008) proposed to use the Kalman filter and smoother techniques based on a *lower-dimensional transformation* of the original data vector X_t.

Without loss of generality, consider the generic data-rich DSGE model introduced in section 2. The first-order dynamics of errors e_t allow us to rewrite the system (2), (7)-(8) in state-space form as follows:

$$\tilde{X}_t = \underbrace{\left[\mathbf{\Lambda}(\boldsymbol{\theta}) \;\vdots\; -\mathbf{\Psi}\mathbf{\Lambda}(\boldsymbol{\theta})\right]}_{\tilde{\mathbf{\Lambda}}} \underbrace{\begin{bmatrix} S_t \\ S_{t-1} \end{bmatrix}}_{\tilde{F}_t} + v_t \tag{51}$$

$$\tilde{F}_t = \underbrace{\begin{bmatrix} \mathbf{G}(\boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{I} & \mathbf{0} \end{bmatrix}}_{\tilde{\mathbf{G}}} \tilde{F}_{t-1} + \underbrace{\begin{bmatrix} \mathbf{H}(\boldsymbol{\theta}) \\ \mathbf{0} \end{bmatrix}}_{\tilde{\mathbf{H}}} \varepsilon_t, \tag{52}$$

where we denoted $\tilde{X}_t = X_t - \mathbf{\Psi} X_{t-1}$. Collect all the matrices in $\mathbf{\Theta} = \left\{\mathbf{\Lambda}, \mathbf{\Psi}, \mathbf{R}, \tilde{\mathbf{G}}, \tilde{\mathbf{H}}, \mathbf{Q}\right\}$. Suppose that the proposed lower-dimensional transformation of data vector $\tilde{X}_t$ is implemented by some $J \times J$ invertible matrix $\mathbf{A}$ such that $\tilde{X}_t^* = \mathbf{A}\tilde{X}_t$, $t = 1..T$. Also, suppose that we partition $\tilde{X}_t^*$ and $\mathbf{A}$ as below:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}^L \\ \mathbf{A}^H \end{bmatrix}, \qquad \tilde{X}_t^* = \begin{bmatrix} \tilde{X}_t^L \\ \tilde{X}_t^H \end{bmatrix}, \qquad \text{where } \tilde{X}_t^L = \mathbf{A}^L \tilde{X}_t,\ \tilde{X}_t^H = \mathbf{A}^H \tilde{X}_t, \tag{53}$$

with matrices $\mathbf{A}^L$ and $\mathbf{A}^H$ being $m \times J$ and $(J-m) \times J$, $m < J$.

Jungbacker and Koopman (2008) are able to show (Lemma 1, Lemma 2) that you can find a suitable matrix $\mathbf{A}$ such that $\tilde{X}_t^L$ and $\tilde{X}_t^H$ are uncorrelated and only the low-dimensional sub-vector $\tilde{X}_t^L$ depends on DSGE states $\tilde{F}_t$:

$$\begin{aligned} \tilde{X}_t^L &= \mathbf{A}^L \tilde{\mathbf{\Lambda}} \tilde{F}_t + v_t^L, \\ \tilde{X}_t^H &= v_t^H, \end{aligned} \qquad \begin{bmatrix} v_t^L \\ v_t^H \end{bmatrix} \sim iidN\left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{\Sigma}_L & \mathbf{0} \\ \mathbf{0} & \mathbf{\Sigma}_H \end{bmatrix} \right), \tag{54}$$

where $\mathbf{\Sigma}_L = \mathbf{A}^L \mathbf{R} \mathbf{A}^{L\prime}$ and $\mathbf{\Sigma}_H = \mathbf{A}^H \mathbf{R} \mathbf{A}^{H\prime}$. Moreover, they show that the knowledge of a high-dimensional matrix $\mathbf{A}^H$ and a data vector $\tilde{X}_t^H$ is not required to estimate the DSGE states $\tilde{F}_t$ and to compute the likelihood of the original model.

In terms of matrix $\mathbf{A}^L$, Jungbacker and Koopman prove that it should be of the form:

$$\mathbf{A}^L = \mathbf{C} \bar{\mathbf{\Lambda}}' \mathbf{R}^{-1}, \tag{55}$$

for some invertible $m \times m$ matrix $\mathbf{C}$ and $J \times m$ matrix $\bar{\mathbf{\Lambda}}$, columns of which form a basis of the column space of $\tilde{\mathbf{\Lambda}}$. In practice, they recommend setting $\bar{\mathbf{\Lambda}} = \tilde{\mathbf{\Lambda}}$ and $\mathbf{C} = \left(\tilde{\mathbf{\Lambda}}' \mathbf{R}^{-1} \tilde{\mathbf{\Lambda}} \right)^{-1}$ in case the matrix of factor loadings $\tilde{\mathbf{\Lambda}}$ has full column rank.

Now that we know $\mathbf{A}^L$ we can sample states $\tilde{F}_t$ using the Carter-Kohn (1994) forward-backward algorithm applied to a lower-dimensional model

$$\tilde{X}_t^L = \mathbf{A}^L \tilde{\mathbf{\Lambda}} \tilde{F}_t + v_t^L, \qquad v_t^L \sim iid\ N(\mathbf{0}, \mathbf{\Sigma}_L) \tag{56}$$

$$\tilde{F}_t = \tilde{\mathbf{G}}\tilde{F}_{t-1} + \tilde{\mathbf{H}}\varepsilon_t, \qquad \varepsilon_t \sim iid\ N(\mathbf{0}, \mathbf{Q}(\boldsymbol{\theta})). \tag{57}$$

We can also compute the log-likelihood of data $L(\tilde{X} \mid \mathbf{\Theta})$ as

$$L(\tilde{X} \mid \mathbf{\Theta}) = c + L(\tilde{X}^L \mid \mathbf{\Theta}) - \frac{T}{2}\log\frac{|\mathbf{R}|}{|\mathbf{\Sigma}_L|} - \frac{1}{2}\sum_{t=1}^{T}\hat{v}_t'\mathbf{R}^{-1}\hat{v}_t, \tag{58}$$

where $c = -\frac{1}{2}(J-m)T\log(2\pi)$ and $\hat{v}_t = \tilde{X}_t - \left[\bar{\mathbf{\Lambda}}\left(\bar{\mathbf{\Lambda}}'\mathbf{R}^{-1}\bar{\mathbf{\Lambda}}\right)\bar{\mathbf{\Lambda}}'\mathbf{R}^{-1}\right]\tilde{X}_t$. The term $L(\tilde{X}^L \mid \mathbf{\Theta})$ is the log-likelihood of the transformed data evaluated by using a Kalman filter during the forward pass of the Carter-Kohn algorithm on the low-dimensional model (56)-(57).

In the ensuing empirical analysis of a data-rich DSGE model, we have applied the Jungbacker-Koopman algorithm presented in this section to improve the speed of computations. To get a sense of CPU time gains, we have also estimated the model – though on fewer draws – without the speed-up and have found that the “improved” estimation of the data-rich DSGE model runs 2.5 times faster. The CPU gains reported by Jungbacker and Koopman (2008) for a dynamic factor model of a size similar to our data-rich DSGE model are about 11 times faster. Differences in time savings are due to the significant chunk of time that it takes to solve numerically the underlying DSGE model in the data-rich DSGE model estimation, a step absent in the DFM estimation and not affected by the Jungbacker-Koopman speed-up.

4 Data and Transformations

To estimate the data-rich DSGE model, we employ a large panel of U.S. quarterly macroeconomic and financial time series compiled by Stock and Watson (2008).[6] The panel covers 1959:Q1 – 2006:Q4, however, our sample in this chapter spans only 1984:Q1 – 2005:Q4. We focus on this later period primarily for two reasons: (i) to avoid dealing with the issue of the Great Moderation[7]; and (ii) to concentrate on a period with a relatively stable monetary policy regime.

Our data set consists of 12 *core series* that measure specific DSGE model concepts and 77 *non-core* informational series that load on all DSGE states and may contain useful information about the aggregate state of the economy. The core series include three measures of real output (real GDP, the index of total industrial production and the index of industrial production: manufacturing), three measures of price inflation (GDP deflator inflation, personal consumption expenditure (PCE) deflator inflation, and CPI inflation), three indicators of the nominal interest rates (the federal funds rate, the 3-month T-bill rate and the yield on AAA-rated corporate bonds), and three series measuring the inverse velocity of money (IVM based on the M1 aggregate and the M2 aggregate and IVM based on the adjusted monetary base). The 77 non-core series include the measures of real activity, labor market variables, housing indicators, prices and wages, financial variables (interest rate spreads, exchange rate depreciations, credit

[6] The data set is available online at:
http://www.princeton.edu/~mwatson/ddisk/hendryfestschrift_replicationfiles_April28_2008.zip
[7] The "Great Moderation" refers to a decline in the volatility of output and inflation observed in the U.S. since the mid-1980s until the recent financial crisis. For evidence and implications, please see Bernanke (2004), Stock and Watson (2002c), Kim and Nelson (1999a), and McConnell and Perez-Quiros (2000). The last two papers argue that a break in the volatility of U.S. GDP growth occurred in 1984:Q1.

stocks, stock returns) and, together with appropriate transformations to eliminate trends, are described in Appendix C.

Most of the core series are computed based on the raw indicators from Stock and Watson (2008) database and from the Fred-II database[8] maintained by the Federal Reserve Bank of St. Louis (database mnemonics are in italics). To obtain three measures of *real per-capita output*, we take real GDP (*SW2008::GDP251*), total industrial production (*SW2008::IPS10*) and industrial production in the manufacturing sector (*SW2008::IPS43*), and divide each series by the civilian non-institutional population (*Fred-II::CNP16OV*). We then take the natural logarithm and extract the linear trend by an OLS regression. The resulting detrended series are multiplied by 100 to convert them to percentage deviations from respective means. The *inflation* measures are computed as the first difference of the natural logarithm of the GDP deflator (*SW2008::GDP272A*), of the PCE deflator (*SW2008::GDP273A*), and of the Consumer Price Index – All Items (*SW2008::CPIAUCSL*), all multiplied by 400 to get to the annualized percentages. Our indicators of the *nominal interest rate* are (i) the effective federal funds rate (*SW2008::FYFF*), (ii) the 3-month U.S. Treasury bill rate in the secondary market (*SW2008::FYGM3*) and (iii) the yield on Moody's AAA-rated corporate bonds (*SW2008::FYAAAC*). We use a simple 3-month average to obtain quarterly annualized interest rates from monthly raw data.

To generate the appropriate *inverse money velocities*, we take three monetary aggregates: the sweep-adjusted money stock M1 (*CDJ::M1S*), the sweep-adjusted money

---

[8] The Fred-II database is available online at: http://research.stlouisfed.org/fred2/

stock M2 (*CDJ::M2S*) and the monetary base adjusted for changes in reserve requirements (*SW2008::FMFBA*). The sweep-adjusted stocks M1S and M2S are provided by Cynamon, Dutkowsky and Jones (2006)[9] and correct the distortionary impact (on the conventional measures M1 and M2) of the financial innovation that started in the early 1990s. These distortions take the form of underreporting of actual transactions balances and arise because of retail sweep programs and commercial demand deposit sweep programs, in which U.S. banks move a portion of funds from their customer demand deposits or other checkable deposits into instruments with zero reserve requirements. Since our DSGE model does not have any explicit open- economy context, we further adjust the monetary base FMFBA by deducting the amount of U.S. dollar currency held physically outside the United States.[10] We take M1S, M2S and the adjusted FMFBA, divide each series by the nominal GDP (*Fred-II::GDP*) to obtain the respective inverse velocities of money. For each IVM, we take the natural logarithm of the M/GDP ratio and scale it by 100. Finally, we remove the linear deterministic trend from the IVM based on M1S.

Because measurement equations (41) are modeled without intercepts, we estimate the data-rich DSGE model on a demeaned data set. Also, in line with standard practice in the factor literature, we standardize each time series so that its sample variance is equal to unity (however, we do not scale the core series when estimating the data-rich DSGE model).

---

[9] Sweep-adjusted money stocks are available online at: http://www.sweepmeasures.com.

[10] Federal Reserve Board: Flow of Funds Accounts of the United States: Z.1 Statistical Release for March 12, 2009 (available at http://www.federalreserve.gov/releases/z1/20090312/). Table L.204 “Checkable Deposits and Currency”, line 23 (Rest of the world: Currency), unique identifier: Z1/Z1/FL263025003.Q

5 Empirical Results

In this section, we conduct the empirical analysis of the regular and the data-rich DSGE model. We begin by discussing the choice of the prior distributions of model parameters and then describe the posterior estimates of deep structural parameters in both models. Second, we compare the estimated DSGE state variables from our data-rich and from the regular DSGE model. Finally, we explore the differences that the two models imply about the sources of business cycle fluctuations and about the propagation of structural innovations, notably the monetary policy and technology shocks, to the measures of real output, inflation, interest rates and the real money balances.

5.1 Priors

Since we estimate the regular DSGE model (130) and the data-rich DSGE model (40)-(42) using Bayesian techniques, we have to provide prior distributions for both models' parameters.

In our data-rich DSGE model, we have two groups of parameters: state-space model parameters comprising matrices $\mathbf{\Lambda}$, $\mathbf{\Psi}$ and $\mathbf{R}$, and deep structural parameters $\boldsymbol{\theta}$ of an underlying DSGE model. The prior for the state-space matrices is elicited differently for the core and the non-core data indicators contained in X_t. Let $\mathbf{\Lambda}_k$ and R_{kk} be the factor loadings and a variance of the measurement error innovation for the k^{th} measurement equation, $k = 1..J$.

Regarding the non-core measurement equations, the prior for $(\mathbf{\Lambda}_k, R_{kk})$ and for Ψ_{kk} is defined as follows. Similarly to Boivin and Giannoni (2006) and Kose, Otrok and

Whiteman (2008), we assume a joint Normal-InverseGamma prior distribution for $\left(\mathbf{\Lambda}_k, R_{kk}\right)$ so that $R_{kk} \sim IG_2(s_0, \nu_0)$ with location parameter $s_0 = 0.001$ and degrees of freedom $\nu_0 = 3$, and the prior mean of factor loadings is centered around the vector of zeros $\mathbf{\Lambda}_k \mid R_{kk} \sim \; N(\mathbf{\Lambda}_{k,0}, R_{kk}\mathbf{M}_0^{-1})$ with $\mathbf{\Lambda}_{k,0} = \mathbf{0}$ and $\mathbf{M}_0 = \mathbf{I}_N$. The prior for the k^{th} measurement equation's autocorrelation Ψ_{kk}, all k , is $N(0,1)$. We are making it perfectly tight, however, because there could be data series with stochastic trends we seek to capture with potentially highly persistent DSGE states-factors and not with highly persistent measurement errors. This implies that all measurement errors are *iid* mean-zero normal random variables.

In contrast, the prior distribution for the factor loadings in the core measurement equations follows the scheme explained in example (44). Instead of hypothetical "output" and "inflation" groups, we substitute four categories of the core series: real output, inflation, the nominal interest rate, and the inverse velocity of money, with three specific measures within each category, as described in the Data section. The joint prior distribution is still Normal-Inverse-Gamma $(\mathbf{\Lambda}_{k,0}, \mathbf{M}_0, s_0, \nu_o)$, but now, for each of the core series, the prior mean of the factor loadings $\mathbf{\Lambda}_{k,0}$ is centered at the regular-DSGE-model-implied factor loadings of a corresponding DSGE model variable (real output $\hat{Y}_t$, inflation $\hat{\pi}_t$, the nominal interest rate $\hat{R}_t$ or the inverse money velocity $\hat{\hat{M}}_t - \hat{Y}_t$), evaluated at the current draw of deep structural parameters $\mathbf{\theta}$. The covariance scaling matrix $\mathbf{M}_0$ is assumed diagonal $\mathbf{M}_0 = diag(\mathbf{\Omega}(\mathbf{\theta}))$, where $\mathbf{\Omega}(\mathbf{\theta})$ is the unconditional covariance matrix of the DSGE model state variables evaluated at a current draw of $\boldsymbol{\theta}$. $\mathbf{M}_0$ is the same across all core measurement equations. This choice implies that the prior will be tighter for the loadings on more volatile DSGE states. A similar approach is pursued in Schorfheide, Sill and Kryshko (2010) reproduced as Chapter 3 in this dissertation. The scale s_0 and degrees of freedom ν_0 are the same as for the parameters in the non-core measurement equations above. Finally, as argued in section 3.1, we use a degenerate prior for real GDP, GDP deflator inflation, the federal funds rate and the IVM based on the M2S monetary aggregate.

Our choice of prior distribution for the deep structural parameters of a DSGE model broadly follows Aruoba and Schorfheide (2009). We keep the same prior for the regular and for the data-rich DSGE models that we estimate below. A subset of these parameters that are *fixed* in estimation is reported in Table D1. We choose to have a logarithmic utility of household consumption by fixing $\gamma = 1$. We set the depreciation rate of capital δ to 0.014, which is the average quarterly ratio of the depreciation of fixed assets to the stock of these fixed assets in 1959-2005 (NIPA-FAT11 for stocks, NIPA-FAT13 for depreciation of fixed assets and consumer durables). The steady-state annualized inflation rate π_A is fixed at 2.5 percent – the average GDP deflator inflation in our sample. We implicitly impose the Fischer equation and let the steady-state annualized real interest rate r_A be equal to 2.84 percent. This value is obtained as the average federal funds interest rate in our sample minus π_A. Households' discount factor is therefore $\beta = 1/(1 + r_A/400)$.

We also introduce several normalizations. We normalize to 1 the steady-state real output Y_* and steady-state money demand shock χ_*. We use the average log inverse velocity of money (log[M2S/GDP]) in our sample to pin down $\log(\bar{M}_*/Y_*)$. Finally, as in Aruoba and Schorfheide (2009), we fix $\log(H_*/Y_*)$ to -3.5. This number is derived from the average inverse labor productivity in the data. In our sample, on average a worker produces roughly $33 of real GDP per hour. Hence, average H/Y in the data is $1/33$. From the average share of government spending (consumption plus investment) in nominal GDP, we calibrate g_* to be 1.2.

We also want our data-rich DSGE model to be broadly consistent – in terms of the conduct of monetary policy – with the other regular DSGE models estimated on post-1983 data. Therefore, we shut down "data-richness" for a moment and estimate our DSGE model on just three standard observables: real GDP, GDP deflator inflation and the federal funds rate. The resulting estimates of the Taylor (1993) rule coefficients were: $\psi_1 = 1.82$, $\psi_2 = 0.18$ and $\rho_R = 0.78$. In the estimation of the data-rich DSGE model, we set the policy rule coefficients to these values. This procedure is similar in spirit to Boivin and Giannoni (2006), who assume that the policy rate R_t is measured in the data by the federal funds rate without an error. This assumption guarantees that the estimated monetary policy rule coefficients will not drift far away from the conventional post-1983 values documented in the literature.

Despite detrending performed on all three measures of real per capita output, they are still highly persistent. To strike a balance between the observed output persistence and the need to have stationarity in the model, we fix the autocorrelation of the technology shock ρ_Z at 0.98. In the intermediate goods-producing sector, we further assume no fixed costs ($\tilde{F} = 0$) and the absence of static indexation for non-optimizing firms ($\pi_{**} = 1$).

The prior distributions for other parameters are summarized in Table D2. The prior for the steady-state related parameters represents the view that the capital share of α in a Cobb-Douglas production function of intermediate goods firms is about 0.3 and that the average markup these firms charge is about 15 percent. The prior for the Calvo (1983) probability ζ controlling nominal price rigidity is quite agnostic and spans the range of values consistent with fairly rigid and fairly flexible prices. As in Del Negro and Schorfheide (2008), the prior density for the price indexation parameter ι is close to uniform on a unit interval. Parameter ν_m controlling the interest-rate elasticity of money demand is a priori distributed according to a Gamma distribution with mean 20 and standard deviation 5. The existing literature (e.g., Aruoba, Schorfheide 2009, Levin, Onatsky, Williams and Williams 2005, and Christiano, Eichenbaum and Evans 2005) documents fairly large estimates of the money demand elasticity ranging from 10 to 25. The 90 percent interval for the investment adjustment cost parameter S'' spans values that Christiano, Eichenbaum, Evans (2005) find when matching DSGE and vector autoregression impulse response functions. The priors for the parameters determining the exogenous shock processes are taken from Aruoba and Schorfheide (2009). They reflect the belief that the money demand and government spending shocks are quite persistent.

5.2 Posteriors: Regular vs. Data-Rich DSGE Model

Using the Gibbs sampler with the Metropolis step outlined in section 3.1, we estimate the data-rich DSGE model. In addition, we have also estimated the regular DSGE model using standard Bayesian techniques (Random Walk Metropolis-Hastings algorithm, see An and Schorfheide, 2007). The underlying theoretical New Keynesian core is the same as in the data-rich DSGE model. The difference comes in the measurement equation (41): we keep only four core observable data series (real GDP, GDP deflator inflation, the federal funds interest rate and the inverse velocity of money based on the M2S aggregate), impose the factor loadings as in (130) and assume perfect measurement of all four model concepts (see the notes to Table D3, p.76).

The only parameters of direct interest here are the deep structural parameters $\boldsymbol{\theta}$ of an underlying DSGE model, and we report the posterior means and 90 percent credible intervals of these in the columns of Table D3. We find the capital share of output and the average price markup to be in line with estimates from regular – few observables, perfect measurement – DSGE estimation. We find little evidence of dynamic indexation by intermediate goods firms in both versions of the model. The implied average duration of nominal price contracts is about $1/(1-0.797)$ = 4.9 quarters. On the one hand, this is close to what Aruoba and Schorfheide (2009) find in their money-in-the-utility specification of a DSGE model and what Del Negro and Schorfheide (2008) document under the "standard" agnostic prior about nominal price rigidities (their Table 6, p. 1206). On the other hand, this is much higher than the price contracts duration of about 3 quarters found by Smets and Wouters (2007) and Schorfheide, Sill and Kryshko (2010).

In the context of a data-rich DSGE model similar to ours, Boivin and Giannoni's (2006) estimates imply that the firms change prices very slowly – on average once per at least 7 quarters. The 4.9 quarters found in the data-rich version is quite higher than the duration of price contracts documented for the regular DSGE model ($1/(1-0.759)=4.15$ quarters). The implication of this difference is that the implied slope of the New Keynesian Phillips curve[11] measuring the elasticity of current inflation to real marginal costs (and to real output) falls from 0.0745 to 0.0517 as we move from the perfect measurement, few observables to a richer data set in estimation of the same underlying DSGE model. This means, for example, that the cost of disinflation associated with achieving a 1 percent reduction in the rate of inflation at the expense of tolerating negative real output growth, as predicted by the data-rich DSGE model, turns out to be more sizable than the output cost of disinflation predicted by the traditional regular DSGE model.

As anticipated, we have obtained a fairly high elasticity of money demand. Our estimate of ν_m in the data-rich DSGE model case implies that a 100-basis-points increase in the interest rate leads to a 3.2 percent decline in real money balances. A very large estimate of the investment adjustment cost parameter (30.8 in data-rich versus 11.1 in the

[11] We say *implied* slope because our underlying theoretical DSGE model is linearized around positive steady-state inflation rate ($\pi_A = 2.5\%$) and assumes the absence of static price indexation by the non-optimizing intermediate goods firms ($\pi_{**} = 1$). This implies that we have a dynamic New Keynesian Phillips curve with additional lags of real marginal costs $\widehat{MC}_t$. In a more conventional model where the non-optimizing intermediate goods firms index their prices to the steady-state inflation rate ($\pi_{**} = \pi_* = = 1 + \pi_A/400$), the NK Phillps curve features only current marginal costs, the coefficient next to which γ_{mc} we report:

$$\hat{\pi}_t = \gamma_{\pi 1}\hat{\pi}_{t-1} + \gamma_{\pi 2}E_t(\hat{\pi}_{t+1}) + \gamma_{mc}\widehat{MC}_t$$

where $\gamma_{\pi 1} = \iota/(1+\beta\iota)$, $\gamma_{\pi 2} = \beta/(1+\beta\iota)$ and $\gamma_{mc} = (1-\zeta)(1-\zeta\beta)/(\zeta(1+\beta\iota))$.

regular DSGE model), as Aruoba and Schorfheide (2009) argue, has something to do with the need to reduce the volatility of the return to capital and to dampen its effect on marginal costs, which in turn affect current inflation through the New Keynesian Phillips curve relationship. This is reasonable given that in our data-rich DSGE model, the industrial production measures of real output are more volatile than the GDP-based measure, while the volatilities of inflation measures are fairly similar. In both models, the money demand shock χ_t turns out to be highly serially correlated, and the persistence of the government spending shock g_t is high as well, but more moderate. In the data-rich environment, this is hardly surprising, since these shocks are now the common factors for a large sub-panel of non-core informational series, many of which are fairly persistent.

5.3 Estimated States: Regular vs. Data-Rich DSGE Model

Our empirical analysis proceeds by plotting the estimated DSGE state variables from our data-rich DSGE model and from the regular DSGE model.

Figure D1 depicts the posterior means and 90 percent credible intervals of the estimated data-rich DSGE model states. These include three endogenous variables (model inflation $\hat{\pi}_t$, the nominal interest rate $\hat{R}_t$ and real household consumption $\hat{X}_t$) and three structural AR(1) shocks (government spending g_t, money demand χ_t and neutral technology Z_t). It is these states that are included in measurement equation (41) with potentially non-zero loadings. The figure depicts as well the smoothed versions of these same variables in a regular DSGE model estimation derived by Kalman smoother at posterior mean of the deep structural parameters.

Four observations stand out. First, all three structural disturbances exhibit large swings and prolonged deviations from zero capturing the persistent low-frequency movements in the data. Second, the estimated data-rich DSGE model states are much *smoother* than their counterparts in the regular DSGE model. The intuition is straightforward. In the data-rich context, the model states are the common components of a large panel of data, and they have to capture well not only a few core macro series (as is the case in the regular DSGE model), but also very many non-core informational series.

The third observation is that the money demand shock χ_t appears to be very different in the data-rich versus the regular DSGE model estimation. The underlying reason is that in the case of the regular DSGE model, it was mainly responsible for capturing the dynamics of the inverse money velocity based on M2S in the small 4-series data set. Once we allow for the rich panel of macro and financial observables, χ_t helps explain other series as well (for example, housing variables and non-GDP measures of real output – see Table D4), yet at the cost of the fit for the IVM_M2S. The fourth observation is a counterfactual behavior of government spending shock g_t and real consumption $\hat{X}_t$ during recessions: the former tends to fall and the latter to rise when times are bad. In reality, of course, it is the other way around: as a recession unfolds, real consumption falls and government purchases are usually intensified to mitigate the negative impact of the recession on aggregate demand. The estimated path of g_t would make sense, however, if we think of it as a general aggregate demand shock not specifically connected to government purchases. In spite of our DSGE model being able to track well the total output dynamics, it cannot properly discriminate the components, in particular $\hat{X}_t$. The solution would seem to be to enlarge the model by incorporating, say, an investment-specific technology shock a la Greenwood, Hercowitz and Krusell (1998) and to make the real consumption in the data one of the core observables, as for example is done in Smets and Wouters (2007) and Boivin and Giannoni (2006).

5.4 Sources of Business Cycle Fluctuations

Another dimension along which the data-rich DSGE model and the regular (few observables, perfect measurement) DSGE model differ relates to the sizes of estimated standard deviations of the exogenous shocks driving business cycle in our model economy. From inspecting Table D3, one can observe that all standard deviations (except for σ_R) are getting smaller when we move from the regular to the data-rich case. In part, this is due to the fact that in the data-rich DSGE model we allow for the measurement error (or the theoretical gap between a particular model concept and a data indicator) so that a portion of fluctuations in all observables is accounted for by this indicator-specific component. This conclusion is further confirmed by inspecting Figure D1 that depicts the posterior means and 90 percent credible intervals for all three shocks – which are a subset of the DSGE state variables. As the figure shows, the estimated shocks in the data-rich DSGE model case seem to have smaller amplitude of fluctuations and are much smoother than their regular DSGE model counterparts.

As the sizes and the estimated time paths of exogenous shocks vary, the two models are also telling us quite different stories about the sources of business cycle fluctuations. When we assume the one-to-one data indicator – model concept correspondence and the perfect measurement, the four structural shocks are required to explain all fluctuations in the small 4–variable data set containing one measure of the real output, inflation, interest rate and the inverse money velocity. As we allow for multiple indicators and for the indicator-specific measurement error (or the theretical gap) and go for a richer data set, the results (see Table D5) suggest that the importance of some structural shocks may have been overstated.

Table D5 presents the unconditional variance decomposition of the core macro series for the regular and the data-rich DSGE models. Two overall conclusions stand out. First, the estimated indicator-specific measurement errors/theoretical gaps seem to account for a significant share of fluctuations in the core macro series considered, ranging from 4 to 82 percent. Second, as we move from the regular to the data-rich DSGE model, the role of technology innovations in generating fluctuations in real output, inflation and the interest rates is noticeably reduced.

Beginning with the real output, the diminished role of TFP shocks is partially compensated by the higher importance of the government spending shocks ranging from 10 to 17 percent. The increased role of the money demand shocks accounting for about 30 percent of unconditional variance of industrial production (IP) and IP: Manufacturing suggests that the IP's behavior over the business cycle is markedly different from that of the real GDP. From 2 to 4 percent of fluctuations in the measures of real output are due to the monetary policy innovations, a modest increase from 1 percent found in the regular DSGE model.

For the various theoretically distinct measures of inflation, the reduced role of TFP shocks is documented mostly on account of the non-negligible (19-36 percent) contribution of the idiosyncratic-specific component. In part, the lower contribution of technology innovations is taken over by the money demand shocks: they explain 3.1 – 3.5 percent of fluctuations in the PCE deflator inflation and the CPI inflation as compared to zero in the regular – perfect measurement, few observables – case.

Looking at the variance decomposition for the interest rates, we observe that the share of technology shocks has fallen from 96 percent in the regular to 67-82 percent in the data-rich DSGE model. The importance of the indicator-specific measurement error (theoretical gap) components, though, remained quite low. At the same time, we document a much higher contribution of the monetary policy innovations in generating fluctuations in interest rates. In the regular case – when the interest rate was assumed to be perfectly measured just by the federal funds rate – the monetary policy shocks accounted for only 4 percent of the unconditional variance. Once we allow for several noisy indicators of the interest rates, the contribution of the monetary policy shocks has risen to 14-17 percent.

When we assumed that the inverse money velocity is properly measured in the data by the single series – the IVM based on M2S aggregate, the major drivers of its fluctuations over the business cycle were the money demand shocks (about 60 percent) and technological innovations (29 percent), with contribution of the monetary policy shocks being essentially zero. After we moved to a data-rich environment and added to the list two measures of the IVM – one based on M1S aggregate and another based on the adjusted money base – the picture has changed dramatically. The role of the shocks to money demand has fallen considerably to 3 percent (IVM_MBase), 6 percent (IVM_M2S) and 17 percent (IVM_M1S), whereas the contribution to the unconditional variance of technology shocks has increased to 40 percent, though only for the inverse velocities based on M1 and monetary base. For the IVM_M2S, it is the indicator-specific "measurement error" that has become the major driver of fluctuations (82 percent) suggesting that our theoretical DSGE model captures the comovements in the real output and M2S balances quite poorly and is probably misspecified along this dimension. As expected, the results reveal a much greater role (10 percent) of the monetary policy in generating fluctuations of the IVM based on monetary base. This makes perfect sense given that the monetary base is the most fluid aggregate and is more interest-rate-sensitive than M1 and M2 aggregates.

5.5 Impulse Response Analysis

One of the key appealing features of DSGE models is that the researchers and policymakers can use modern macroeconomic theory to interpret and predict the comovement of aggregate macro time series over the business cycle. Therefore, in this subsection we focus on propagating all structural innovations (government spending, money demand, monetary policy and technology) in both the regular DSGE model and the data-rich DSGE model with a view to generate and compare the predictions for the key macroeconomic observables. By construction, in the regular DSGE model we are limited to obtain these predictions only for four primary series – real GDP, GDP deflator inflation, fed funds rate and real M2S, assumed to measure perfectly the corresponding model concepts. In the data-rich DSGE model, though, we could trace the dynamic effects of the same shocks to additional data indicators measuring real output, inflation, interest rates and real money balances. We defer the discussion of the impact of structural shocks on the non-core data variables in the data-rich DSGE model to Chapter 2.

In Figure D2, we present the impulse response functions (IRFs) of the four primary macro observables: real GDP, GDP deflator inflation, fed funds rate and real M2S – to four one-standard-deviation structural shocks. A positive 1-std *government spending innovation* is associated with 60 to 80 basis points increase in real GDP on impact. Since the government finances its additional purchases through borrowing in the open market, it diverts part of the resources and partially "crowds out" private consumption and investment. Heavier borrowing raises nominal short-term interest rate by 2 to 5.5 basis points (b.p.) and inhibits private investment even more, which in turn leads to declining return on capital and lower marginal costs. The latter explains the negative effect (15-30 b.p.) of g_t on GDP deflator inflation that we observe on impact. Finally, high interest rates raise the opportunity cost of holding money and households reduce their real money balances. As can be seen from Figure D2, the regular DSGE model clearly overstates the expansionary impact of government spending on real GDP by about 20 basis points and also overestimates the negative effect on GDP deflator inflation by 15 basis points (which is twice as the size of the effect in the data-rich DSGE model). At the same time, the impact of crowding out on the fed funds rate is clearly understated: the data-rich DSGE model predicts 5.5 b.p. increase at the 5th-quarter peak, while the regular DSGE model yields only 2 b.p. increase peaking in 2 years after the initial shock.

The 2nd row of Figure D2 depicts the IRFs to the *money demand innovation*. It should be noted that in our theoretical New Keynesian model the money term enters the equilibrium conditions only in single place – in money demand equation (85). And the central bank is always assumed to supply enough money balances to satisfy all demand from households given current nominal interest rate. Because of that, the money balances are block exogenous and the money demand shocks – while raising or lowering M_t – do not affect either real output, or inflation or the interest rate in equilibrium. This is exactly true for the regular DSGE model, IRFs of which show positive response of the real M2S to one-std money demand shock and zero response of all other variables. This is approximately true in the data-rich DSGE model, but only for the four primary observables shown. The IRFs for the other noisy measures of real output, inflation, interest rate and real money balances (not shown) are non-zero and generally follow the patterns depicted by the thick blue line, though on a higher-scale grid: a positive money demand innovation raises real output contemporaneously, dampens prices and leads to the standard liquidity effect (lower interest rates associated with higher real money balances). The regular DSGE model differs from the data-rich one in that the former seems to overstate by a wide margin (roughly 45 basis points) the contemporaneous positive effect of the elevated money demand on real M2S.

Let us now turn to the effects of *monetary policy innovation*, which are summarized in the 3rd row of Figure D2 and in Figure D3. A contractionary monetary policy shock corresponds to 60 (regular) – 75 (data-rich) basis points increase in the federal funds rate. Both versions of the DSGE models predict that the real GDP and the GDP deflator inflation will fall by 40-50 b.p. and 25-30 b.p., respectively, before returning to their trend paths. As the nominal policy rate rises and the opportunity costs of holding money for households increase, we observe a strong liquidity effect associated with falling real money balances (50 b.p. in the regular and 72 b.p. in the data-rich DSGE model). Also, high interest rates make the saving motive and buying more bonds temporarily a more attractive option. This raises households' marginal utility of consumption and discourages current spending in favor of the future consumption. Because the household faces investment adjustment costs and cannot adjust investment quickly, and government spending in the model is exogenous, the lower consumption leads to a fall in aggregate demand. The firms respond to lower demand in part by contracting real output and in part by reducing the optimal price. Hence, the aggregate price level falls, but not as much given nominal rigidities in the intermediate goods-producing sector. Notice that despite some on-impact differences, the responses of all variables to the monetary policy innovation remain very similar and quantitatively close in the regular and the data-rich DSGE models.

The real challenge is revealed in Figure D3. The IRFs of the other measures of the real output and inflation to the monetary policy innovation produce puzzling results. For example, industrial production: total and industrial production: manufacturing actually rise following a contractionary monetary policy shock, at least on impact. By the same token, the PCE deflator inflation and CPI inflation react positively to monetary tightening, despite GDP deflator inflation – the primary inflation measure – responding negatively as prescribed by theory. We discuss further the potential reasons for that and show how to deal with these puzzling results in Chapter 2. For now, we would just like to note that these puzzles may indicate the potential misspecification of our DSGE model.

We plot the effects of a positive *technology innovation* in row 4 of Figure D2 and in Figure D4 (other core series). Following positive TFP shock, the real GDP broadly increases, as our economy becomes more productive and the firms find it optimal to produce more. Both models generate the hump-shaped positive IRFs; the regular DSGE model predicts that the maximal impact on real GDP of 75 basis points is achieved at the 14^{th}-quarter peak, while the data-rich DSGE model's response is more persistent, but is twice as low and peaks roughly at the 23^{rd} quarter. New demand come primarily from higher capital investment, reflecting much better future return on capital, and also from additional household consumption fueled by greater income. The higher output on the supply side plus improved efficiency implies a downward pressure on prices (GDP deflator inflation falls by 52 basis points in the data-rich versus 90 basis points in the regular DSGE model). The increase in real GDP above steady state and the fall of inflation below target level, under the estimated monetary policy Taylor rule, requires the Fed to move the policy rate in opposite directions. The fact that the Fed actually lowers the policy rate means that the falling prices effect dominates. Declining interest rate boosts real output even more, which in turn raises further the return on capital. As the positive impact of technological innovation dissipates, this higher return, through the future marginal costs channel, fuels inflationary expectations that ultimately translate into contemporaneous upward price pressures. The Fed reacts by increasing the policy rate, which explains the observed hump in the fed funds interest rate IRF. Given temporarily lower interest rates, households choose to hold, with some lag, relatively higher real money balances. A part of the growing money demand comes endogenously from the elevated level of economic activity. A general observation from comparing the IRFs from the regular and the data-rich DSGE models is that the regular DSGE model tends to overestimate all effects of TFP shocks, though on impact they might not be too different.

Looking at the responses of the alternative measures of real output, inflation, interest rates and real money balances to the positive TFP shock (Figure D4), we generally conclude that they remain qualitatively similar to the reactions of primary data indicators and we don't observe puzzles as documented above for the effects of monetary tightening. The measures of industrial production tend to rise, although more slowly than the real GDP, the price inflations tend to fall though the magnitude of the on-impact effect is twice as low. The 3-month T-bill rate and the AAA bond yield broadly follow the path of the federal funds rate, with bond yield falling slower and lagging roughly 4 quarters. The measures of real money balances respond by and large positively and with a hump, yet the initial responses of the real M1S and the real monetary base remain negative for two quarters in a row.

6 Conclusions

In a growing body of literature that estimates macroeconomic DSGE models, two assumptions remain very common: (i) that a particular model concept is perfectly measured by a single data series without an error, and (ii) that all relevant information to estimate the state and the parameters of the economy is summarized by a few observable data indicators, usually equal to the number of structural shocks in the model. In this chapter, we relaxed these two assumptions and estimated a version of the monetary DSGE model with standard New Keynesian core on a richer data set. This so called *data-rich DSGE model* can be seen as a combination of a regular DSGE model and a dynamic factor model in which factors are the economic state variables of the DSGE model and the transition of factors is governed by a DSGE model solution.

We used the post-1983 U.S. data on real output, inflation, nominal interest rates, measures of inverse money velocity and a large panel of the other informational macroeconomic and financial series to estimate and compare the new data-rich DSGE model with a regular – few observables, perfect measurement – DSGE model, both sharing the same theoretical core. The estimation involved Bayesian MCMC methods. Because of the data set's high panel dimension, the likelihood-based estimation of the data-rich DSGE model was computationally very challenging. To reduce the costs, we employed a novel speedup as in Jungbacker and Koopman (2008) and achieved the computational time savings of 60 percent.

We documented that the data-rich DSGE model generates a higher duration of the Calvo price contracts and a lower implied slope of the New Keynesian Phillips curve measuring the elasticity of current inflation to real marginal costs. As we moved from the regular to the data-rich DSGE model, we found that: (i) the role of technology innovations in generating fluctuations in real output, inflation and the interest rates is noticeably reduced; and that (ii) the contribution of monetary policy shocks to cyclical fluctuations of the interest rates increased from 4 to 14-17 percent. Regarding dynamic propagation, we established that (i) despite some slight on-impact differences, the responses of all primary observables (real GDP, GDP deflator inflation, fed funds rate and real M2) to the monetary policy innovation remain theoretically plausible and quantitatively close in the regular and in the data-rich DSGE models; and that (ii) the regular DSGE model tended to overestimate all effects of TFP shocks, though on impact they might not have been too different. Finally, we found some puzzling results for the responses of the industrial production, the PCE deflator inflation and the CPI inflation to monetary tightening, which may indicate the potential misspecification of our theoretical DSGE model. We plan to address and discuss these issues and puzzles further in Chapter 2.

Appendix A. DSGE Model

Appendix A1. First-Order Conditions of Household

The problem of each household j is to maximize the utility function (9) subject to budget constraint (11) and capital accumulation equation (12) for all t. Associate Lagrange multipliers $\lambda_t(j)$ and $Q_t(j)$ with constraints (11) and (12), respectively. Then, the First Order Conditions with respect to $x_t(j)$, $h_t(j)$, $m_{t+1}(j)$, $i_t(j)$, $k_{t+1}(j)$ and $b_{t+1}(j)$ are:

$$\lambda_t(j) = \frac{U'(x_t(j))}{P_t} \tag{59}$$

$$\lambda_t(j) = \frac{A}{P_t W_t} \tag{60}$$

$$\frac{U'(x_t(j))}{P_t} = \beta E_t \left\{ \frac{\chi_{t+1}}{P_{t+1}} \left(\frac{A}{Z_*^{1/(1-\alpha)}} \right)^{(1-\nu_m)} \left(\frac{m_{t+1}(j)}{P_{t+1}} \right)^{-\nu_m} + \frac{U'(x_{t+1}(j))}{P_{t+1}} \right\} \tag{61}$$

$$\begin{aligned} 1 = \mu_t(j) \left[1 - S\left(\frac{i_t(j)}{i_{t-1}(j)} \right) - S'\left(\frac{i_t(j)}{i_{t-1}(j)} \right) \frac{i_t(j)}{i_{t-1}(j)} \right] + \\ + \beta E_t \left\{ \mu_{t+1}(j) \frac{U'(x_{t+1}(j))}{U'(x_t(j))} S'\left(\frac{i_{t+1}(j)}{i_t(j)} \right) \left[\frac{i_{t+1}(j)}{i_t(j)} \right]^2 \right\} \end{aligned} \tag{62}$$

$$\mu_t(j) = \beta E_t \left\{ \frac{U'(x_{t+1}(j))}{U'(x_t(j))} \left(R^k_{t+1} + \mu_{t+1}(j)(1-\delta) \right) \right\} \tag{63}$$

$$1 = \beta E_t \left\{ \frac{U'(x_{t+1}(j))}{U'(x_t(j))} \frac{R_t}{\pi_{t+1}} \right\}, \tag{64}$$

where $\pi_t = P_t / P_{t-1}$ denotes inflation and where we have substituted out Lagrange multiplier $\lambda_t(j)$ with its equivalent expression using marginal utility of consumption and have introduced the normalized shadow price of installed capital $\mu_t(j) = \dfrac{Q_t(j)}{U'(x_t(j))}$.

We do not take first order conditions with respect to A-D securities holdings $a_{t+1}(j)$ explicitly, because we make use of the result due to Erceg, Henderson and Levin (2000). This result says that under the assumption of complete markets for A-D securities and under the additive separability of labor and money balances in household's utility, the equilibrium price of A-D securities will be such that optimal consumption will not depend on idiosyncratic shocks. Hence, all households will share the same marginal utility of consumption, and given (59), Lagrange multiplier $\lambda_t(j)$ will also be the same across all households: $\lambda_t(j) = \lambda_t$, all j and t. This implies that in equilibrium all households will choose the same consumption, money and bond holdings, investment and capital. Note that we don't have wage rigidity in this model – therefore the choice of optimal labor will also be same. This implies that we can safely drop index j from all equilibrium conditions of households and proceed accordingly.

The first two FOCs could be combined to yield labor supply equation relating real wage to marginal rate of substitution between consumption and labor. (61) is an Euler equation for money holdings, which together with (64) – an Euler equation for bond holdings – implies household's optimal demand for real money balances. Equation (62) determines the law of motion for shadow price of installed capital. If there were no investment adjustment costs, this price will be equal to 1, which is standard in neoclassical growth model. Also note that if we were to have an investment specific technology shock, this shadow price will be equal to relative price of capital in consumption units. Equation (63) is an Euler equation for capital holdings. The shadow cost of purchasing one unit of capital today should be equal to the real return from renting it to firms plus the tomorrow's resale value of capital that has not yet depreciated.

Appendix A2. First-Order Conditions of Intermediate Goods Firm

Monopolistically competitive intermediate goods producer i, which is allowed to re-optimize, chooses the optimal price $P_t^o(i)$ that maximizes discounted stream of profits subject to optimal demand from final good producers:

$$\begin{aligned} &\max_{P_t^o(i)} \quad E_t\left\{\sum_{s=0}^{\infty}(\zeta\beta)^s\,\Xi_{t+s|t}^{p}\,(P_t^o(i)\pi_{t+s|t}^{adj}-P_{t+s}MC_{t+s})Y_{t+s}(i)\right\} \\ &\text{s.t.} \quad Y_{t+s}(i)=\left[\frac{P_t^o(i)\pi_{t+s|t}^{adj}}{P_{t+s}}\right]^{-\frac{(1+\lambda)}{\lambda}} Y_{t+s}, \qquad s=0,1,2,.... \end{aligned} \tag{65}$$

First, obtain an expression for $\dfrac{\partial Y_{t+s}(i)}{\partial P_t^o(i)}$:

$$\frac{\partial Y_{t+s}(i)}{\partial P_t^o(i)}=-\frac{(1+\lambda)}{\lambda}\left[\frac{P_t^o(i)\pi_{t+s|t}^{adj}}{P_{t+s}}\right]^{-\frac{(1+\lambda)}{\lambda}-1}\frac{\pi_{t+s|t}^{adj}}{P_{t+s}}Y_{t+s}=-\left(\frac{1+\lambda}{\lambda}\right)\frac{Y_{t+s}(i)}{P_t^o(i)}. \tag{66}$$

Now the first order condition for the problem (65), where we will plug optimal demand $Y_{t+s}(i)$ into the objective function and assume interior solution, is:

$$E_t\left\{\sum_{s=0}^{\infty}(\zeta\beta)^s\,\Xi_{t+s|t}^{p}\left[\pi_{t+s|t}^{adj}\left(Y_{t+s}(i)+P_t^o(i)\frac{\partial Y_{t+s}(i)}{\partial P_t^o(i)}\right)-P_{t+s}MC_{t+s}\frac{\partial Y_{t+s}(i)}{\partial P_t^o(i)}\right]\right\}=0. \tag{67}$$

Consider expression inside square brackets:

$$
\begin{aligned}
\left[\cdots\right] \quad &= \pi^{adj}_{t+s|t} Y_{t+s}(i) + \left(P^o_t(i)\pi^{adj}_{t+s|t} - P_{t+s}MC_{t+s} \right) \frac{\partial Y_{t+s}(i)}{\partial P^o_t(i)} = \\
&= \pi^{adj}_{t+s|t} Y_{t+s}(i) - \left(P^o_t(i)\pi^{adj}_{t+s|t} - P_{t+s}MC_{t+s} \right) \frac{(1+\lambda)}{\lambda} \frac{Y_{t+s}(i)}{P^o_t(i)} = \\
&= \frac{Y_{t+s}(i)}{P^o_t(i)} \left(P^o_t(i)\pi^{adj}_{t+s|t} - \frac{(1+\lambda)}{\lambda} \left(P^o_t(i)\pi^{adj}_{t+s|t} - P_{t+s}MC_{t+s} \right) \right) = \\
&= \frac{1}{\lambda} \frac{Y_{t+s}(i)}{P^o_t(i)} \left((\lambda - 1 - \lambda) P^o_t(i)\pi^{adj}_{t+s|t} + (1+\lambda) P_{t+s}MC_{t+s} \right).
\end{aligned}
$$

Cancelling out $1/\lambda \neq 0$ and multiplying (67) by -1, we could rewrite the FOC as follows:

$$
E_t \left\{ \sum_{s=0}^{\infty} (\zeta\beta)^s \Xi^p_{t+s|t} \frac{Y_{t+s}(i)}{P^o_t(i)} \left[P^o_t(i)\pi^{adj}_{t+s|t} - (1+\lambda) P_{t+s}MC_{t+s} \right] \right\} = 0. \qquad (68)
$$

Remark 1: Since for $s > 0$, $\pi^{adj}_{t+s|t} = \prod_{l=1}^{s} \pi^{\iota}_{t+l-1} \pi^{(1-\iota)}_{**}$ and $\pi^{adj}_{(t+1)+s|(t+1)} = \prod_{l=1}^{s} \pi^{\iota}_{(t+1)+l-1} \pi^{(1-\iota)}_{**} =$

$= \pi^{\iota}_{t+1}\pi^{\iota}_{t+2}...\pi^{\iota}_{t+s}\pi^{(1-\iota)s}_{**}$, it follows that $\pi^{adj}_{t+(s+1)|t} = \prod_{l=1}^{s+1} \pi^{\iota}_{t+l-1}\pi^{(1-\iota)}_{**} = \pi^{\iota}_t \pi^{\iota}_{t+1}\pi^{\iota}_{t+2}...\pi^{\iota}_{t+s}\pi^{(1-\iota)(s+1)}_{**} =$

$= \left[\pi^{\iota}_t \pi^{(1-\iota)}_{**} \right] \pi^{adj}_{(t+1)+s|(t+1)}$.

Remark 2: Since for $s > 0$, $\Xi^p_{t+s|t} = \frac{\lambda_{t+s}}{\lambda_t}$ and so $\Xi^p_{t+(s+1)|t} = \frac{\lambda_{t+s+1}}{\lambda_t}$, it follows that

$\Xi^p_{(t+1)+s|(t+1)} = \frac{\lambda_{t+1+s}}{\lambda_{t+1}} \frac{\lambda_t}{\lambda_t} = \; = \Xi^p_{t+(s+1)|t} \big/ \Xi^p_{t+1|t}$ and that $\Xi^p_{t+(s+1)|t} = \Xi^p_{t+1|t} \Xi^p_{(t+1)+s|(t+1)}$.

Remark 3: Notice that given expression for an optimal demand for good i in (65), $Y_{t+(s+1)}(i) \neq Y_{(t+1)+s}(i)$. However, using result from Remark 1, we obtain:

$$Y_{t+(s+1)}(i) = \left[\frac{P^o_{t+1}(i)}{P^o_{t+1}(i)} \frac{P^o_t(i)\pi^{adj}_{t+(s+1)|t}}{P_{t+(s+1)}} \right]^{-\frac{(1+\lambda)}{\lambda}} Y_{t+(s+1)} =$$

$$= \left[\frac{P^o_t(i)}{P^o_{t+1}(i)} \right]^{-\frac{(1+\lambda)}{\lambda}} \left[\frac{P^o_{t+1}(i)}{P_{(t+1)+s}} \left[\pi^{\iota}_t \pi^{(1-\iota)}_{**} \right] \pi^{adj}_{(t+1)+s|(t+1)} \right]^{-\frac{(1+\lambda)}{\lambda}} Y_{t+1+s} =$$

$$= \left[\frac{P^o_t(i)}{P^o_{t+1}(i)} \right]^{-\frac{(1+\lambda)}{\lambda}} \left[\pi^{\iota}_t \pi^{(1-\iota)}_{**} \right]^{-\frac{(1+\lambda)}{\lambda}} Y_{(t+1)+s}(i)$$

To express FOC (68) recursively, we define two auxiliary variables:

$$f^{(1)}_t = E_t \left\{ \sum_{s=0}^{\infty} (\zeta\beta)^s \Xi^p_{t+s|t} Y_{t+s}(i) \pi^{adj}_{t+s|t} \right\} \tag{69}$$

$$f^{(2)}_t = E_t \left\{ \sum_{s=0}^{\infty} (\zeta\beta)^s \Xi^p_{t+s|t} Y_{t+s}(i) \frac{P_{t+s}}{P^o_t(i)} MC_{t+s} \right\}, \tag{70}$$

so that FOC becomes:

$$f^{(1)}_t = (1+\lambda) f^{(2)}_t. \tag{71}$$

Recalling that $\Xi^p_{t|t} = 1$, $\pi^{adj}_{t|t} = 1$ and using results from Remarks 1, 2 and 3, we can rewrite (69) as:

$$f^{(1)}_t = Y_t(i) + E_t \left\{ \sum_{k=0}^{\infty} (\zeta\beta)^{k+1} \Xi^p_{t+(k+1)|t} Y_{t+(k+1)}(i) \pi^{adj}_{t+(k+1)|t} \right\} =$$

$$= Y_t(i) + (\zeta\beta) E_t \left\{ \sum_{k=0}^{\infty} (\zeta\beta)^k (\Xi^p_{t+1|t} \Xi^p_{(t+1)+k|(t+1)}) \left[\frac{P^o_t(i)}{P^o_{t+1}(i)} \right]^{-\frac{(1+\lambda)}{\lambda}} \times \right.$$

$$\left. \times \left[\pi^{\iota}_t \pi^{(1-\iota)}_{**} \right]^{-\frac{(1+\lambda)}{\lambda}} Y_{(t+1)+k}(i) \left[\pi^{\iota}_t \pi^{(1-\iota)}_{**} \right] \pi^{adj}_{(t+1)+k|(t+1)} \right\} =$$

$$= Y_t(i) + \zeta\beta \left[\pi^{\iota}_t \pi^{(1-\iota)}_{**} \right]^{-\frac{1}{\lambda}} E_t \left\{ \left[\frac{P^o_t(i)}{P^o_{t+1}(i)} \right]^{-\frac{(1+\lambda)}{\lambda}} \Xi^p_{t+1|t} \sum_{k=0}^{\infty} (\zeta\beta)^k \Xi^p_{(t+1)+k|(t+1)} Y_{(t+1)+k}(i) \pi^{adj}_{(t+1)+k|(t+1)} \right\} =$$

$$=\left[\frac{P_t^o(i)}{P_t}\right]^{-\frac{(1+\lambda)}{\lambda}} Y_t+\zeta\beta\left[\pi_t^{\iota}\pi_{**}^{(1-\iota)}\right]^{-\frac{1}{\lambda}} E_t\left\{\left[\frac{P_t^o(i)}{P_{t+1}^o(i)}\right]^{-\frac{(1+\lambda)}{\lambda}} \Xi_{t+1|t}^p f_{t+1}^{(1)}\right\}. \tag{72}$$

Similarly, the recursion for $f_t^{(2)}$ becomes:

$$f_t^{(2)}=Y_t(i)\frac{P_t MC_t}{P_t^o(i)}+(\zeta\beta)E_t\left\{\sum_{k=0}^{\infty}(\zeta\beta)^k(\Xi_{t+1|t}^p\Xi_{(t+1)+k|(t+1)}^p)\left[\frac{P_t^o(i)}{P_{t+1}^o(i)}\right]^{-\frac{(1+\lambda)}{\lambda}}\times\right.$$

$$\left.\times\left[\pi_t^{\iota}\pi_{**}^{(1-\iota)}\right]^{-\frac{(1+\lambda)}{\lambda}} Y_{(t+1)+k}(i)\frac{P_{t+1+k}}{P_t^o(i)}MC_{t+1+k}\frac{P_{t+1}^o(i)}{P_{t+1}^o(i)}\right\}=$$

$$=\left[\frac{P_t^o(i)}{P_t}\right]^{-\frac{(1+\lambda)}{\lambda}-1} MC_tY_t+\zeta\beta\left[\pi_t^{\iota}\pi_{**}^{(1-\iota)}\right]^{-\frac{(1+\lambda)}{\lambda}} E_t\left\{\left[\frac{P_t^o(i)}{P_{t+1}^o(i)}\right]^{-\frac{(1+\lambda)}{\lambda}-1} \Xi_{t+1|t}^p f_{t+1}^{(2)}\right\}. \tag{73}$$

In summary, the first order conditions of the problem (27) boil down to these three equations:

$$f_t^{(1)}=\left(p_t^o\right)^{-\frac{(1+\lambda)}{\lambda}} Y_t+\zeta\beta\left(\pi_t^{\iota}\pi_{**}^{(1-\iota)}\right)^{-\frac{1}{\lambda}} E_t\left\{\left(\frac{p_t^o}{p_{t+1}^o\pi_{t+1}}\right)^{-\frac{(1+\lambda)}{\lambda}} \Xi_{t+1|t}^p f_{t+1}^{(1)}\right\} \tag{74}$$

$$f_t^{(2)}=\left(p_t^o\right)^{-\frac{(1+\lambda)}{\lambda}-1} MC_tY_t+\zeta\beta\left(\pi_t^{\iota}\pi_{**}^{(1-\iota)}\right)^{-\frac{(1+\lambda)}{\lambda}} E_t\left\{\left(\frac{p_t^o}{p_{t+1}^o\pi_{t+1}}\right)^{-\frac{(1+\lambda)}{\lambda}-1} \Xi_{t+1|t}^p f_{t+1}^{(2)}\right\} \tag{75}$$

$$f_t^{(1)}=(1+\lambda)f_t^{(2)}, \tag{76}$$

where we have defined the optimal price relative to the price level $p_t^o=\frac{P_t^o}{P_t}$ and $\pi_t=\frac{P_t}{P_{t-1}}$.

Appendix A3. Evolution of Price Dispersion

Aggregate price dispersion across intermediate goods firms is captured by variable $D_t = \int_0^1 \left(\frac{P_t(i)}{P_t} \right)^{-\frac{(1+\lambda)}{\lambda}} di$. By properties of Calvo pricing, $P_t(i)$ is equal to optimal price P_t^o with probability $1-\zeta$ (optimizing firms) and is equal to $\left[\pi_{t-1}^{\iota} \pi_{**}^{(1-\iota)} \right] P_{t-1}(i)$ with probability ζ (non-optimizing firms). Therefore, by definition of D_t we have:

$$D_t = \int_0^1 \left(\frac{P_t(i)}{P_t} \right)^{-\frac{(1+\lambda)}{\lambda}} di = (1-\zeta) \left(\frac{P_t^o}{P_t} \right)^{-\frac{(1+\lambda)}{\lambda}} + \zeta \left[\pi_{t-1}^{\iota} \pi_{**}^{(1-\iota)} \right]^{-\frac{(1+\lambda)}{\lambda}} \int_0^1 \left(\frac{P_{t-1}(i)}{P_t} \right)^{-\frac{(1+\lambda)}{\lambda}} di =$$

$$= (1-\zeta) \left(\frac{P_t^o}{P_t} \right)^{-\frac{(1+\lambda)}{\lambda}} + \zeta \left[\pi_{t-1}^{\iota} \pi_{**}^{(1-\iota)} \right]^{-\frac{(1+\lambda)}{\lambda}} \left(\frac{P_{t-1}}{P_t} \right)^{-\frac{(1+\lambda)}{\lambda}} \int_0^1 \left(\frac{P_{t-1}(i)}{P_{t-1}} \right)^{-\frac{(1+\lambda)}{\lambda}} di$$

The last line implies:

$$D_t = (1-\zeta) \left(\frac{P_t^o}{P_t} \right)^{-\frac{(1+\lambda)}{\lambda}} + \zeta \left[\left(\frac{\pi_{t-1}}{\pi_t} \right)^{\iota} \left(\frac{\pi_{**}}{\pi_t} \right)^{(1-\iota)} \right]^{-\frac{(1+\lambda)}{\lambda}} D_{t-1} \tag{77}$$

Appendix A4. Equilibrium Conditions and Aggregate Disturbances

We define equilibrium in our economy in a standard way. It is determined by the optimality conditions and laws of motion summarized below:

(1) Households' optimality conditions

$$U'(x_t) = \frac{A}{W_t} \tag{78}$$

$$\frac{U'(x_t)}{P_t} = \beta E_t \left\{ \frac{\chi_{t+1}}{P_{t+1}} \left(\frac{A}{Z_*^{1/(1-\alpha)}} \right)^{(1-\nu_m)} \left(\frac{m_{t+1}}{P_{t+1}} \right)^{-\nu_m} + \frac{U'(x_{t+1})}{P_{t+1}} \right\} \tag{79}$$

$$1=\mu_t\left[1-S\left(\frac{i_t}{i_{t-1}}\right)-S'\left(\frac{i_t}{i_{t-1}}\right)\frac{i_t}{i_{t-1}}\right]+\beta E_t\left\{\mu_{t+1}\frac{U'(x_{t+1})}{U'(x_t)}S'\left(\frac{i_{t+1}}{i_t}\right)\left[\frac{i_{t+1}}{i_t}\right]^2\right\} \quad (80)$$

$$\mu_t=\beta E_t\left\{\frac{U'(x_{t+1})}{U'(x_t)}\left(R_{t+1}^k+\mu_{t+1}(1-\delta)\right)\right\} \quad (81)$$

$$1=\beta E_t\left\{\frac{U'(x_{t+1})}{U'(x_t)}\frac{R_t}{\pi_{t+1}}\right\} \quad (82)$$

$$k_{t+1}=(1-\delta)k_t+\left[1-S\left(\frac{i_t}{i_{t-1}}\right)\right]i_t \quad (83)$$

$$\Xi_{t+1|t}^p=\frac{U'(x_{t+1})}{U'(x_t)}\frac{1}{\pi_{t+1}} \quad (84)$$

Note that (79) and (82) imply money demand equation[12]:

$$\left(\bar{M}_t\right)^{\nu_m}=\left(\frac{m_{t+1}}{P_t}\right)^{\nu_m}=\frac{\beta R_t}{U'(x_t)(R_t-1)}E_t\left\{\left(\frac{A}{Z_*^{1/(1-\alpha)}}\right)^{(1-\nu_m)}\frac{\chi_{t+1}}{\pi_{t+1}^{(1-\nu_m)}}\right\}. \quad (85)$$

(2) <u>Firms' optimality conditions</u>

$$K_t=\frac{\alpha}{1-\alpha}\frac{W_t}{R_t^k}H_t \quad (86)$$

$$MC_t=\left(\frac{1}{\alpha}\right)^{\alpha}\left(\frac{1}{1-\alpha}\right)^{(1-\alpha)}\frac{W_t^{1-\alpha}\left(R_t^k\right)^{\alpha}}{Z_t} \quad (87)$$

$$f_t^{(1)}=\left(p_t^o\right)^{-\frac{(1+\lambda)}{\lambda}}Y_t+\zeta\beta\left(\pi_t^{\iota}\pi_{**}^{(1-\iota)}\right)^{-\frac{1}{\lambda}}E_t\left\{\left(\frac{p_t^o}{p_{t+1}^o\pi_{t+1}}\right)^{-\frac{(1+\lambda)}{\lambda}}\Xi_{t+1|t}^p f_{t+1}^{(1)}\right\} \quad (88)$$

[12] We deflate nominal money stock m_{t+1} by P_t (and not P_{t+1}) since it has been chosen in period *t* based on realization of period *t* disturbances. We denote corresponding real money balances by $\bar{M}_{t+1}=m_{t+1}/P_t$.

$$f_t^{(2)} = \left(p_t^o\right)^{-\frac{(1+\lambda)}{\lambda}-1} MC_t Y_t + \zeta\beta\left(\pi_t^{\iota}\pi_{**}^{(1-\iota)}\right)^{-\frac{(1+\lambda)}{\lambda}} E_t\left\{\left(\frac{p_t^o}{p_{t+1}^o\pi_{t+1}}\right)^{-\frac{(1+\lambda)}{\lambda}-1} \Xi_{t+1|t}^p f_{t+1}^{(2)}\right\} \quad (89)$$

$$f_t^{(1)} = (1+\lambda) f_t^{(2)} \quad (90)$$

$$\pi_t = \left[(1-\zeta)\left(\pi_t p_t^o\right)^{-\frac{1}{\lambda}} + \zeta\left(\pi_{t-1}^{\iota}\pi_{**}^{(1-\iota)}\right)^{-\frac{1}{\lambda}}\right]^{-\lambda}, \quad (91)$$

where we have denoted $p_t^o = P_t^o / P_t$ and where equilibrium requires $K_t = k_t$, $H_t = h_t$.

(3) Taylor rule

$$\frac{R_t}{R_*} = \left(\frac{R_{t-1}}{R_*}\right)^{\rho_R}\left(\left(\frac{\pi_t}{\pi_*}\right)^{\psi_1}\left(\frac{Y_t}{Y_*}\right)^{\psi_2}\right)^{(1-\rho_R)} e^{\varepsilon_{R,t}}, \qquad \text{where } \varepsilon_{R,t} \sim N(0,\sigma_R^2) \quad (92)$$

(4) Aggregate demand and supply

$$X_t + I_t + \left(1 - \frac{1}{g_t}\right)Y_t = Y_t \quad (93)$$

$$Y_t = \frac{1}{D_t}(Z_t K_t^{\alpha} H_t^{1-\alpha} - \tilde{F}) \quad (94)$$

where equilibrium requires that $X_t = x_t$ and $I_t = i_t$, and that:

$$D_t = \zeta\left[\left(\frac{\pi_{t-1}}{\pi_t}\right)^{\iota}\left(\frac{\pi_{**}}{\pi_t}\right)^{(1-\iota)}\right]^{-\frac{(1+\lambda)}{\lambda}} D_{t-1} + (1-\zeta)\left[p_t^o\right]^{-\frac{(1+\lambda)}{\lambda}}. \quad (95)$$

(5) Aggregate disturbances (technology, money demand, government spending and monetary policy):

$$\ln Z_t = (1-\rho_Z)\ln Z_* + \rho_Z \ln Z_{t-1} + \varepsilon_{Z,t} \quad (96)$$

$$\ln \chi_t = (1-\rho_\chi)\ln \chi_* + \rho_\chi \ln \chi_{t-1} + \varepsilon_{\chi,t} \tag{97}$$

$$\ln g_t = (1-\rho_g)\ln g_* + \rho_g \ln g_{t-1} + \varepsilon_{g,t}\,, \tag{98}$$

where it is understood that innovations to the above laws of motion, as well as the monetary policy shock $\varepsilon_{R,t}$, are *iid* $N(0,\sigma_i^2)$ random variables, $i \in \{Z, \chi, g, R\}$.

Appendix A5. Steady State and Log-Linearized Equilibrium Conditions

In what follows we specialize the household's utility to be constant-relative-risk-aversion function:

$$U(x_t) = B\frac{x_t^{1-\gamma}}{1-\gamma}.$$

In addition, for any generic variable V_t the corresponding "star" variable V_* denotes its steady state value and "hat" variable stands for log-deviation from steady state: $\hat{V}_t = \ln(V_t/V_*)$

Steady State Conditions

$$R_* = \frac{\pi_*}{\beta}$$

$$R_*^k = \frac{1}{\beta} + \delta - 1$$

$$p_*^o = \left(\frac{1}{1-\zeta} - \frac{\zeta}{1-\zeta}\left(\frac{\pi_{**}}{\pi_*}\right)^{-\frac{1-\iota}{\lambda}}\right)^{-\lambda}$$

$$D_* = \frac{1-\zeta}{1-\zeta\left(\frac{\pi_{**}}{\pi_*}\right)^{-\frac{(1+\lambda)}{\lambda}(1-\iota)}}\left(p_*^o\right)^{-\frac{(1+\lambda)}{\lambda}}$$

$$\bar{Y}_* = Y_* D_*$$

$$K_* = \frac{\alpha p_*^o(\bar{Y}_* + \tilde{F})}{(1+\lambda)R_*^k}\left(\frac{1-\zeta\beta\left(\frac{\pi_{**}}{\pi_*}\right)^{-\frac{(1+\lambda)}{\lambda}(1-\iota)}}{1-\zeta\beta\left(\frac{\pi_{**}}{\pi_*}\right)^{-\frac{(1-\iota)}{\lambda}}}\right)$$

$$Z_* = \frac{(\bar{Y}_* + \tilde{F})}{K_*^\alpha H_*^{1-\alpha}}$$

$$I_* = \delta K_*$$

$$W_* = \frac{1-\alpha}{\alpha}\frac{K_*}{H_*}R_*^k$$

$$X_* + I_* + \left(1-\frac{1}{g_*}\right)Y_* = Y_*$$

$$A = \frac{1}{\bar{M}_*}\left(\frac{\chi_* W_* \pi_*^{\nu_m}}{(R_*-1)Z_*^{\frac{1-\nu_m}{1-\alpha}}}\right)^{\frac{1}{\nu_m}}$$

$$BX_*^{-\gamma} = \frac{A}{W_*}$$

Log-Linearized Equilibrium Conditions

<u>Households</u> $\hat{W}_t = \gamma \hat{X}_t$

$$\hat{I}_t = \frac{1}{1+\beta}\hat{I}_{t-1} + \frac{\beta}{1+\beta}\hat{I}_{t+1} + \frac{1}{S''(1+\beta)}\hat{\mu}_t$$

$$-\gamma\hat{X}_t = -\gamma\hat{X}_{t+1} + (\hat{R}_t - \hat{\pi}_{t+1})$$

$$\hat{\mu}_t - \gamma\hat{X}_t = \beta(1-\delta)\hat{\mu}_{t+1} - \gamma\hat{X}_{t+1} + \beta R_*^k \hat{R}_{t+1}^k$$

$$\hat{K}_{t+1} = (1-\delta)\hat{K}_t + \delta\hat{I}_t$$

$$\hat{\Xi}_{t|t-1}^p = -\gamma(\hat{X}_t - \hat{X}_{t-1}) - \hat{\pi}_t$$

$$\nu_m \hat{\bar{M}}_{t+1} = \gamma\hat{X}_t + \nu_m\hat{\chi}_{t+1} - (1-\nu_m)\hat{\pi}_{t+1} - \frac{1}{R_* - 1}\hat{R}_t$$

<u>Firms</u>

$$\hat{K}_t = \hat{H}_t + \hat{W}_t - \hat{R}_t^k$$

$$\hat{MC}_t = (1-\alpha)\hat{W}_t + \alpha\hat{R}_t^k - \hat{Z}_t$$

$$\hat{f}_t^{(1)} = \hat{f}_t^{(2)}$$

$$\hat{f}_t^{(1)} = (1-C_1)\left(-\frac{1+\lambda}{\lambda}\hat{p}_t^o + \hat{Y}_t\right) + C_1\left(-\frac{\iota}{\lambda}\hat{\pi}_t + \frac{1+\lambda}{\lambda}[-\hat{p}_t^o + \hat{\pi}_{t+1} + \hat{p}_{t+1}^o] + \hat{\Xi}_{t+1|t}^p + \hat{f}_{t+1}^{(1)}\right)$$

$$\hat{f}_t^{(2)} = (1-C_2)\left(-\left(\frac{1+\lambda}{\lambda}+1\right)\hat{p}_t^o + \hat{Y}_t + \hat{MC}_t\right) +$$
$$C_2\left(-\frac{\iota(1+\lambda)}{\lambda}\hat{\pi}_t + \left(\frac{1+\lambda}{\lambda}+1\right)[-\hat{p}_t^o + \hat{\pi}_{t+1} + \hat{p}_{t+1}^o] + \hat{\Xi}_{t+1|t}^p + \hat{f}_{t+1}^{(2)}\right)$$

$$\hat{p}_t^o = (C_3 - 1)\hat{\pi}_t - C_3\iota\zeta\left(\frac{\pi_{**}}{\pi_*}\right)^{-\frac{1-\iota}{\lambda}}\hat{\pi}_{t-1},$$

where $C_1 = \zeta\beta\left(\frac{\pi_{**}}{\pi_*}\right)^{-\frac{1-\iota}{\lambda}}$, $C_2 = \zeta\beta\left(\frac{\pi_{**}}{\pi_*}\right)^{-\frac{1+\lambda}{\lambda}(1-\iota)}$, $C_3 = \frac{1}{1-\zeta}\left(p_*^o\right)^{\frac{1}{\lambda}}$.

<u>Taylor Rule</u>

$$\hat{R}_t = \rho_R\hat{R}_{t-1} + (1-\rho_R)(\psi_1\hat{\pi}_t + \psi_2\hat{Y}_t) + \varepsilon_{R,t}$$

Aggregate Demand and Supply $\hat{\hat{Y}}_t = \hat{Y}_t + \hat{D}_t$

$$\hat{\hat{Y}}_t = \left(1 + \frac{\tilde{F}}{\bar{\bar{Y}}_*}\right)(\hat{Z}_t + \alpha \hat{K}_t + (1-\alpha)\hat{H}_t)$$

$$\hat{D}_t = \left(-\frac{p_*^0}{D_*}\frac{1+\lambda}{\lambda}(1-\zeta)\right)\hat{p}_t^o + \zeta\left(\frac{\pi_{**}}{\pi_*}\right)^{-\frac{1+\lambda}{\lambda}(1-\iota)}\left(\hat{D}_{t-1} + \frac{1+\lambda}{\lambda}\hat{\pi}_t - \frac{\iota(1+\lambda)}{\lambda}\hat{\pi}_{t-1}\right)$$

$$\hat{Y}_t = \frac{X_*}{X_* + I_*}\hat{X}_t + \frac{I_*}{X_* + I_*}\hat{I}_t + \hat{g}_t$$

Aggregate Disturbances

$$\hat{Z}_t = \rho_Z \hat{Z}_{t-1} + \varepsilon_{Z,t}, \qquad \varepsilon_{Z,t} \sim iid\ N(0, \sigma_Z^2)$$

$$\hat{\chi}_t = \rho_\chi \hat{\chi}_{t-1} + \varepsilon_{\chi,t}, \qquad \varepsilon_{\chi,t} \sim iid\ N(0, \sigma_\chi^2)$$

$$\hat{g}_t = \rho_g \hat{g}_{t-1} + \varepsilon_{g,t}, \qquad \varepsilon_{g,t} \sim iid\ N(0, \sigma_g^2)$$

$$\varepsilon_{R,t} \sim iid\ N(0, \sigma_R^2)$$

Appendix B. Details of Markov Chain Monte Carlo Algorithm

Appendix B1. Data-Rich DSGE Model: Gibbs Sampler: Step 2.2.a): Generating Unobserved States S^T

To sample the unobserved states S^T from $p(S^T \mid \Gamma, \boldsymbol{\theta}; X^T)$, given the state-space model parameters Γ and the structural DSGE model parameters $\boldsymbol{\theta}$, we will use the Carter-Kohn (1994) forward-backward algorithm. We begin by quasi-differencing the measurement equation (41)

$$X_t = \boldsymbol{\Lambda}(\boldsymbol{\theta}) S_t + e_t \tag{99}$$

to obtain the *iid* normal errors: $(\mathbf{I} - \boldsymbol{\Psi} L) X_t = (\mathbf{I} - \boldsymbol{\Psi} L) \boldsymbol{\Lambda}(\boldsymbol{\theta}) S_t + v_t$. Since the matrix polynomial multiplying S_t is of order 1, we can stack the additional lag of S_t and rewrite our linear Gaussian state-space system as follows:

$$\tilde{X}_t = \underbrace{\left[\boldsymbol{\Lambda}(\boldsymbol{\theta}) \;\vdots\; -\boldsymbol{\Psi}\boldsymbol{\Lambda}(\boldsymbol{\theta}) \right]}_{\tilde{\boldsymbol{\Lambda}}} \cdot \begin{bmatrix} S_t \\ S_{t-1} \end{bmatrix} + v_t \tag{100}$$

$$\underbrace{\begin{bmatrix} S_t \\ S_{t-1} \end{bmatrix}}_{\tilde{S}_t} = \underbrace{\begin{bmatrix} \mathbf{G}(\boldsymbol{\theta}) & \mathbf{0} \\ \mathbf{I} & \mathbf{0} \end{bmatrix}}_{\tilde{\mathbf{G}}} \underbrace{\begin{bmatrix} S_{t-1} \\ S_{t-2} \end{bmatrix}}_{\tilde{S}_{t-1}} + \underbrace{\begin{bmatrix} \mathbf{H}(\boldsymbol{\theta}) \\ \mathbf{0} \end{bmatrix}}_{\tilde{\mathbf{H}}} \varepsilon_t, \tag{101}$$

or more compactly:

$$\tilde{X}_t = \tilde{\boldsymbol{\Lambda}} \tilde{S}_t + v_t \tag{102}$$

$$\tilde{S}_t = \tilde{\mathbf{G}} \tilde{S}_{t-1} + \tilde{\mathbf{H}} \varepsilon_t \tag{103}$$

where $\tilde{X}_t = X_t - \boldsymbol{\Psi} X_{t-1}$, $v_t \sim iid\ N(\mathbf{0}, \mathbf{R})$, and $\varepsilon_t \sim iid\ N(\mathbf{0}, \mathbf{Q}(\boldsymbol{\theta}))$. For convenience, collect all the parameter matrices in $\Xi = \{\tilde{\boldsymbol{\Lambda}}, \mathbf{R}, \tilde{\mathbf{G}}, \tilde{\mathbf{H}}, \mathbf{Q}(\boldsymbol{\theta})\}$.

As in Carter-Kohn (1994), we first apply the Kalman filter to the state-space system (102)-(103) to generate the filtered DSGE states $\tilde{S}_{t|t}$ and their covariance matrices $\tilde{\mathbf{P}}_{t|t}$, for $t = 1..T$ (forward pass of the algorithm):

$$\text{prediction} \quad \begin{cases} \tilde{S}_{t+1|t} = \tilde{\mathbf{G}} \tilde{S}_{t|t} \\ \tilde{\mathbf{P}}_{t+1|t} = \tilde{\mathbf{G}} \tilde{\mathbf{P}}_{t|t} \tilde{\mathbf{G}}' + (\tilde{\mathbf{H}} \mathbf{Q}(\boldsymbol{\theta}) \tilde{\mathbf{H}}') \\ \eta_{t+1|t} = \tilde{X}_t - \tilde{\boldsymbol{\Lambda}} \tilde{S}_{t+1|t} \\ f_{t+1|t} = \tilde{\boldsymbol{\Lambda}} \tilde{\mathbf{P}}_{t+1|t} \tilde{\boldsymbol{\Lambda}}' + \mathbf{R} \end{cases} \tag{104}$$

$$\text{updating} \quad \begin{cases} \tilde{S}_{t+1|t+1} = \tilde{S}_{t+1|t} + \mathbf{K}_t \eta_{t+1|t} \\ \tilde{\mathbf{P}}_{t+1|t+1} = \tilde{\mathbf{P}}_{t+1|t} - \mathbf{K}_t \tilde{\boldsymbol{\Lambda}} \tilde{\mathbf{P}}_{t+1|t} \end{cases} \tag{105}$$

where $\mathbf{K}_t = \tilde{\mathbf{P}}_{t+1|t} \tilde{\boldsymbol{\Lambda}}' f_{t+1|t}^{-1}$ is the Kalman gain and $\eta_{t+1|t}$ is the period t prediction error. Second, starting from $\tilde{S}_{T|T}$ and $\tilde{\mathbf{P}}_{T|T}$, we roll back in time and draw the elements of S^T from a sequence of conditional Gaussian distributions. We draw $\tilde{S}_T$ from its conditional distribution given parameters Ξ and data $\tilde{X}^T$

$$\tilde{S}_T \mid \Xi, \tilde{X}^T \sim N(\tilde{S}_{T|T}, \tilde{\mathbf{P}}_{T|T}). \tag{106}$$

We generate $\tilde{S}_t$ for $t = T-1, T-2, ..., 1$ by proceeding backwards and by drawing from

$$\tilde{S}_t \mid \tilde{S}^*_{t+1}, \Xi, \tilde{X}^t \sim N(\tilde{S}_{t|t,\tilde{S}^*_{t+1}}, \tilde{\mathbf{P}}_{t|t,\tilde{S}^*_{t+1}}), \tag{107}$$

where $\tilde{X}^t = \left\{ \tilde{X}_1, ..., \tilde{X}_t \right\}$ and

$$\tilde{S}_{t|t,\tilde{S}^*_{t+1}} = \tilde{S}_{t|t} + \tilde{\mathbf{P}}_{t|t}\tilde{\mathbf{G}}^{*\prime}\left[\tilde{\mathbf{G}}^*\tilde{\mathbf{P}}_{t|t}\tilde{\mathbf{G}}^{*\prime} + \mathbf{Q}^*\right]^{-1}\left(\tilde{S}^*_{t+1} - \tilde{\mathbf{G}}^*\tilde{S}_{t|t}\right) \tag{108}$$

$$\tilde{\mathbf{P}}_{t|t,\tilde{S}^*_{t+1}} = \tilde{\mathbf{P}}_{t|t} - \tilde{\mathbf{P}}_{t|t}\tilde{\mathbf{G}}^{*\prime}\left[\tilde{\mathbf{G}}^*\tilde{\mathbf{P}}_{t|t}\tilde{\mathbf{G}}^{*\prime} + \mathbf{Q}^*\right]^{-1}\tilde{\mathbf{G}}^*\tilde{\mathbf{P}}_{t|t}\,. \tag{109}$$

Notice that the covariance matrix $\mathbf{\Sigma}_u$ of the error term $u_t = \tilde{\mathbf{H}}\varepsilon_t$ in state transition equation (103) is singular:

$$\mathbf{\Sigma}_u = E(u_t u_t') = E(\tilde{\mathbf{H}}\varepsilon_t\varepsilon_t'\tilde{\mathbf{H}}') = \left[\begin{array}{c|c} \mathbf{H}(\boldsymbol{\theta})\mathbf{Q}(\boldsymbol{\theta})\mathbf{H}(\boldsymbol{\theta})' & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{0} \end{array}\right] \tag{110}$$

Therefore, we use the approach of Kim and Nelson (1999b, p. 194-196) and condition the distribution of $\tilde{S}_t$ on only a non-identity-related part of $\tilde{S}_{t+1}$ (namely $\tilde{S}^*_{t+1}$) that corresponds to the non-singular upper-left corner of $\mathbf{\Sigma}_u$ (otherwise, if we conditioned on full state vector $\tilde{S}_{t+1}$, we would be unable to draw $\tilde{S}_t$, since the covariance matrix in (107) would be singular). This requires that

$$\tilde{S}^*_{t+1} = \mathbf{M}\tilde{S}_{t+1}, \quad \tilde{\mathbf{G}}^* = \mathbf{M}\tilde{\mathbf{G}}, \quad \mathbf{Q}^* = \mathbf{M}\mathbf{\Sigma}_u\mathbf{M}', \tag{111}$$

where $\mathbf{M}$ is the appropriate selection matrix consisting of 0s and 1s.

To initialize the Kalman filter (104)-(105), we set $\tilde{S}_{0|0}$ and $\tilde{\mathbf{P}}_{0|0}$ to the unconditional mean and covariance of the DSGE states $\tilde{S}_t$.

Appendix B2. Data-Rich DSGE Model: Gibbs Sampler: Step 2.2.b): Generating State-Space Parameters Γ

To sample the state-space parameters $\Gamma = \{\mathbf{\Lambda}, \mathbf{R}, \mathbf{\Psi}\}$ from $p(\Gamma \mid S^T, \boldsymbol{\theta}; X^T)$ given the unobserved DSGE states S^T and the structural DSGE model parameters $\boldsymbol{\theta}$, we use the approach of Chib and Greenberg (1994). Due to diagonality of $\mathbf{R}$ and $\mathbf{\Psi}$, and conditional on known unobserved states S^T, the equations (41)-(42) represent a collection of the linear regressions with AR(1) errors, with k^{th} equation given by

$$X_{k,t} = \mathbf{\Lambda}_k' S_t + e_{k,t} \tag{112}$$

$$e_{k,t} = \Psi_{kk} e_{k,t-1} + v_{k,t}, \qquad v_{k,t} \sim iid\ N(0, R_{kk}) \tag{113}$$

where $\mathbf{\Lambda}_k'$ is a $1 \times N$ vector and a k^{th} row of $\mathbf{\Lambda}$. Therefore in what follows we will draw the elements in Γ equation by equation for $k = 1..J$.

For each $\left(\mathbf{\Lambda}_k, R_{kk}, \Psi_{kk}\right)$, we consider the following conjugate prior distribution:

$$\begin{aligned} p(\mathbf{\Lambda}_k, R_{kk}, \Psi_{kk}) &= p(\mathbf{\Lambda}_k, R_{kk}) p(\Psi_{kk}) = \\ &= NIG_2(\mathbf{\Lambda}_k, R_{kk} \mid \mathbf{\Lambda}_{k,0}; \mathbf{M}_{k,0}; s_0; \nu_0) \times N(\Psi_{kk} \mid \Psi_0, \sigma_{\Psi,0}^2) \mathbf{1}_{\{|\Psi_{kk}|<1\}}, \end{aligned} \tag{114}$$

in which we set the parameters of Normal-Inverse-Gamma-2 density to $s_0 = 0.001$, $\nu_0 = 3$ and $\mathbf{\Lambda}_{k,0}$, $\mathbf{M}_{k,0}$ may in general depend on $\mathbf{\theta}$, and where we take $\Psi_0 = 0$ and $\sigma_{\Psi,0}^2 = 1$.

Conditional posterior density of $\left(\mathbf{\Lambda}_k, \boldsymbol{R}_{kk}\right)$: The posterior density is of the form

$$p(\mathbf{\Lambda}_k, R_{kk} \mid \Psi_{kk}; S^T, \mathbf{\theta}, X^T) \propto p(X_k^T \mid S^T, \mathbf{\Lambda}_k, R_{kk}, \Psi_{kk}, \mathbf{\theta}) p(\mathbf{\Lambda}_k, R_{kk}). \tag{115}$$

Define

$$X_{k,t}^* = X_{k,t} - \Psi_{kk} X_{k,t-1} \qquad S_t^* = S_t - \Psi_{kk} S_{t-1} \tag{116}$$

and rewrite (112)-(113) as a linear regression:

$$X_{k,t}^* = \mathbf{\Lambda}_k' S_t^* + v_{k,t}. \tag{117}$$

Define $T \times 1$ matrix $X_k^* = [X_{k,1}^*, X_{k,2}^*, ..., X_{k,T}^*]'$ and $T \times N$ matrix $S^* = [S_1^*, S_2^*, ..., S_T^*]'$ and rewrite (117) in matrix form:

$$X_k^* = S^* \mathbf{\Lambda}_k + v_k \tag{118}$$

It can be shown (Chib, Greenberg 1994, Bauwens, Lubrano, Richard 1999, Theorem 2.22, p. 57) that the likelihood of (118) is proportional to a Normal-Inverse-Gamma-2 density defined as

$$p(X_k^T \mid S^T, \mathbf{\Lambda}_k, R_{kk}, \Psi_{kk}, \boldsymbol{\theta}) \propto p_{NIG_2}(\mathbf{\Lambda}_k, R_{kk} \mid \hat{\mathbf{\Lambda}}_k, (S^{*\prime} S^*), s, T-N-2), \tag{119}$$

where[13]

$$\hat{\mathbf{\Lambda}}_k = \left(S^{*\prime} S^*\right)^{-1} S^{*\prime} X_k^* \tag{120}$$

$$s = X_k^{*\prime} \left(\mathbf{I}_T - S^* \left(S^{*\prime} S^*\right)^{-1} S^{*\prime} \right) X_k^* = X_k^{*\prime} \left(X_k^* - S^* \hat{\mathbf{\Lambda}}_k \right) \tag{121}$$

$$\begin{aligned} p_{NIG_2}(\boldsymbol{\beta}, \sigma^2 \mid \boldsymbol{\mu}, \mathbf{M}, s, \nu) = C_{Ng}^{-1}(\mathbf{M}, s, \nu; p)\left(\sigma^2\right)^{-\frac{1}{2}(\nu+p+2)} \times \\ \times \exp\left\{ -\frac{1}{2\sigma^2}\left[s + (\boldsymbol{\beta}-\boldsymbol{\mu})'\mathbf{M}(\boldsymbol{\beta}-\boldsymbol{\mu}) \right] \right\} \end{aligned} \tag{122}$$

Since the assumed prior $p(\mathbf{\Lambda}_k, R_{kk})$ is also of Normal-Inverse-Gamma-2 form, by Theorem 2.24 (Bauwens, Lubrano, Richard 1999, p. 56-61) we deduce:

$$\begin{aligned} p(\mathbf{\Lambda}_k, R_{kk} \mid \Psi_{kk}; S^T, \boldsymbol{\theta}, X^T) \propto p_{NIG_2}(\mathbf{\Lambda}_k, R_{kk} \mid \hat{\mathbf{\Lambda}}_k, (S^{*\prime} S^*), s, T-N-2) \times \\ \times\ p_{NIG_2}(\mathbf{\Lambda}_k, R_{kk} \mid \mathbf{\Lambda}_{k,0}, \mathbf{M}_{k,0}, s_0, \nu_0) \end{aligned}$$

---

[13] Normalization constant in p_{NIG_2} is $C_{Ng}(\mathbf{M}, s, \nu; p) = \text{Gamma}\left(\frac{\nu}{2}\right)\left(\frac{2}{s}\right)^{\nu/2} (2\pi)^{p/2} |\mathbf{M}|^{-\frac{1}{2}}$, where $p = \dim \boldsymbol{\beta}$.

$$\propto p_{NIG_2}(\mathbf{\Lambda}_k, R_{kk} \mid \bar{\mathbf{\Lambda}}_k, \bar{\mathbf{M}}_k, \bar{s}, \bar{\nu}), \tag{123}$$

with parameters given by

$$\bar{\mathbf{M}}_k = \mathbf{M}_{k,0} + \left(S^{*\prime} S^*\right)$$

$$\bar{\mathbf{\Lambda}}_k = \bar{\mathbf{M}}_k^{-1}\left(\mathbf{M}_{k,0}\mathbf{\Lambda}_{k,0} + (S^{*\prime}S^*)\hat{\mathbf{\Lambda}}_k\right)$$

$$\bar{s} = s_0 + s + \left(\mathbf{\Lambda}_{k,0} - \hat{\mathbf{\Lambda}}_k\right)' \left[\mathbf{M}_{k,0}^{-1} + \left(S^{*\prime}S^*\right)^{-1}\right]^{-1} \left(\mathbf{\Lambda}_{k,0} - \hat{\mathbf{\Lambda}}_k\right)$$

$$\bar{\nu} = \nu_0 + T \, .$$

The alternative equivalent expression for $\bar{s}$ used in computations is

$$\bar{s} = s_0 + s + \mathbf{\Lambda}_{k,0}'\mathbf{M}_{k,0}\mathbf{\Lambda}_{k,0} + \hat{\mathbf{\Lambda}}_k'\left(S^{*\prime}S^*\right)\hat{\mathbf{\Lambda}}_k - \bar{\mathbf{\Lambda}}_k'\bar{\mathbf{M}}_k\bar{\mathbf{\Lambda}}_k$$

The resulting conditional posterior density of $\left(\mathbf{\Lambda}_k, R_{kk}\right)$ is Normal-Inverse-Gamma-2, and we sample the loadings $\mathbf{\Lambda}_k$ and the variance of measurement error R_{kk} sequentially from:

$$\begin{array}{ll} R_{kk} \mid \Psi_{kk}; S^T, \boldsymbol{\theta}, X^T & \sim IG_2(\bar{s}, \bar{\nu}) \\ \mathbf{\Lambda}_k \mid R_{kk}, \Psi_{kk}; S^T, \boldsymbol{\theta}, X^T & \sim N_N(\bar{\mathbf{\Lambda}}_k, R_{kk}\bar{\mathbf{M}}_k^{-1}) \end{array} \tag{124}$$

Conditional posterior density of Ψ_{kk} : The posterior density is of the form

$$p(\Psi_{kk} \mid \mathbf{\Lambda}_k, R_{kk}; S^T, \boldsymbol{\theta}, X^T) \propto p(X_k^T \mid S^T, \mathbf{\Lambda}_k, R_{kk}, \Psi_{kk}, \boldsymbol{\theta})\, p(\Psi_{kk}) \tag{125}$$

Similar to what we did above, we define

$$e_{k,t} = X_{k,t} - \mathbf{\Lambda}_k' S_t \qquad e_k = \begin{bmatrix} e_{k,2} \\ \vdots \\ e_{k,T} \end{bmatrix} \qquad e_{k,-1} = \begin{bmatrix} e_{k,1} \\ \vdots \\ e_{k,T-1} \end{bmatrix} \tag{126}$$

and rewrite (113) in matrix form:

$$e_k = e_{k,-1}\Psi_{kk} + v_k \tag{127}$$

Because now we only care about the autocorrelation parameter Ψ_{kk}, the likelihood function in (125) is proportional to the normal density

$$\begin{aligned} p(X_k^T \mid S^T, \mathbf{\Lambda}_k, R_{kk}, \Psi_{kk}, \mathbf{\theta}) \quad & \propto \exp\left\{-\frac{1}{2R_{kk}}(e_k - e_{k,-1}\Psi_{kk})'(e_k - e_{k,-1}\Psi_{kk})\right\} \\ & \propto \exp\left\{-\frac{1}{2R_{kk}}(\Psi_{kk} - \hat{\Psi}_{kk})'\left(e'_{k,-1}e_{k,-1}\right)(\Psi_{kk} - \hat{\Psi}_{kk})\right\} \end{aligned} \tag{128}$$

with $\hat{\Psi}_{kk} = \left(e'_{k,-1}e_{k,-1}\right)^{-1} e'_{k,-1}e_k$. Provided that the prior for Ψ_{kk} is truncated normal with mean Ψ_0 and variance $\sigma^2_{\Psi,0}$, the conditional posterior density is proportional to a product of two normals:

$$\begin{aligned} p(\Psi_{kk} \mid \mathbf{\Lambda}_k, R_{kk}; S^T, \mathbf{\theta}, X^T) \quad & \propto \exp\left\{-\frac{1}{2R_{kk}}(\Psi_{kk} - \hat{\Psi}_{kk})'\left(e'_{k,-1}e_{k,-1}\right)(\Psi_{kk} - \hat{\Psi}_{kk})\right\} \\ & \propto \exp\left\{-\frac{1}{2\sigma^2_{\Psi,0}}(\Psi_{kk} - \Psi_0)^2\right\} \times \mathbf{1}_{\{|\Psi_{kk}|<1\}} \end{aligned}$$

This implies that the conditional posterior of Ψ_{kk} is (truncated) normal $N(\bar{\Psi}_{kk}, \bar{V}_{\Psi_{kk}}) \times \mathbf{1}_{\{|\Psi_{kk}|<1\}}$ with

$$\begin{aligned} \bar{V}_{\Psi_{kk}} &= \left(\left[R_{kk}\left(e'_{k,-1}e_{k,-1}\right)^{-1}\right]^{-1} + \left(\sigma^2_{\Psi,0}\right)^{-1}\right)^{-1} \\ \bar{\Psi}_{kk} &= \bar{V}_{\Psi_{kk}}\left(\left[R_{kk}\left(e'_{k,-1}e_{k,-1}\right)^{-1}\right]^{-1}\hat{\Psi}_{kk} + \left(\sigma^2_{\Psi,0}\right)^{-1}\Psi_0\right) \end{aligned} \tag{129}$$

Appendix C. Data: Description and Transformations

#	Short Name	SW Mnemonic	Trans Code	Description
Core Series				
	Real Output			
1.	RGDP		4	Real Per-capita Gross Domestic Product
2.	IP_TOTAL		4	Per-capita Industrial Production Index: Total
3.	IP_MFG		4	Per-capita Industrial Production Index: Manufacturing
	Inflation			
4.	PGDP		4	GDP Deflator Inflation
5.	PCED		4	Personal Consumption Expenditure Deflator Inflation
6.	CPI_ALL		4	Consumer Price Index (All Items) Inflation
	Nominal Interest Rate			
7.	FedFunds		4	Interest Rate: Federal Funds (effective), % per annum
8.	TBill_3m		4	Interest Rate: U.S. Treasury bills, secondary market, 3 month, % per annum
9.	AAABond		4	Bond Yield: Moody's AAA Corporate, % per annum
	Inverse Velocity of Money (M/Y)			
10.	IVM_M1S_det		4	Inverse Velocity of Money based on M1S aggregate
11.	IVM_M2S		4	Inverse Velocity of Money based on M2S aggregate
12.	IVM_MBase_bar		4	Inverse Velocity of Money based on adjusted Monetary Base
Non-Core Series				
	Output and Components			
1.	IP_CONS_DBLE	IPS13	3*	INDUSTRIAL PRODUCTION INDEX - DURABLE CONSUMER GOODS
2.	IP_CONS_NONDBLE	IPS18	3*	INDUSTRIAL PRODUCTION INDEX - NONDURABLE CONSUMER GOODS
3.	IP_BUS_EQPT	IPS25	3*	INDUSTRIAL PRODUCTION INDEX - BUSINESS EQUIPMENT
4.	IP_DBLE_MATS	IPS34	3*	INDUSTRIAL PRODUCTION INDEX - DURABLE GOODS MATERIALS
5.	IP_NONDBLE_MATS	IPS38	3*	INDUSTRIAL PRODUCTION INDEX - NONDURABLE GOODS MATERIALS
6.	IP_FUELS	IPS306	3*	INDUSTRIAL PRODUCTION INDEX - FUELS
7.	PMP	PMP	0	NAPM PRODUCTION INDEX (PERCENT)
8.	RCONS	GDP252	3*	Real Personal Consumption Expenditures, Quantity Index (2000=100) , SAAR
9.	RCONS_DUR	GDP253	3*	Real Personal Consumption Expenditures - Durable Goods , Quantity Index (2000=100), SAAR
10.	RCONS_SERV	GDP255	3*	Real Personal Consumption Expenditures - Services, Quantity Index (2000=100) , SAAR
11.	REXPORTS	GDP263	3*	Real Exports, Quantity Index (2000=100) , SAAR
12.	RIMPORTS	GDP264	3*	Real Imports, Quantity Index (2000=100) , SAAR
13.	RGOV	GDP265	3*	Real Government Consumption Expenditures & Gross Investment, Quantity Index (2000=100), SAAR
	Labor Market			
14.	EMP_MINING	CES006	3*	EMPLOYEES, NONFARM - MINING
15.	EMP_CONST	CES011	3*	EMPLOYEES, NONFARM - CONSTRUCTION
16.	EMP_DBLE_GDS	CES017	3*	EMPLOYEES, NONFARM - DURABLE GOODS
17.	EMP_NONDBLES	CES033	3*	EMPLOYEES, NONFARM - NONDURABLE GOODS
18.	EMP_SERVICES	CES046	3*	EMPLOYEES, NONFARM - SERVICE-PROVIDING
19.	EMP_TTU	CES048	3*	EMPLOYEES, NONFARM - TRADE, TRANSPORT, UTILITIES
20.	EMP_WHOLESALE	CES049	3*	EMPLOYEES, NONFARM - WHOLESALE TRADE
21.	EMP_RETAIL	CES053	3*	EMPLOYEES, NONFARM - RETAIL TRADE
22.	EMP_FIRE	CES088	3	EMPLOYEES, NONFARM - FINANCIAL ACTIVITIES
23.	EMP_GOVT	CES140	3	EMPLOYEES, NONFARM - GOVERNMENT
24.	URATE_ALL	LHUR	0	UNEMPLOYMENT RATE: ALL WORKERS, 16 YEARS & OVER (%,SA)
25.	U_DURATION	LHU680	0	UNEMPLOY.BY DURATION: AVERAGE(MEAN)DURATION IN WEEKS (SA)
26.	U_L5WKS	LHU5	3	UNEMPLOY.BY DURATION: PERSONS UNEMPL.LESS THAN 5 WKS (THOUS.,SA)
27.	U_5_14WKS	LHU14	3	UNEMPLOY.BY DURATION: PERSONS UNEMPL.5 TO 14 WKS (THOUS.,SA)
28.	U_M15WKS	LHU15	3	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 WKS + (THOUS.,SA)
29.	U_15_26WKS	LHU26	3	UNEMPLOY.BY DURATION: PERSONS UNEMPL.15 TO 26 WKS (THOUS.,SA)
30.	U_M27WKS	LHU27	3	UNEMPLOY.BY DURATION: PERSONS UNEMPL.27 WKS + (THOUS,SA)
31.	HOURS_AVG	CES151	0	AVG WKLY HOURS, PROD WRKRS, NONFARM - GOODS-PRODUCING
	Housing			
32.	HSTARTS_NE	HSNE	1	HOUSING STARTS:NORTHEAST (THOUS.U.)S.A.
33.	HSTARTS_MW	HSMW	1	HOUSING STARTS:MIDWEST(THOUS.U.)S.A.
34.	HSTARTS_SOU	HSSOU	1	HOUSING STARTS:SOUTH (THOUS.U.)S.A.
35.	HSTARTS_WST	HSWST	1	HOUSING STARTS:WEST (THOUS.U.)S.A.

35.	HSTARTS_WST	HSWST	1	HOUSING STARTS:WEST (THOUS.U.)S.A.
36.	RRESINV	GDP261	3*	Real Gross Private Domestic Investment - Residential, Quantity Index (2000=100), SAAR
	Financial Variables			
37.	SFYGM6	Sfygm6	0	fygm6-fygm3 fygm6: INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,6-MO.(% PER ANN,NSA) fygm3: INTEREST RATE: U.S.TREASURY BILLS,SEC MKT,3-MO.(% PER ANN,NSA)
38.	SFYGT1	Sfygt1	0	fygt1-fygm3 fygt1: INTEREST RATE: U.S.TREASURY CONST MATURITIES,1-YR.(% PER ANN,NSA)
39.	SFYGT10	Sfygt10	0	fygt10-fygm3 fygt10: INTEREST RATE: U.S.TREASURY CONST MATURITIES,10-YR.(% PER ANN,NSA)
40.	SFYBAAC	sFYBAAC	0	FYBAAC-Fygt10 FYBAAC: BOND YIELD: MOODY'S BAA CORPORATE (% PER ANNUM)
41.	BUS_LOANS	BUSLOANS	3	Commercial and Industrial Loans at All Commercial Banks (FRED) Billions $ (SA)
42.	CONS_CREDIT	CCINRV	3*	CONSUMER CREDIT OUTSTANDING - NONREVOLVING(G19)
43.	DLOG_EXR_US	EXRUS	2	UNITED STATES;EFFECTIVE EXCHANGE RATE(MERM)(INDEX NO.)
44.	DLOG_EXR_CHF	EXRSW	2	FOREIGN EXCHANGE RATE: SWITZERLAND (SWISS FRANC PER U.S.$)
45.	DLOG_EXR_YEN	EXRJAN	2	FOREIGN EXCHANGE RATE: JAPAN (YEN PER U.S.$)
46.	DLOG_EXR_GBP	EXRUK	2	FOREIGN EXCHANGE RATE: UNITED KINGDOM (CENTS PER POUND)
47.	DLOG_EXR_CAN	EXRCAN	2	FOREIGN EXCHANGE RATE: CANADA (CANADIAN $ PER U.S.$)
48.	DLOG_SP500	FSPCOM	2	S&P'S COMMON STOCK PRICE INDEX: COMPOSITE (1941-43=10)
49.	DLOG_SP_IND	FSPIN	2	S&P'S COMMON STOCK PRICE INDEX: INDUSTRIALS (1941-43=10)
50.	DLOG_DJIA	FSDJ	2	COMMON STOCK PRICES: DOW JONES INDUSTRIAL AVERAGE
	Investment, Inventories, Orders			
51.	NAPMI	PMI	0	PURCHASING MANAGERS' INDEX (SA)
52.	NAPM_NEW_ORDRS	PMNO	0	NAPM NEW ORDERS INDEX (PERCENT)
53.	NAPM_VENDOR_DEL	PMDEL	0	NAPM VENDOR DELIVERIES INDEX (PERCENT)
54.	NAPM_INVENTORIES	PMNV	0	NAPM INVENTORIES INDEX (PERCENT)
55.	RINV_GDP	GDP256	3*	Real Gross Private Domestic Investment, Quantity Index (2000=100) , SAAR
56.	RNONRESINV_STRUCT	GDP259	1	Real Gross Private Domestic Investment - Nonresidential - Structures, Quantity Index (2000=100), SAA
57.	RNONRESINV_BEQUIPT	GDP260	3*	Real Gross Private Domestic Investment - Nonresidential - Equipment & Software
	Prices and Wages			
58.	RAHE_CONST	CES277R	3*	REAL AVG HRLY EARNINGS, PROD WRKRS, NONFARM - CONSTRUCTION (CES277/PI071)
59.	RAHE_MFG	CES278R	3	REAL AVG HRLY EARNINGS, PROD WRKRS, NONFARM - MFG (CES278/PI071)
60.	P_COM	PSCCOMR	2	Real SPOT MARKET PRICE INDEX:BLS & CRB: ALL COMMODITIES(1967=100) (PSCCOM/PCEPIL PSCCOM: SPOT MARKET PRICE INDEX:BLS & CRB: ALL COMMODITIES(1967=100) PCEPILFE: PCE Price Index Less Food and Energy (SA) Fred
61.	P_OIL	PW561R	2	PPI Crude (Relative to Core PCE) (pw561/PCEPiLFE) pw561: PRODUCER PRICE INDEX: CRUDE PETROLEUM (82=100,NSA)
62.	P_NAPM_COM	PMCP	2	NAPM COMMODITY PRICES INDEX (PERCENT)
63.	RCOMP_HOUR	LBPUR7	1*	REAL COMPENSATION PER HOUR,EMPLOYEES:NONFARM BUSINESS(82=100,SA)
64.	ULC	LBLCPU	1*	UNIT LABOR COST: NONFARM BUSINESS SEC (1982=100,SA)
65.	PCED_DUR	GDP274A	2	Personal Consumption Expenditures: Durable goods Price Index
66.	PCED_NDUR	GDP275A	2	Personal Consumption Expenditures: Nondurable goods Price Index
67.	PCED_SERV	GDP276A	2	Personal Consumption Expenditures: Services Price Index
68.	PINV_GDP	GDP277A	2	Gross private domestic investment Price Index
69.	PINV_NRES_STRUCT	GDP280A	2	GPDI Price Index: Structures
70.	PINV_NRES_EQP	GDP281A	2	GPDI Price Index: Equipment and software Price Index
71.	PINV_RES	GDP282A	2	GPDI Price Index: Residential Price Index
72.	PEXPORTS	GDP284A	2	GDP: Exports Price Index
73.	PIMPORTS	GDP285A	2	GDP: Imports Price Index
74.	PGOV	GDP286A	2	Government consumption expenditures and gross investment Price Index
	Other			
75.	UTL11	UTL11	0	CAPACITY UTILIZATION - MANUFACTURING (SIC)
76.	UMICH_CONS	HHSNTN	1	U. OF MICH. INDEX OF CONSUMER EXPECTATIONS(BCD-83)
77.	LABOR_PROD	LBOUT	1*	OUTPUT PER HOUR ALL PERSONS: BUSINESS SEC(1982=100,SA)

Notes: Transformation codes: 0 – nothing; 1 – log(); 2 – dlog(); 3 – log of the ratio of subaggregate to aggregate; 4 – transformation described in the main text, pp. 29. Asterisk (*) indicates the transformed variable has been further linearly detrended.

Source of data: Stock and Watson (2008), “Forecasting in Dynamic Factor Models Subject to Structural Instability,” available online at http://www.princeton.edu/~mwatson/ddisk/hendryfestschrift_replicationfiles_April28_2008.zip

Full sample available: 1959:Q1-2006:Q4. Sample used in estimation: 1984:Q1-2005:Q4.

All series available at monthly frequency have been converted to quarterly by simple averaging in native units.

Appendix D. Tables and Figures

Table D1. Data-Rich DSGE Model: Parameters Fixed During Estimation - Calibration and Normalization

Parameter Name	Mnemonics	Value
Depreciation rate	δ	0.014
Risk aversion in HH utility function	γ	1
Money demand shock in steady state	χ_*	1
Share of govt spending in steady state	g_*	1.2
Fixed costs in production	F	0
MP rule: response to inflation	ψ_1	1.82
MP rule: response to output gap	ψ_2	0.18
MP rule: int rate smoothing parameter	ρ_R	0.78
Persistence: TFP shock	ρ_Z	0.98
Steady state inflation (in % pa)	π_A	2.5
Steady state real interest rate (in % pa)	r_A	2.84
Price indexation parameter	π_{**}	1
Steady state real GDP	Y_*	1
Log inverse velocity of money in SS	$\log(M_* / Y_*)$	0.778
Steady state of log average inverse labor productivity	$\log(H_* / Y_*)$	–3.5
Transformations: $\beta = \frac{1}{1 + r_A/400}$; $\pi_* = 1 + \frac{\pi_A}{400}$		

Table D2. Data-Rich DSGE Model: Prior Distributions

Parameter Name		Domain	Density	Para 1	Para 2
Firms					
Share of capital	α	[0;1)	Beta	0.3	0.025
Average economy wide markup	λ	*R+*	Gamma	0.15	0.01
$1-\zeta$ prob of reoptimizing firm's price	ζ	[0;1)	Beta	0.6	0.15
Indexation parameter	ι	[0;1)	Beta	0.5	0.25
Households					
Elasticity of money demand	ν_m	*R+*	Gamma	20	5
Investment adjustment cost parameter	S''	*R+*	Gamma	5.0	2.5
Shocks					
Persistence: govt spending process	ρ_g	[0;1)	Beta	0.8	0.1
Persistence: money demand shock	ρ_χ	[0;1)	Beta	0.8	0.1
Stdev: govt spending process	σ_g	*R+*	InvGamma	1	4
Stdev: money demand shock	σ_χ	*R+*	InvGamma	1	4
Stdev: monetary policy shock	σ_R	*R+*	InvGamma	0.5	4
Stdev: TFP shock	σ_Z	*R+*	InvGamma	1	4

Notes: Para 1 and Para 2 are (i) the means and the standard deviations for Beta, Gamma, and Normal distributions; (ii) the upper and the lower bound of support for the Uniform distribution; (iii) s and ν for the Inverse Gamma distribution, where $p_{IG}(\sigma \mid s,\nu) \propto \sigma^{-\nu-1}\exp(-\nu s^2/2\sigma^2)$.

Table D3. Data-Rich DSGE Model: Posterior Estimates

Parameter Name		Regular DSGE model Mean	Regular DSGE model 90% CI	Data-Rich DSGE model Mean	Data-Rich DSGE model 90% CI
Firms					
Share of capital	α	0.282	[0.269, 0.296]	0.2766	[0.266, 0.292]
Average economy wide markup	λ	0.15	[0.133, 1.166]	0.134	[0.117, 0.154]
$1-\zeta$ prob of reoptimizing firm's price	ζ	0.759	[0.709, 0.809]	0.797	[0.777, 0.819]
Indexation parameter	ι	0.05	[0.00, 0.101]	0.0326	[0.001, 0.0636]
Households					
Elasticity of money demand	ν_m	25.943	[19.581, 31.65]	23.199	[17.13, 31.27]
Investment adjustment cost parameter	S''	11.079	[6.299, 15.683]	30.754	[26.506, 35.29]
Shocks					
Persistence: govt spending process	ρ_g	0.886	[0.85, 0.92]	0.870	[0.839, 0.909]
Persistence: money demand shock	ρ_χ	0.974	[0.958, 0.992]	0.961	[0.936, 0.981]
Stdev: govt spending process	σ_g	1.227	[1.062, 1.388]	0.851	[0.605, 1.238]
Stdev: money demand shock	σ_χ	0.865	[0.757, 0.972]	0.396	[0.327, 0.464]
Stdev: monetary policy shock	σ_R	0.199	[0.175, 0.223]	0.2404	[0.211, 0.275]
Stdev: TFP shock	σ_Z	0.557	[0.471, 0.639]	0.375	[0.322, 0.439]
Implied Slope of NK Phillips Curve	κ	0.0745		0.0517	

Notes: Results labeled **"Regular DSGE model"** refer to the standard Bayesian estimation of the same underlying theoretical DSGE model as presented in the main text, but only on 4 core observable data series (real GDP, GDP deflator inflation, the federal funds rate and the inverse velocity of money based on the M2S aggregate) assumed to be perfectly measured. In terms of the state-space representation (40)-(42), this means that the vector of data X_t contains just these 4 core observables, the factor loadings $\mathbf{\Lambda}$ are restricted as below, and there are no measurement errors e_t:

$$\underbrace{\begin{bmatrix} \text{RealGDP}_t \\ \text{GDP\_Def\_Inflation}_t \\ \text{FedFundsRate}_t \\ \text{IVM\_M2S}_t \end{bmatrix}}_{X_t} = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 4 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 4 & 0 & \cdots & 0 \\ -1 & 0 & 0 & 1 & \cdots & 0 \end{bmatrix}}_{\mathbf{\Lambda}} \cdot \underbrace{\begin{bmatrix} \hat{Y}_t \\ \hat{\pi}_t \\ \hat{R}_t \\ \hat{M}_t \\ \vdots \end{bmatrix}}_{S_t} \tag{130}$$

Table D4. Data-Rich DSGE Model: Summary of the Unconditional Variance Decomposition

iid Measurement Errors; Dataset = DFM3.txt
on average, 20K draws, 4K burn-in

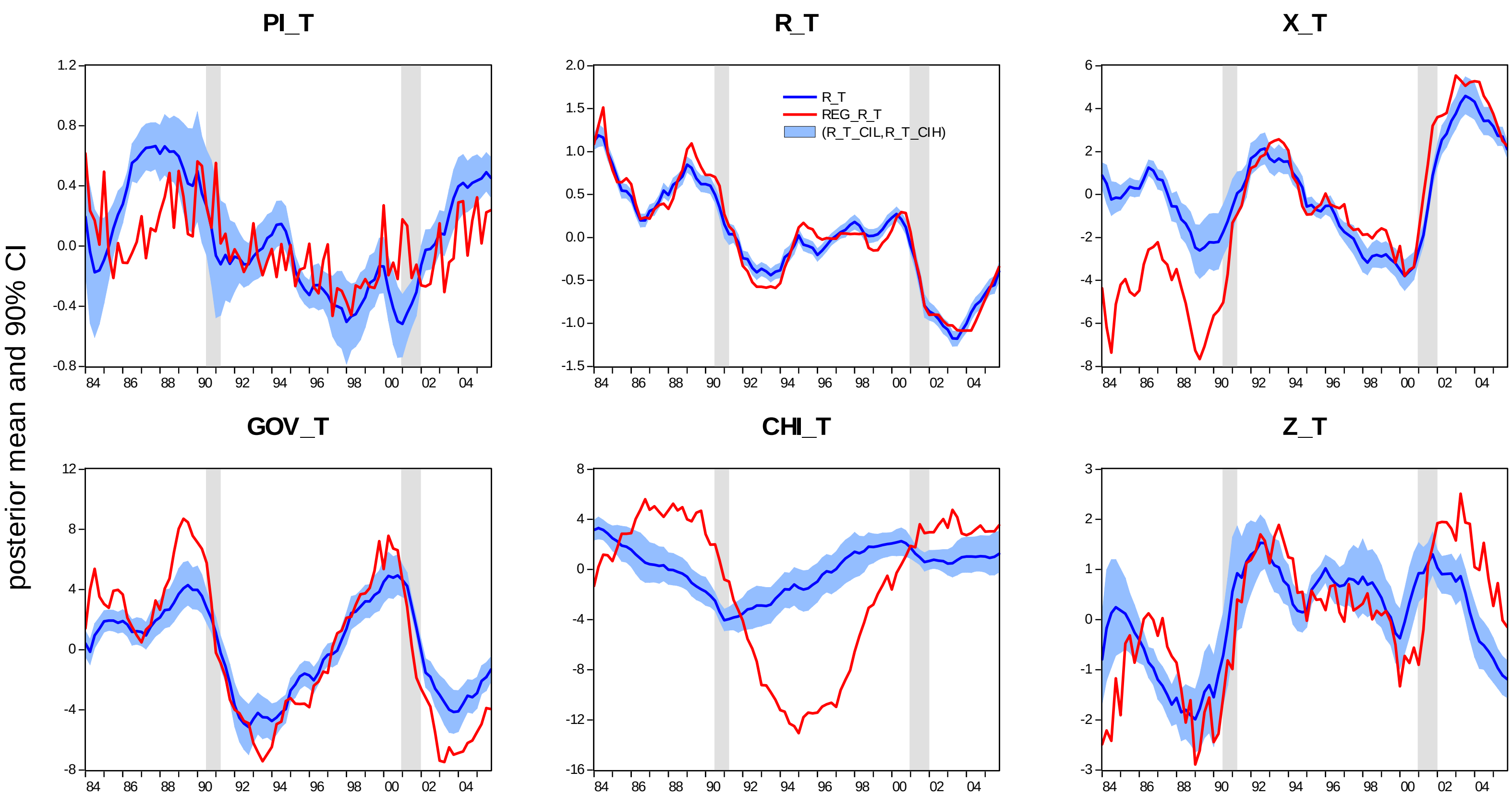
	GOV	CHI	MP	Z	All Shocks	Error term
Core Variables	**0.05**	**0.08**	**0.06**	**0.56**	**0.749**	**0.251**
Real output	0.14	0.21	0.03	0.48	0.852	0.148
Inflation	0.01	0.02	0.01	0.70	0.733	0.267
Interest rates	0.01	0.00	0.15	0.76	0.925	0.075
Money velocities	0.07	0.09	0.04	0.29	0.489	0.512
Non-Core Variables	**0.09**	**0.13**	**0.06**	**0.45**	**0.719**	**0.281**
Output and components	0.07	0.27	0.08	0.45	0.873	0.127
Labor market	0.19	0.14	0.06	0.46	0.848	0.152
Investment, inventories, orders	0.10	0.13	0.02	0.63	0.882	0.118
Housing	0.04	0.26	0.07	0.42	0.794	0.206
Prices and wages	0.03	0.05	0.04	0.45	0.568	0.432
Financial variables	0.06	0.03	0.05	0.32	0.451	0.549
Other	0.02	0.12	0.09	0.64	0.866	0.134

Table D5. Data-Rich DSGE vs. Regular DSGE Model: Unconditional Variance Decomposition

		GOV	CHI	MP	Z	All Shocks	Measurement Error
Regular DSGE:	**Real GDP**	0.099	0.000	0.012	0.889	1.000	-
Data-Rich DSGE:	**Real GDP**	0.081	0.000	0.040	0.648	0.770	0.230
	IP Total	0.167	0.308	0.021	0.395	0.891	0.110
	IP Manufacturing	0.166	0.317	0.020	0.392	0.894	0.106
Regular DSGE:	**GDP Def inflation**	0.020	0.000	0.009	0.970	1.000	-
Data-Rich DSGE:	**GDP Def inflation**	0.011	0.000	0.011	0.789	0.811	0.189
	PCE Def inflation	0.004	0.035	0.003	0.703	0.745	0.255
	CPI ALL Inflation	0.005	0.031	0.006	0.600	0.642	0.358
Regular DSGE:	**Fed Funds rate**	0.001	0.000	0.040	0.959	1.000	-
Data-Rich DSGE:	**Fed Funds rate**	0.004	0.000	0.135	0.817	0.956	0.044
	3m T-Bill rate	0.007	0.003	0.160	0.788	0.958	0.042
	AAA Bond yield	0.013	0.008	0.168	0.672	0.861	0.139
Regular DSGE:	**IVM_M2S**	0.117	0.596	0.001	0.286	1.000	-
Data-Rich DSGE:	**IVM_M1S_det**	0.055	0.174	0.016	0.404	0.648	0.352
	IVM_M2S	0.042	0.063	0.003	0.071	0.178	0.822
	IVM_MBASE_bar	0.099	0.031	0.104	0.406	0.639	0.361

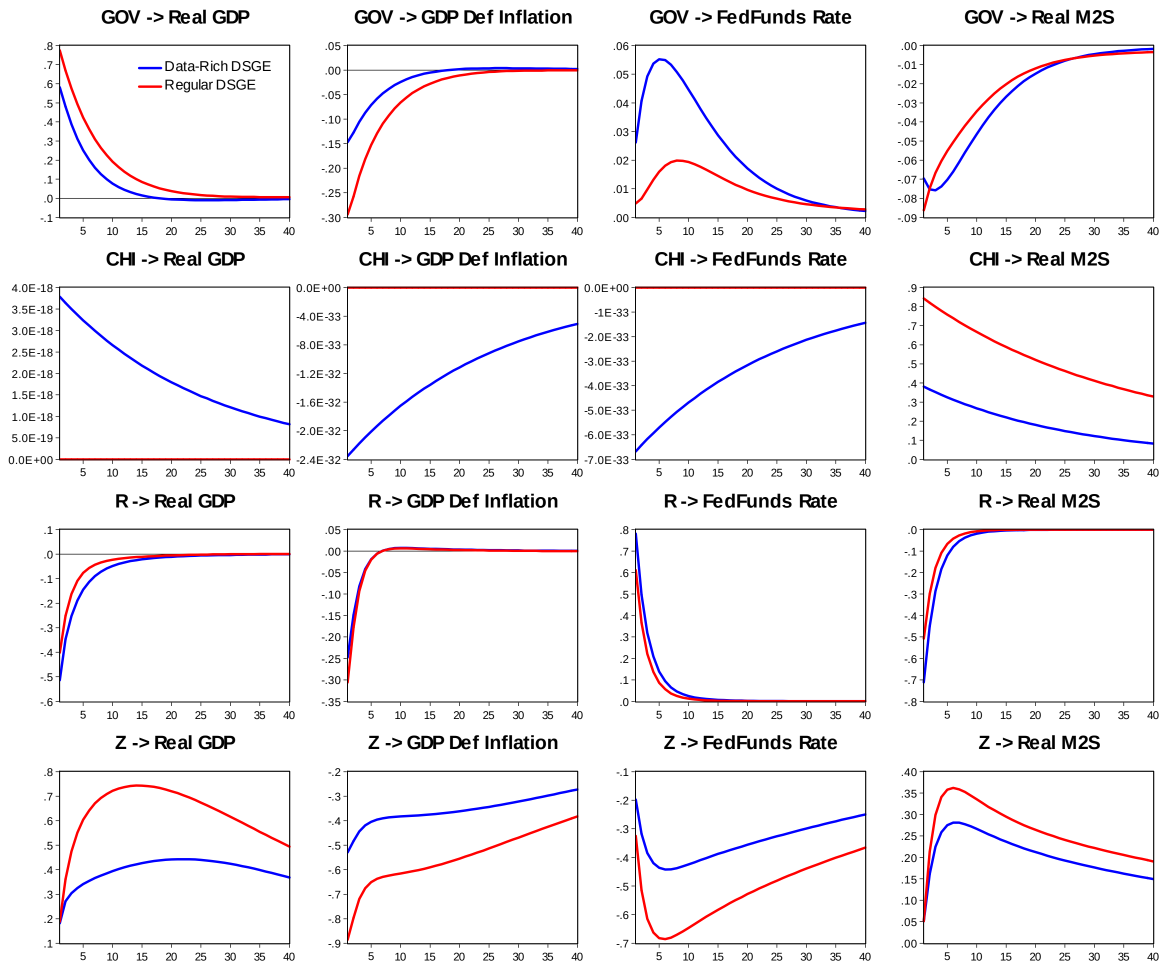
Notes: Structural shocks are GOV - government spending, CHI - money demand, MP - monetary policy, and Z - neutral technology.
Data-Rich DSGE Model: *iid* errors; dataset = dfm3.txt; algorithm: Jungbacker-Koopman; 20K draws, 4K burn-in; VD: posterior mean
Regular DSGE Model: no measurement errors; dataset = 4 primary observables; 100K draws, 20K burn-in; VD: posterior mean

Figure D1. Data-Rich DSGE Model (iid errors): Estimated Model States



Notes: Figure depicts the posterior means and 90% credible intervals of the data-rich DSGE model state variables (blue line and bands): inflation (PI_T, π_t), nominal interest rate (R_T, R_t), real consumption (X_T, x_t), government spending shock (GOV_T, g_t), money demand shock (CHI_T, χ_t), and neutral technology shock (Z_T, Z_t). Red line corresponds to the smoothed versions of the same variables in a *regular* DSGE model estimation derived by Kalman smoother at posterior mean of deep structural parameters (see notes to Table D3 for definition of "regular DSGE estimation").

Figure D2. Impulse Responses to Structural Shocks: Primary Observables



Shocks: GOV - government spending; CHI - money demand; R - monetary policy; Z - technology

Figure D3. Impact of Monetary Policy Innovation on Core Macro Series: Regular vs. Data-Rich DSGE Model

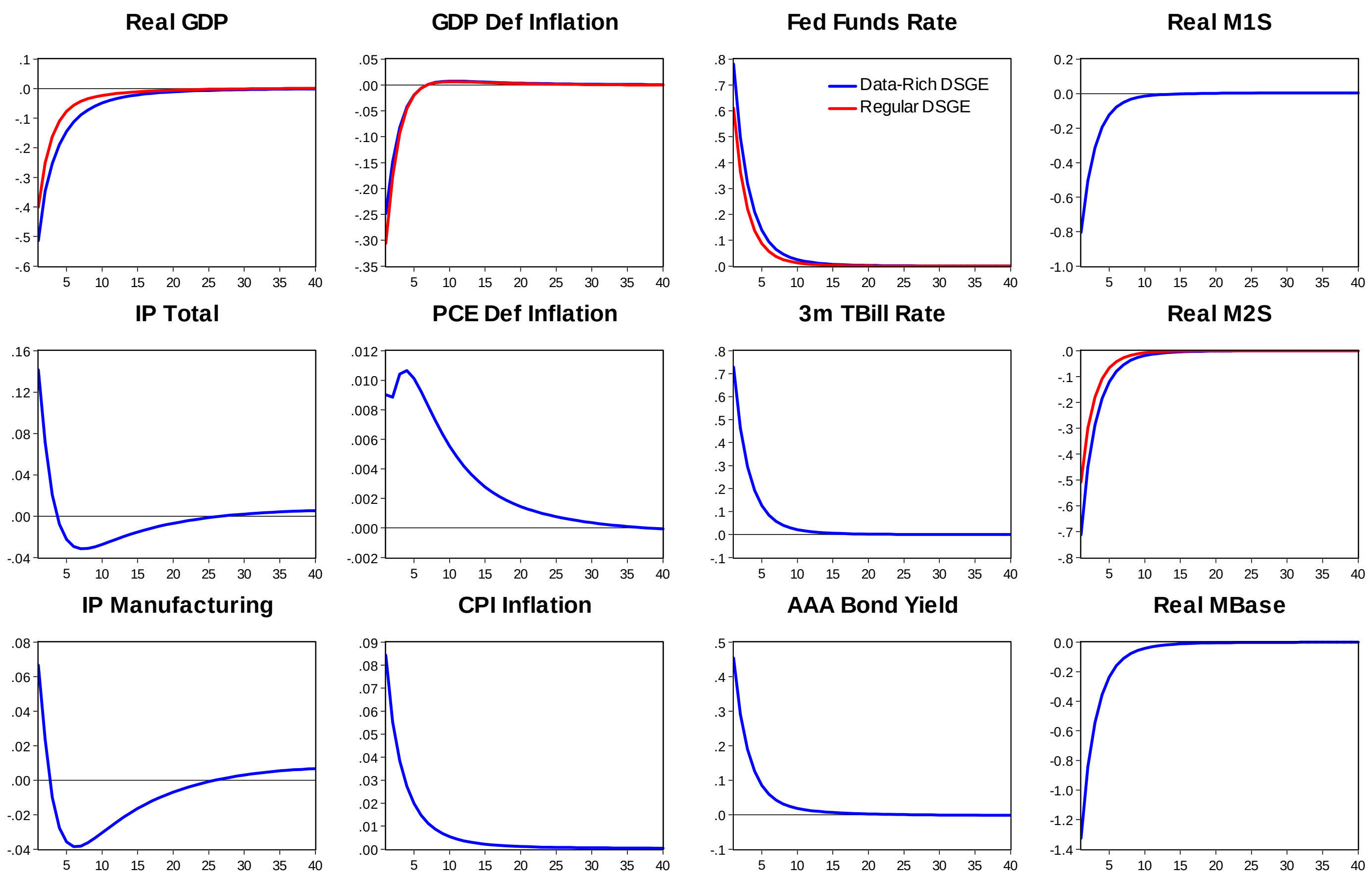
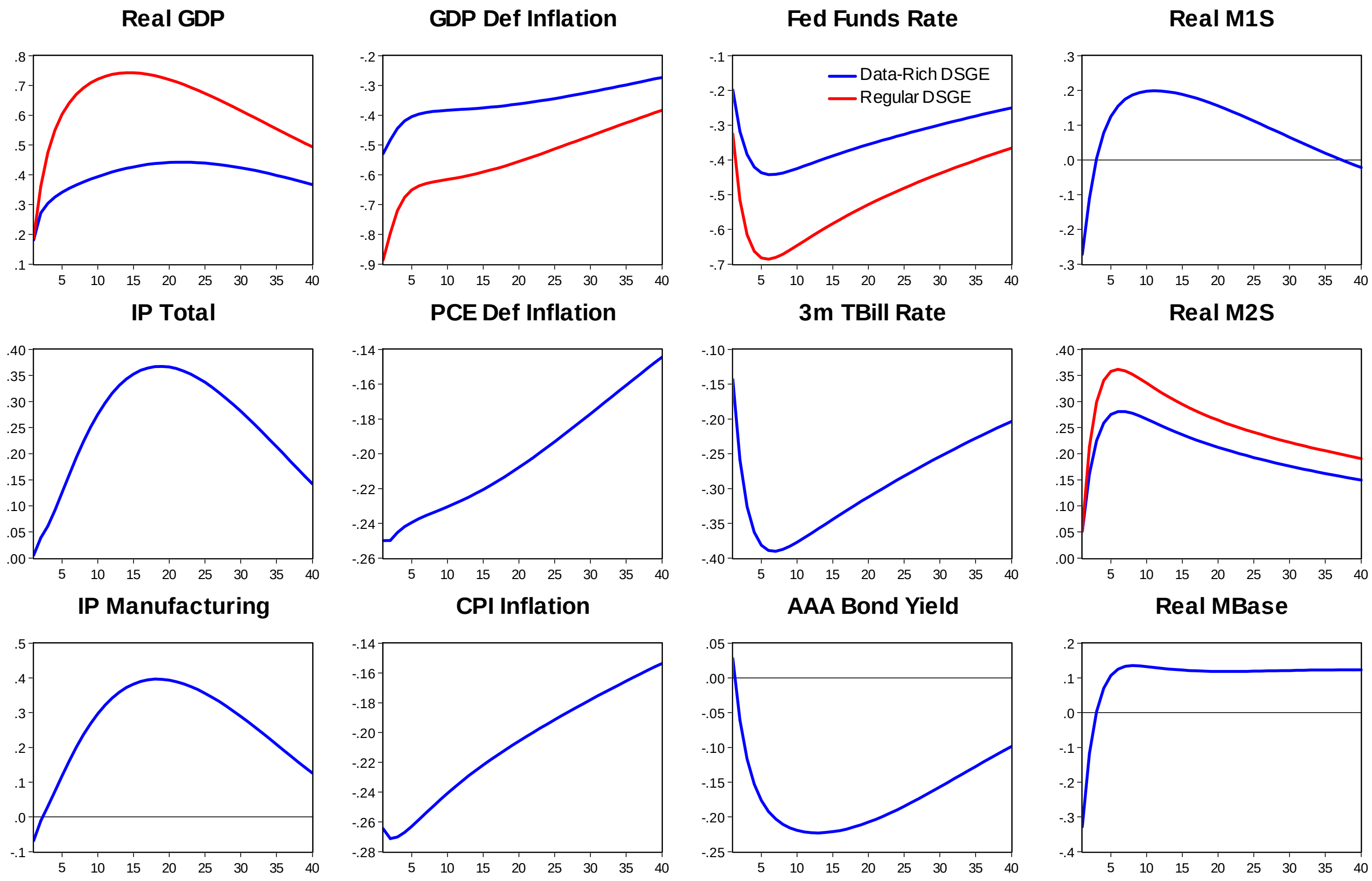


Figure D4. Impact of Technology Innovation on Core Macro Series: Regular vs. Data-Rich DSGE Model



CHAPTER 2. DATA-RICH DSGE AND DYNAMIC FACTOR MODELS

1 Introduction

Dynamic factor models (DFM) and dynamic stochastic general equilibrium (DSGE) models are widely used for empirical research in macroeconomics. The traditional areas of DFM application are the construction of coincident and leading indicators (e.g., Stock and Watson 1989, Altissimo et al. 2001) and the forecasting of macro time series (Stock and Watson 1999, 2002a, b; Forni, Hallin, Lippi and Reichlin 2003; Boivin and Ng 2005). DFMs are also used for real-time monitoring (Giannone, Reichlin, Small 2008; Aruoba, Diebold, and Scotti 2009), in monetary policy applications (e.g., the Factor Augmented VAR approach of Bernanke, Boivin, and Eliasz 2005, Stock and Watson 2005) and in the study of international business cycles (Kose, Otrok, Whiteman 2003, 2008; Del Negro and Otrok 2008). The micro-founded optimization-based DSGE models primarily focus on understanding the sources of business cycle fluctuations and on assessing the importance of nominal rigidities and various types of frictions in the economy. Recently, they appear to have been able to replicate well many salient features of the data (e.g., Christiano, Eichenbaum, and Evans 2005; Smets and Wouters 2003, 2007). As a result, the versions of DSGE models extended to open economy and multisector contexts are increasingly used as tools for projections and policy analysis at major central banks (Adolfson et al. 2007, 2008; Edge, Kiley and Laforte 2009; Coenen, McAdam and Straub 2008).

The empirical factor literature argues that the co-movement of large panels of macroeconomic and financial data can be captured by relatively few common unobserved factors. Early work by Sargent and Sims (1977) found that the dynamic index model with two indices fits well the real variables in their panel. Giannone, Reichlin and Sala (2004) claim that the number of common shocks, or, in their terminology, the stochastic dimension of the U.S. economy, is two. Based on recent theoretical work developing more formal number-of-factors criteria, several authors (e.g., Bai and Ng 2007; Hallin and Liška 2007; Stock and Watson 2005) have argued for a higher number of dynamic factors that drive large U.S. macroeconomic panels – ranging from four to seven.

The dynamics in DSGE models are also often governed by a handful of state variables and exogenous processes such as preference and/or technology shocks. Boivin and Giannoni (2006) combine a DSGE and a factor model into a data-rich DSGE model, in which DSGE states are factors and factor dynamics are subject to DSGE model implied restrictions. They argue that the richer information coming from large macroeconomic and financial panels can provide better estimates of the DSGE states and of the structural shocks driving the economy. On top of that, Boivin and Giannoni (2006) showed – and we confirm their conclusions in Chapter 1 – that the data-rich DSGE model delivers different estimates of deep structural parameters of the model compared to standard non-data-rich estimation.

In this chapter, we take both a data-rich DSGE model and an empirical dynamic factor model to the same rich data set, and ask: How similar or different would be the latent empirical factors extracted by a factor model versus the estimated data-rich DSGE model states? Do they span a common factor space? Or – in other words – can we predict the true estimated DFM latent factors from the DSGE model states with a fair amount of accuracy? We ask this question for three reasons. First, the factor spaces comparison may serve as a useful tool for evaluating a DSGE model. Recent research has shown that misspecification remains a concern for valid inference in DSGE models (Del Negro, Schorfheide, Smets and Wouters 2007 – DSSW hereafter). If a DSGE model is taken to a particular small set of observables, misspecification often manifests itself through the inferior fit. Dynamic factor models usually fit well and perform well in forecasting. So if it turns out that the spaces spanned by two models are close, that is good news for a DSGE model. This means that a DSGE model overall captures the sources of co-movement in the large panel of data as a sort of a core, and that the differences in fit between a data-rich DSGE model and a DFM are potentially due to restricted factor loadings in the former. Second, it is well known that the latent common components extracted by dynamic factor models from the large panels of data do not mean much in general. That's one of the biggest weaknesses of DFMs. If factor spaces in two models are closely aligned, this facilitates the economic interpretation of a dynamic factor model, since the empirical factors become isomorphic to the DSGE model state variables with clear economic meaning. Third, if factor spaces are close, we are able to propagate the structural shocks in otherwise completely non-structural dynamic factor model to obtain predictions for a broad range of macro series of interest.[14] This way of doing policy analysis is more reliable, because, on top of the impulse responses derived in the data-rich DSGE model, which might be misspecified, we are able to generate a second set of responses to the same shocks in the context of a factor model that is primarily data-driven and fits better.

We compare a data-rich DSGE model with a standard New Keynesian core to an empirical dynamic factor model by estimating both on a rich panel of U.S. macroeconomic and financial data compiled by Stock and Watson (2008). The specific version of the data-rich DSGE model is taken from Chapter 1. The estimation involves Bayesian Markov Chain Monte Carlo (MCMC) methods.

We find that the spaces spanned by the empirical factors and by the data-rich DSGE model states are very close meaning that, using a collection of linear regressions, we are able to predict the true estimated factors from the DSGE states fairly accurately. Given the accuracy, we can use this predictive link to map in every period the impact of any structural DSGE shock on the data-rich DSGE states into the empirical factors. We then multiply the responses of empirical factors by the DFM factor loadings to generate the impulse responses of data indicators to structural shocks. Applying this procedure, we propagate monetary policy and technology innovations in an otherwise non-structural dynamic factor model to obtain predictions for many more series than just a handful of

[14] This is similar in spirit to the Factor Augmented VAR approach (originally implemented by Bernanke, Boivin and Eliasz (2005) and also by Stock and Watson (2005) to study the impact of monetary policy shocks on a large panel of macro data) and similar to the structural factor model of Forni, Giannone, Lippi and Reichlin (2007). The paper by Bäurle (2008) is the closest work related to the analysis in this chapter. It offers a method to incorporate the prior information from a DSGE model in estimation of a dynamic factor model and analyzes the impact of the monetary policy shocks on both the factors and selected data series.

traditional macro variables, including measures of real activity, price indices, labor market indicators, interest rate spreads, money and credit stocks, and exchange rates. For instance, contractionary monetary policy realistically leads to a decline in housing starts and in residential investment, to a hump-shaped positive response of the unemployment rate peaking in the 5th quarter after the shock before returning to normal, to the negative rates of commodity price inflation, to a widening of interest rate spreads, to a contraction of consumer credit and to an appreciation of the dollar – despite the fact that our DSGE model does not model these features explicitly.

The chapter is organized as follows. In Section 2 we present the variant of a dynamic factor model and a quick snapshot of the data-rich DSGE model to be used in empirical analysis. Our econometric methodology to estimate two models is discussed in Section 3. Section 4 describes our data set and transformations. In Section 5 we proceed by conducting the empirical analysis. We begin by discussing the choice of the prior distributions of dynamic factor model's parameters. Second, we analyze the estimated empirical factors and the posterior estimates of the DSGE model state variables and explore how well they are able to capture the co-movements in the data. Third, we compare the spaces spanned by the latent empirical factors and by the data-rich DSGE model state variables. Finally, we use the proximity of the factor spaces to propagate the monetary policy and technology innovations in an otherwise non-structural dynamic factor model to obtain the predictions for the macro series of interest. Section 6 concludes.

2 Two Models

In this section, we begin by describing the variant of a dynamic factor model. Then, we present a quick snapshot of the data-rich DSGE model with a New Keynesian core to be estimated on the same large panel of macro and financial series.

2.1 Dynamic Factor Model

We choose to work with the version of the dynamic factor model as originally developed by Geweke (1977) and Sargent and Sims (1977) and recently used by Stock and Watson (2005). If the forecasting performance is a correct guide to choose the appropriate factor model specification, the literature remains rather inconclusive in that respect. For example, Forni, Hallin, Lippi and Reichlin (2003) found supportive results for the generalized dynamic factor specification over the static factor specification, while Boivin and Ng (2005) documented little differences for the competing factor specifications.

Let F_t denote the $N \times 1$ vector of common unobserved factors that are related to a $J \times 1$ large[15] ($J \gg N$) panel of macroeconomic and financial data X_t according to the following factor model:

$$X_t = \mathbf{\Lambda} F_t + e_t \tag{131}$$

$$F_t = \mathbf{G} F_{t-1} + \eta_t, \qquad \eta_t \sim iid\ N(\mathbf{0}, \mathbf{Q}) \tag{132}$$

$$e_t = \mathbf{\Psi} e_{t-1} + v_t, \qquad v_t \sim iid\ N(\mathbf{0}, \mathbf{R}), \tag{133}$$

[15] A typical panel includes from one to two hundred series: e.g. Stock and Watson's (2005) database has $J = 132$, while in Giannone, Reichlin and Sala (2004) $J = 190$. The number of common factors is usually in single digits.

where $\mathbf{\Lambda}$ is the $J \times N$ matrix of factor loadings, e_t is the idiosyncratic errors allowed to be serially correlated, $\mathbf{G}$ is the $N \times N$ matrix that governs common factor dynamics and η_t is the vector of stochastic innovations. The factors and idiosyncratic errors are assumed to be uncorrelated at all leads and lags: $E(F_t e_{i,s}) = 0$, all i, t and s. As in Stock and Watson (2005), we assume that matrices $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{\Psi}$ are diagonal, which implies we have an *exact* dynamic factor model: $E(e_{i,t} e_{j,s}) = 0$, $i \neq j$, all t and s. This is in contrast to the *approximate* DFM of Chamberlain and Rothschild (1983) that relaxes this assumption and allows for some correlation across idiosyncratic errors $e_{i,t}$ and $e_{j,t}$, $i \neq j$. As written, the model is already in static form, since data series X_t load only on contemporaneous factors and not on their lags.[16]

2.2 Data-Rich DSGE Model

The specific version of the data-rich DSGE model that we choose to work with in this chapter is taken from Chapter 1, Section 2.

Its New Keynesian business cycle core features capital as the factor of production, nominal rigidities in price setting, and investment adjustment costs. The real money stock enters households' utility in additively separable fashion. The economy is populated by households, final and intermediate goods-producing firms and a central bank (monetary authority). A representative household works, consumes, saves, holds money balances

[16] In general, a measurement equation is often written as $X_t = \lambda(L) f_t + e_t$, with data loading on current and lagged dynamic factors f_t. However, assuming $\lambda(L)$ has at most p lags, and defining $F_t = (f_t', \ldots, f_{t-p}')'$, we can rewrite it as (131). Here F_t is the vector of static factors as opposed to dynamic factors f_t. To make things simpler, in the model (131)-(133), however, the static and dynamic factors coincide.

and accumulates capital. It consumes the final output manufactured by perfectly competitive final good firms. The final good producers produce by combining a continuum of differentiated intermediate goods supplied by monopolistically competitive intermediate goods firms. To manufacture their output, intermediate goods producers hire labor and capital services from households. Also, when optimizing their prices, intermediate goods firms face the nominal price rigidity a la Calvo (1983), and those firms that are unable to re-optimize may index their price to lagged inflation. Monetary policy is conducted by the central bank setting the one-period nominal interest rate on public debt via a Taylor-type interest rate feedback rule. Given the interest rate, the central bank supplies enough nominal money balances to meet equilibrium demand from households.

In Chapter 1, Section 2 we have shown that if $\boldsymbol{\theta}$ is the vector of deep structural parameters characterizing preferences and technology in our DSGE model and ε_t is the vector of exogenous shocks, then the equilibrium dynamics of the data-rich DSGE model can be summarized by the transition equation of the non-redundant DSGE model state variables S_t:

$$S_t = \mathbf{G}(\boldsymbol{\theta}) S_{t-1} + \mathbf{H}(\boldsymbol{\theta}) \varepsilon_t, \qquad \text{where } \varepsilon_t \sim iid\ N(0, \mathbf{Q}(\boldsymbol{\theta})) \tag{134}$$

and the collection of measurement equations connecting the core macro series X_t^F and the non-core informational macro series X_t^S to the DSGE model states:

$$\underbrace{\begin{bmatrix} X_t^F \\ \hline X_t^S \end{bmatrix}}_{X_t} = \underbrace{\begin{bmatrix} \boldsymbol{\Lambda}_{\mathbf{F}}(\boldsymbol{\theta}) \\ \hline \boldsymbol{\Lambda}_{\mathbf{S}} \end{bmatrix}}_{\boldsymbol{\Lambda}(\boldsymbol{\theta})} S_t + \underbrace{\begin{bmatrix} e_t^F \\ \hline e_t^S \end{bmatrix}}_{e_t}, \tag{135}$$

where the measurement errors e_t may be serially correlated, but uncorrelated across different data indicators ($\mathbf{\Psi}$, $\mathbf{R}$ are diagonal):

$$e_t = \mathbf{\Psi} e_{t-1} + v_t, \qquad v_t \sim iid\ N(\mathbf{0}, \mathbf{R}). \tag{136}$$

Notice that the state-space representation of the data-rich DSGE model (134)-(136) is very much like the dynamic factor model (131)-(133) in which transition of the unobserved factors is governed by a DSGE model solution and where some factor loadings are restricted by the economic meaning of the DSGE model concepts.

3 Econometric Methodology

This section discusses the estimation techniques for the two models considered in this chapter. First, we refer the reader to Chapter 1 on the details about a Markov Chain Monte Carlo algorithm to estimate the data-rich DSGE model, including the choice of the prior for factor loadings. Second, we describe the Gibbs sampler to estimate a dynamic factor model.

3.1 Estimation of the Data-Rich DSGE Model

We refer the reader to Chapter 1, Section 3.1 and Chapter 1's appendices regarding the implementation details of the MCMC algorithm to estimate our data-rich DSGE model.

3.2 Estimation of the Dynamic Factor Model

Consider the original dynamic factor model described in section 2.1:

$$X_t = \mathbf{\Lambda} F_t + e_t \tag{137}$$

$$F_t = \mathbf{G} F_{t-1} + \eta_t, \qquad \eta_t \sim iid\ N(\mathbf{0}, \mathbf{Q}) \tag{138}$$

$$e_t = \mathbf{\Psi} e_{t-1} + v_t, \qquad v_t \sim iid\ N(\mathbf{0}, \mathbf{R}). \tag{139}$$

Let us collect the state-space matrices into $\Gamma = \{\mathbf{\Lambda}, \mathbf{\Psi}, \mathbf{R}, \mathbf{G}\}$ and the latent empirical factors into $F^T = \{F_1, F_2, \ldots, F_T\}$. Similar to the data-rich DSGE model (134)-(136), (137)-(139) is a linear Gaussian state-space model, and we are interested in joint inference about model parameters Γ and latent factors F^T. Unlike in the data-rich DSGE model, though, we no longer have deep structural parameters determining the behavior of matrices in transition equation (138).

We sidestep the problem of a proper dimension of factor space by assuming that $\dim(F_t) = N = 6$, the number of non-redundant model states in the data-rich DSGE model. In contrast, the dynamic factor literature has devoted considerable attention to developing the objective criteria that would determine the proper number of static factors by trading the fit against complexity (Bai and Ng, 2002) and of dynamic factors (e.g., Bai and Ng 2007, Hallin and Liska 2007, Amengual and Watson 2007, Stock and Watson 2005) in DFMs similar to the one above. However, our choice is indirectly supported by the work of Stock and Watson (2005) and Jungbacker and Koopman (2008), who, roughly based on these criteria, find seven dynamic and seven static factors driving a similar panel of macro and financial data.

A principal components analysis of the data set X^T reveals that our choice for the number of factors is not an unreasonable one. As Table F1 demonstrates, the first 6 principal components account for about 75 percent of the variation in the data. The scree plot in Figure F1 shows a very flat slope of the ordered eigenvalues curve when going from the 6th to 7th eigenvalue. Putting in the 7th principal component would add 4.4 percent to the total variance of the data explained, a fairly marginal improvement over the already high cumulative proportion of 75 percent.

Another problem associated with the dynamic factor model (137)-(139) is that the scales and signs of factors F_t and of factor loadings $\mathbf{\Lambda}$ are not separately identified. Regarding scales, take any invertible $N \times N$ matrix $\mathbf{P}$ and notice that the transformed model is observationally equivalent to the original one:

$$X_t = \underbrace{\mathbf{\Lambda P^{-1}}}_{\tilde{\mathbf{\Lambda}}} \underbrace{\mathbf{P}F_t}_{\tilde{F}_t} + e_t \tag{140}$$

$$\underbrace{\mathbf{P}F_t}_{\tilde{F}_t} = \underbrace{\mathbf{PGP^{-1}}}_{\tilde{\mathbf{G}}} \underbrace{\mathbf{P}F_{t-1}}_{\tilde{F}_{t-1}} + \tilde{\eta}_t, \qquad \tilde{\eta}_t \sim iid\ N(\mathbf{0}, \underbrace{\mathbf{PQP'}}_{\tilde{\mathbf{Q}}}) \tag{141}$$

Regarding signs, for the moment think of (137)-(139) as a model with only one factor. Then multiply by -1 the transition equation (138), as well as the factor loading and the factor itself in measurement equation (137). We obtain the new model, yet it is observationally equivalent to the original.

We follow the factor literature (e.g. Geweke and Zhu 1996; Jungbacker and Koopman 2008) and make the following normalization assumptions to tell factors apart from factor loadings: (i) set $\mathbf{Q} = \mathbf{I}_N$ to fix the scale of factors; (ii) require one loading in $\mathbf{\Lambda}$ to be positive for each factor (sign restrictions); and (iii) normalize some factor loadings in $\mathbf{\Lambda}$ to pin down specific factor rotation.

Denote by $\mathbf{\Lambda}_1$ the upper $N \times N$ block of $\mathbf{\Lambda}$ so that $\mathbf{\Lambda} = \left[\mathbf{\Lambda}_1'; \mathbf{\Lambda}_2'\right]'$. One way to implement (ii) and (iii) would be to assume that $\mathbf{\Lambda}_1$ is lower triangular (i.e., $\lambda_{ij} = 0$ for $j > i$, $i = 1, 2, ..., N-1$) with strictly positive diagonal $\lambda_{ii} > 0$, $i = \overline{1, N}$ (see Harvey 1989, p.451). However, our data set in estimation, to be described later in the section Data, will consist of *core* and *non-core* macro and financial series. Furthermore, within the core series we will have four blocks of variables: real output, inflation, the nominal interest rate and the inverse velocity of money, respectively; each block contains several measures of the same concept. For example, the output block comprises real GDP, total industrial production and industrial production in the manufacturing sector; the inflation block includes GDP deflator inflation, CPI inflation and personal consumption expenditures inflation. For this reason, we choose another alternative to implement normalizations (ii) and (iii) – the block-diagonal scheme that to some degree exploits the group structure of the core series in data X_t:

	F_1	F_2	F_3	F_4	F_5	F_6
Real output #1	1	1	+1	0	0	0
Real output #2	1	+1	1	0	0	0
Real output #3	1	1	1	0	0	0
Inflation #1	1	1	0	1	0	0
Inflation #2	+1	1	0	1	0	0
Inflation #3	1	1	0	1	0	0
Interest rate #1	1	1	0	0	+1	0
Interest rate #2	1	1	0	0	1	0
Interest rate #3	1	1	0	0	1	0
IVM #1	1	1	0	0	0	1
IVM #2	1	1	0	0	0	+1
IVM #3	1	1	0	0	0	1
$X^{non-core}$	1	1	1	1	1	1

(142)

where 1s stand for non-zero elements in $\mathbf{\Lambda}$.

We acknowledge that our block-diagonal scheme imposes some overidentifying restrictions on factor loadings beyond those minimally necessary. However, scheme (142) can also be interpreted as a special case of the appealing dynamic hierarchical factor model of Moench, Ng, and Potter (2008), which – on top of aggregate common factors – introduces intermediate block factors and makes use of the block structure of the data.

Now, to estimate the model (137)-(139) under normalizing assumptions (i)-(iii), we again apply the Bayesian MCMC methods as in the estimation of the data-rich DSGE model (Chapter 1, Section 3.1). We construct a Gibbs sampler that iterates on a complete set of known conditional posterior densities to generate draws from the joint posterior distribution $p(\Gamma, F^T \mid X^T)$ of model parameters $\Gamma = \{\mathbf{\Lambda}, \mathbf{\Psi}, \mathbf{R}, \mathbf{G}\}$ and latent factors F^T:

$$p(F^T \mid \Gamma; X^T) \propto p(F^T \mid \Gamma) p(X^T \mid \Gamma, F^T) \tag{143}$$

$$p(\Gamma \mid F^T; X^T) \propto p(\Gamma) p(F^T \mid \Gamma) p(X^T \mid \Gamma, F^T) \tag{144}$$

The main steps of the Gibbs sampler are:

1. Specify initial values $\Gamma^{(0)}$ and $F^{T,(0)}$.
2. Repeat for $g = 1, 2, ..., n_{sim}$

   2.1. Generate latent factors $F^{T,(g)}$ from $p(F^T \mid \Gamma^{(g-1)}; X^T)$ using the Carter-Kohn (1994) forward-backward algorithm;

   2.2. Generate state-space parameters $\Gamma^{(g)}$ from $p(\Gamma \mid F^{T,(g)}; X^T)$ by drawing from a complete set of known conditional densities.
3. Return $\left\{\Gamma^{(g)}, F^{T,(g)}\right\}_{g=1}^{n_{sim}}$

Compared to the MCMC algorithm for the data-rich DSGE model, this Gibbs sampler is easier and it differs in two key respects: (i) we no longer have the complicated Metropolis step, since there are no deep structural parameters $\boldsymbol{\theta}$ coming from the economic model; and (ii) inside Γ, we have to draw matrix $\mathbf{G}$ from the transition equation of factors (in the data-rich DSGE model it was pinned down by numerical solution of a DSGE model given structural parameters $\boldsymbol{\theta}$).

To draw the latent factors F^T from $p(F^T \mid \Gamma; X^T)$, we use the familiar Carter-Kohn (1994) machinery. First, we apply the Kalman filter to the linear Gaussian state-space system (137)-(139) to generate filtered latent factors $\hat{F}_{t|t}$, $t = \overline{1,T}$. Then, starting from $\hat{F}_{T|T}$, we roll back in time along the Kalman smoother recursions and generate

$F^T = \{F_1, F_2, \ldots, F_T\}$ by recursively sampling from a sequence of conditional Gaussian distributions.

To sample from the conditional posterior $p(\Gamma \mid F^T; X^T)$, we notice the following: with diagonality of matrices $\mathbf{\Psi}$ and $\mathbf{R}$ and conditional on factors F^T, (137) and (139) are a set of standard multivariate linear regressions with AR(1) errors and Gaussian innovations ($k = \overline{1, J}$):

$$X_{k,t} = \mathbf{\Lambda}'_k F_t + e_{k,t}, \qquad e_{k,t} = \psi_{kk} e_{k,t-1} + v_{k,t}, \qquad v_{k,t} \sim iid\ N(0, R_{kk}). \qquad (145)$$

Hence, under the conjugate prior $p(\mathbf{\Lambda}, \mathbf{\Psi}, \mathbf{R})$, we can apply the insight of Chib and Greenberg (1994) to derive the conditional posteriors $\left[\mathbf{R} \mid (\mathbf{\Lambda}, \mathbf{\Psi}); \mathbf{G}, F^T, X^T\right]$, $\left[\mathbf{\Lambda} \mid (\mathbf{R}, \mathbf{\Psi}); \mathbf{G}, F^T, X^T\right]$, $\left[\mathbf{\Psi} \mid (\mathbf{\Lambda}, \mathbf{R}); \mathbf{G}, F^T, X^T\right]$ and to sample accordingly.

What remains to be drawn is the transition matrix $\mathbf{G}$. Given factors F^T, the conditional posterior $p(\mathbf{G} \mid (\mathbf{\Lambda}, \mathbf{R}, \mathbf{\Psi}); F^T, X^T)$ can be derived from a VAR(1) in (138):

$$F_t = \mathbf{G} F_{t-1} + \eta_t, \qquad \eta_t \sim iid\ N(\mathbf{0}, \mathbf{I}_N). \qquad (146)$$

We assume the so-called Minnesota prior (Doan, Litterman and Sims, 1984; the specific version comes from Lubik and Schorfheide, 2005) on transition matrix $\mathbf{G}$ and truncate it to the region consistent with the stationarity of (146). We implement our prior by a set of dummy observations that tilt the VAR to a collection of univariate random walks (details are in Appendix E).

To estimate the empirical DFM, in the actual implementation of the Gibbs sampler we have applied the Jungbacker-Koopman computational speed-up presented in

Chapter 1, Section 3.2 and already utilized to improve the speed of computations in the data-rich DSGE model's estimation. We find that the "improved" estimation of the empirical DFM runs 10.5 times faster than the no-speedup estimation, a magnitude consistent with the CPU gains reported by Jungbacker and Koopman (2008) for a DFM of a similar size in their study.

4 Data

To estimate the dynamic factor model and the data-rich DSGE model, we employ a large panel of U.S. quarterly macroeconomic and financial time series compiled by Stock and Watson (2008). The panel covers 1959:Q1 – 2006:Q4, however, our sample in this chapter is restricted only to 1984:Q1 – 2005:Q4 so as to avoid dealing with the issue of the Great Moderation and to concentrate on a period with a relatively stable monetary policy regime.

Our data set is identical to the one employed in Chapter 1 and consists of 12 *core series* that either measure specific DSGE model concepts or are used in the DFM normalization scheme (142), and 77 *non-core* informational series that load on all DSGE states (DFM factors) and may contain useful information about the aggregate state of the economy. The core series include three measures of real output (real GDP, the index of total industrial production and the index of industrial production: manufacturing), three measures of price inflation (GDP deflator inflation, personal consumption expenditure (PCE) deflator inflation, and CPI inflation), three indicators of the nominal interest rates (the federal funds rate, the 3-month T-bill rate and the yield on AAA-rated corporate bonds), and three series measuring the inverse velocity of money (IVM based on the M1 aggregate and the M2 aggregate and IVM based on the adjusted monetary base). The 77 non-core series include the measures of real activity, labor market variables, housing indicators, prices and wages, financial variables (interest rate spreads, exchange rate depreciations, credit stocks, stock returns) and, together with appropriate transformations to eliminate trends, are described in Chapter 1, Appendix C. To save space, we refer the reader to Chapter 1, Section 4 that describes in detail the construction of all data indicators included in our data set.

Because measurement equations (135) and (137) are modeled without intercepts, we estimate a dynamic factor model and a data-rich DSGE model on a demeaned data set. Also, in line with standard practice in the factor literature, we standardize each time series so that its sample variance is equal to unity (however, we do not scale the core series when estimating the data-rich DSGE model).

5 Empirical Analysis

The next step in our analysis is to take a dynamic factor model and a data-rich DSGE model to the data using the MCMC algorithms described above and to present the empirical results. We begin by discussing the choice of the prior distributions of dynamic factor model's parameters. Second, we analyze the estimated empirical factors and the estimates of the DSGE model state variables and explore how well they are able to capture the co-movements in the data. Third, we compare the spaces spanned by the latent empirical factors and by the data-rich DSGE model state variables. Finally, we use the proximity of the factor spaces to propagate the monetary policy and technology innovations in an otherwise non-structural dynamic factor model and obtain the predictions from both models for the core and non-core macro and financial series of interest.

5.1 Priors and Posteriors

Since we estimate the DFM (137)-(139) and the data-rich DSGE model (134)-(136) using Bayesian techniques, we have to provide prior distributions for both models' parameters.

Let us first turn to a dynamic factor model. Let $\mathbf{\Lambda}_k$ and R_{kk} be the factor loadings and a variance of the measurement error innovation for the k^{th} measurement equation, $k = 1..J$. Similarly to Boivin and Giannoni (2006) and Kose, Otrok and Whiteman (2008), we assume a joint Normal-InverseGamma prior distribution for $\left(\mathbf{\Lambda}_k, R_{kk}\right)$ so that $R_{kk} \sim IG_2(s_0, \nu_0)$ with location parameter $s_0 = 0.001$ and degrees of freedom $\nu_0 = 3$, and the prior mean of factor loadings is centered around the vector of zeros $\mathbf{\Lambda}_k \mid R_{kk} \sim N(\mathbf{\Lambda}_{k,0}, R_{kk}\mathbf{M}_0^{-1})$ with $\mathbf{\Lambda}_{k,0} = \mathbf{0}$ and $\mathbf{M}_0 = \mathbf{I}_N$. The prior for the k^{th} measurement equation's autocorrelation Ψ_{kk}, all k, is $N(0,1)$. We are making it perfectly tight, however, because there could be data series with stochastic trends we seek to capture with potentially highly persistent dynamic factors and not with highly persistent measurement errors. This implies that all measurement errors are *iid* mean-zero normal random variables. Finally, as explained in Section 3.2, for the factor transition matrix $\mathbf{G}$, we implement a version of a Minnesota prior (Lubik and Schorfheide, 2005) and tilt the transition equation (138) to a collection of univariate random walks.[17]

In our data-rich DSGE model, we have two groups of parameters: state-space model parameters comprising matrices $\mathbf{\Lambda}$, $\mathbf{\Psi}$ and $\mathbf{R}$, and deep structural parameters $\boldsymbol{\theta}$ of an underlying DSGE model. The prior for the state-space matrices is elicited differently for the core and the non-core data indicators contained in X_t. Regarding the non-core measurement equations, the prior for $\left(\mathbf{\Lambda}_k, R_{kk}\right)$ and for Ψ_{kk} is identical to the one assumed in DFM. The prior distribution for the factor loadings in the core measurement equations follows the same scheme as elaborated in Chapter 1, Section 5.1. Our choice of prior distribution for the deep structural parameters of a DSGE model is exactly identical to the one presented in Chapter 1, Section 5.1.

We use the Gibbs sampler presented above in section 3.2 and the Gibbs sampler with Metropolis step outlined in Chapter 1, Section 3.1 to estimate our empirical dynamic factor model and the data-rich DSGE model, respectively. The only parameters of direct interest are the deep structural parameters $\boldsymbol{\theta}$ of an underlying DSGE model, and we have already discussed them extensively in Chapter 1. We do not discuss the posterior estimates of DFM parameters here either, since we are more interested in comparing factor spaces spanned by the estimated latent factors and by the DSGE model states. However, all the parameter estimates are collected in the technical appendix to this chapter, which is available upon request.

---

[17] The hyperparameters in the actual implementation of the Minnesota prior were set as follows: $\tau = 5$, $d = 0.5$, $\iota = 1$, $w = 1$, $\lambda = 0$, $\mu = 0$. We have also truncated the prior to the region consistent with the stationarity of the factor transition equation.

5.2 Empirical Factors and Estimated DSGE Model States

Our empirical analysis proceeds by plotting the estimated empirical factors extracted by a dynamic factor model and the estimated DSGE state variables from our data-rich DSGE model.

Figure D1 (Chapter 1) depicts the posterior means and 90 percent credible intervals of the estimated data-rich DSGE model states. These include three endogenous variables (model inflation $\hat{\pi}_t$, the nominal interest rate $\hat{R}_t$ and real household consumption $\hat{X}_t$) and three structural AR(1) shocks (government spending g_t, money demand χ_t and neutral technology Z_t). In Chapter 1 we have noted four observations. First, all three structural disturbances exhibit large swings and prolonged deviations from zero capturing the persistent low-frequency movements in the data. Second, the estimated data-rich DSGE model states are much *smoother* than their counterparts in the regular DSGE model, because in the data-rich context, the model states are the common components of a large panel of data, and they have to capture well not only a few core macro series (as is the case in the regular DSGE model), but also very many non-core informational series. The third observation is that the money demand shock χ_t appeared to be very different in the data-rich versus the regular DSGE model estimation, owing primarily to the fact that in the data-rich DSGE model it helped explain housing variables, consumer credit and non-GDP measures of output at the cost of the poorer fit for the IVM_M2S. The fourth observation was a counterfactual behavior of government spending shock and real consumption during recessions: the former tended to fall and the latter to rise when times are bad.

We proceed by discussing the latent empirical factors extracted by our DFM from the same rich data set. Figure F2 plots the posterior means and 90 percent credible intervals of the estimated factors. First, note that unlike the DSGE model states, these factors have in general *no economic interpretation*. This is less true of factors F3-F6, because of the assumed normalization scheme (142). Second, while factors 3 and 5 indeed look much like the data on real output and nominal interest rate, factors 4 and 6 – despite the normalization – do not. This shows that the exclusion normalizations favoring a certain ex-ante meaning of a particular factor are not a sufficient condition to guarantee this meaning ex-post after estimation. The third observation is that the credible intervals for F1 and F2 – the latent factors common to all macro and financial series in the panel – are not uniformly wide or narrow, as is more or less the case for factors F3-F6. During several years prior to 1990-91 recession, the 90 percent credible bands for factor F1 expand, and then quickly shrink after recession is over. The same pattern is observed for factor F2 for several years preceding the 2001 recession. One interpretation of this finding could be that the volatility of these two factors is not constant over time and follows a regime-switching dynamics over the business cycle. Clearly, to have a stronger case, one might like to estimate a DFM on the full postwar sample of available U.S. data.

5.3 How Well Factors Trace Data

Let us now turn to the question of how well the factors and the DSGE states are able to trace the actual data. *A priori* we should expect that the unrestricted dynamic factor model will do a better job on that dimension than the data-rich DSGE model whose cross-equation restrictions might be misspecified and the factor loadings in which might be unduly restricted. And that's indeed what we find and what can be concluded from inspecting Table F2 and Table F3 which present the (posterior mean of) fraction of the unconditional variance of the data series captured by the empirical factors and by the DSGE model states. On average, the data-rich DSGE model states "explain" about 75 percent of variance for the core macro series and 72 percent of variance for the non-core. The latent empirical factors extracted by a DFM are able to account for 95 and 94 percent of the variance for the core and non-core series, respectively. So overall, the empirical factors capture more than the DSGE states.

More specifically, within the core series it is the measures of inflation and of inverse money velocities that are traced relatively more poorly than the real output and nominal interest rates in both models. The same picture is observed in the non-core block of series: price and wage inflation measures and the financial variables in both models tend to have a higher fraction of unconditional variance due to measurement errors. In the data-rich DSGE model, the state variables capture about 15 to 25 percent of the variance in exchange rate depreciations and stock returns, but about 65 to 85 percent of the variance of interest rate spreads and credit stocks. This is not surprising given that our theoretical model does not have New Open-Economy Macroeconomics mechanisms (e.g., Lubik and Schorfheide, 2005 or Adolfson, Laseén, Linde, Villani, 2005, 2008) and does not feature financial intermediation (e.g., Bernanke, Gertler, Gilchrist, 1999). In the dynamic factor model, these percentages are much higher: the latent factors explain about

97-98 percent of the variance of the interest spreads and credit stocks, about 65-82 percent of the variability in exchange rate depreciations and 80-82 percent of stock returns (Table F4). This suggests that our DSGE model is potentially misspecified along this "financial" dimension.

5.4 Comparing Factor Spaces

Up to this point, we have done two things: (i) we have estimated the empirical latent factors in a dynamic factor model and the DSGE states in a data-rich DSGE model; and (ii) we have established that both factors and DSGE states are able to explain a significant portion of the co-movement in the rich panel of U.S. macro and financial series. From Figure D1 (Chapter 1) and Figure F2 we have learned that the states and the factors look quite different; therefore now we come to our central question: can the empirical factors and the estimated DSGE model state variables span the same factor space? Or, in other words, can we predict the true estimated DFM latent factors from the DSGE model states with a fair amount of accuracy?

Let $F_t^{(pm)}$ and $S_t^{(pm)}$ denote the posterior means of the empirical factors and of the data-rich DSGE model state variables. For each latent factor $F_{i,t}^{(pm)}$, we estimate, by Ordinary Least Squares, the following simple linear regression:

$$F_{i,t}^{(pm)} = \beta_{0,i} + \boldsymbol{\beta}'_{1,i} S_t^{(pm)} + u_{i,t} \tag{147}$$

with mean zero and homoscedastic error term $u_{i,t}$. We report the R^2 s for the collection of linear predictive regressions (147) in Table F7. Denoting the OLS estimates by

$\hat{\boldsymbol{\beta}}_0 = [\beta_{0,1},..., \beta_{0,N}]'$ and by $\hat{\boldsymbol{\beta}}_1 = [\boldsymbol{\beta}_{1,1},..., \boldsymbol{\beta}_{1,N}]'$, we then construct the predicted empirical factors $\hat{F}_t^{(pm)}$:

$$\hat{F}_t^{(pm)} = \boldsymbol{\beta}_0 + \hat{\boldsymbol{\beta}}_1 S_t^{(pm)} \tag{148}$$

The Figure F3 overlays true estimated DFM factors $F_t^{(pm)}$ versus those predicted by the DSGE states $\hat{F}_t^{(pm)}$.

From both Table F7 and Figure F3 we can clearly conclude that the DSGE states predict empirical factors really well and therefore the factor spaces spanned by the DSGE model state variables and by the DFM latent factors are very closely aligned. What are the implications of this important finding? First, this implies that a DSGE model indeed captures the essential sources of co-movement in the large panel of data as a sort of a core and that the differences in fit between a data-rich DSGE model and a DFM are potentially due to restricted factor loadings in the former. Second, this also implies a greater degree of comfort about propagation of structural shocks to a wide array of macro and financial series – which is the essence of many policy experiments. Third, the proximity of factor spaces facilitates economic interpretation of a dynamic factor model, as the empirical factors are now isomorphic – through the link (148) – to the DSGE model state variables with clear economic meaning.

5.5 Propagation of Monetary Policy and Technology Innovations

The final and the most appealing implication of the factor spaces proximity in the two models is that it allows us to map the DSGE model state variables into DFM empirical factors every period and therefore propagate any structural shocks from the DSGE model in an *otherwise completely non-structural* dynamic factor model to obtain predictions for a broad range of macro series of interest. Suppose $\mathbf{\Lambda}^{dfm-dsge}$ and $\mathbf{\Lambda}^{dfm}$ denote the posterior means of factor loadings in the data-rich DSGE model (134)-(136) and in the empirical DFM (137)-(139), respectively. Then, for any structural shock $\varepsilon_{i,t}$, we can generate two sets of impulse responses of a large panel of data X_t:

$$\left(\frac{\partial X_{t+h}}{\partial \varepsilon_{i,t}}\right)_{dfm-dsge} = \mathbf{\Lambda}^{dfm-dsge} \times \frac{\partial S_{t+h}}{\partial \varepsilon_{i,t}} \tag{149}$$

$$\left(\frac{\partial X_{t+h}}{\partial \varepsilon_{i,t}}\right)_{dfm} = \mathbf{\Lambda}^{dfm} \times \frac{\partial F_{t+h}}{\partial \varepsilon_{i,t}} = \mathbf{\Lambda}^{dfm}\left[\hat{\boldsymbol{\beta}}_1 \frac{\partial S_{t+h}}{\partial \varepsilon_{i,t}}\right], \tag{150}$$

where $\partial S_{t+h} / \partial \varepsilon_{i,t}$ is computed from the transition equation of the data-rich DSGE model for every horizon $h = 0,1,2,...$ and where we have used the link between S_t and F_t determined by (148).

In what follows we focus on propagating monetary policy $(\varepsilon_{R,t})$ and technology $(\varepsilon_{Z,t})$ innovations in both the data-rich DSGE and the dynamic factor model to generate predictions for the core and non-core macro series. The corresponding impulse response functions (IRFs) are presented in Figure F4, Figure F5, Figure F6 and Figure F7. It is natural to compare our results to findings in two strands of the literature: Factor Augmented Vector Autoregression (FAVAR) literature (e.g. Bernanke, Boivin, Eliasz, 2005; Stock and Watson, 2005) and the regular DSGE literature (e.g. Christiano, Eichenbaum, Evans, 2005; Smets and Wouters, 2003, 2007; DSSW 2007; Aruoba and Schorfheide, 2009; Adolfson, Laseén, Linde, and Villani, 2008). In FAVAR studies, we are able to obtain predictions for a rich panel of U.S. data similar to ours, but only of the monetary policy innovations. In the regular DSGE literature, one can propagate any structural shocks including monetary policy and technology innovations, but to a limited number of core macro variables (e.g., real GDP, consumption, investment, inflation, the interest rate, the wage rate and hours worked in Smets and Wouters, 2007). The framework that we propose in this chapter is able to deliver on both fronts: we are able to compute the responses of the core and non-core variables to both monetary policy and technology shocks. Moreover, we will have two sets of responses: from the data-rich DSGE model, which might be misspecified, and from the dynamic factor model that is primarily data-driven and fits better.

At least from the perspective of monetary policy innovations, we tend to favor the predictions obtained from the empirical dynamic factor model (150). It turns out (we provide evidence below) that the two models' predictions for the non-core variables are fairly close. The responses of the core series, though, seem more plausible in the empirical DFM case, since, for example, channeling the shock through the DFM helps eliminate the puzzling behavior of price inflation observed in the data-rich DSGE model context that we have documented in Chapter 1, Section 5.5.

One general observation from comparing IRFs should be emphasized from the very beginning. The responses of core variables like real GDP, real consumption and investment, and inflation in regular DGSE studies are often hump-shaped, matching well the empirical findings from identified VARs. Our IRFs do not have many humps, because the underlying theoretical DSGE model, as presented in Chapter 1, Section 2.2, abstracts from, say, habit in consumption or variable capital utilization – mechanisms that help get the humps in those often more elaborate models. This, however, can be fixed by replacing the present DSGE model with a more elaborate one.

Let us turn first to the *effects of monetary policy innovation,* which are summarized in Figure F4 and Figure F5. A contractionary monetary policy shock corresponds to 0.75 percent (or 75 basis points) increase in the federal funds rate. As the nominal policy rate rises and the opportunity costs of holding money for households increase, we observe a strong liquidity effect associated with falling real money balances. Also, high interest rates make the saving motive and buying more bonds temporarily a more attractive option. This raises households' marginal utility of consumption and discourages current spending in favor of the future consumption. Because the household faces investment adjustment costs and cannot adjust investment quickly, and government spending in the model is exogenous, the lower consumption leads to a fall in aggregate demand. The firms respond to lower demand in part by contracting real output and in part by reducing the optimal price. Hence, the aggregate price level falls, but not as much given nominal rigidities in the intermediate goods-producing sector.

Why do the monopolistically competitive firms respond to falling demand in part by charging a lower price? The short answer is that because they are able to cut their marginal costs. On the one hand, higher interest rates inhibit investment and the return on capital is falling. On the other hand, firms may now economize on real wages. The market for labor is perfectly competitive, since we assume no wage rigidities. This implies that the real wage is equal to the marginal product of labor, but also that it is equal to the household's marginal rate of substitution between consumption and leisure, as in (78). Since the disutility of labor in our model is fixed, and the marginal utility of consumption is higher, the household accepts lower real wage and the firms are able to pass on their losses in revenues to households by reducing their own wage bills.

Now given lower marginal costs, the New Keynesian Phillips curve suggests we should observe falling aggregate prices and negative rates of inflation (in terms of a deviation from the steady-state inflation). That's what we see in the second column of Figure F4. Notice that channeling the monetary policy shock through the pure dynamic factor model helps correct the so-called "*price puzzle*"[18] for the data-rich-DSGE-model-implied responses of PCE deflator inflation and CPI inflation. Interestingly, a positive response of CPI inflation to a monetary policy contraction is also documented in Stock and Watson (2005), despite the fact that they use a data-rich Factor Augmented VAR. It has been argued (e.g., Bernanke, Boivin and Eliasz, 2005) that the rich information set helps eliminate this sort of anomaly.

As can be seen from the first column of Figure F4, the response of industrial production (IP) to the monetary policy tightening seems counterfactual compared to FAVAR findings (we have documented this finding in Chapter 1, Section 5.5 too). First, this may have something to do with the inherent inertia of IP in responding to monetary policy. It continues to be driven by excessive optimism from the previous phase of the business cycle and it takes time to adjust to new conditions. But once IP falls below the

[18] "Price puzzle" (Sims, 1992) refers to the counterfactual finding in the VAR literature that a measure of prices or inflation responds positively to a contractionary monetary policy shock associated with an unexpected increase in the policy interest rate.

trend, it remains subdued for a long time. Second, this may have something to do with the way the monetary policy shock is identified in the FAVAR literature. By construction, in a Factor Augmented VAR the industrial production is contained in the list of "slow moving" variables, and the identification of the monetary policy shock is achieved by postulating that it does not affect slow variables contemporaneously. Regarding the responses of real GDP, we document that the data-rich DSGE and DFM models disagree about the magnitude of the contraction. The DFM-implied response is almost negligible implying that the costs of disinflation are very small (which is hard to believe), whereas the data-rich-DSGE-model-implied response is about minus 0.5 percent – hump shape aside, a value in the ballpark of findings in the regular DSGE literature.

If we look at the effects of the monetary policy tightening on non-core macro and financial variables (Figure F5), they complete the picture for the core series with details. The real activity measures, such as real consumption of durables, real residential investment and housing starts, broadly decline. Prices go down as well; in particular, we observe negative rates of commodity price inflation and investment deflator inflation. The measures of employment fall (e.g., employment in the services sector) indicating tensions in the labor market, while unemployment gains momentum with a lag before eventually returning to normal. The interest rate spreads (for instance, the 6-month over the 3-month Treasury bill rate) widen considerably, reflecting tighter money market conditions and increased liquidity risks and credit risks. Consumer credit is contracted, in part due to lower demand from borrowers facing higher interest rates and in part owing to the reduced availability of funds. The dollar appreciates, reflecting intensified capital inflows lured by higher returns in the domestic financial market. As a result, both export and import price indices fall, thereby translating – according to the magnitudes in Figure F5 – into a deterioration of the U.S. terms of trade.

Broadly speaking, the reported results are qualitatively very similar to the FAVAR findings of Bernanke, Boivin and Eliasz (2005) and Stock and Watson (2005). Except for the humps, they also accord well with the monetary policy effects on the core variables documented in the regular DSGE literature. On top of that, the responses of the non-core variables seem to provide a reasonable and consistent picture of monetary tightening as well.

We plot the *effects of a positive technology innovation* in Figure F6 (core series) and Figure F7 (non-core series). Following the positive TFP shock, real output broadly increases (although there is a disagreement between the DFM and the data-rich DSGE model as to the response of real GDP), as our economy becomes more productive and the firms find it optimal to produce more. New demand comes primarily from higher capital investment, reflecting much better future return on capital, and also from additional household consumption fueled by greater income. The higher output on the supply side plus improved efficiency implies a downward pressure on prices. Through the lenses of the New Keynesian Phillips curve, the current period inflation is positively related to expected future inflation and to current marginal costs. A positive technology shock has raised production efficiency and reduced the current marginal costs (the elevated real wage resulting from increased labor demand was not enough to prevent that). However, because technology innovation is very persistent, the firms expect future marginal costs and thus future inflation to be lower as well. This anticipation effect, coupled with currently low marginal costs, leads to prices falling now, as is evident from column 2 of the Figure F6.

The increase in real output above steady state and the fall of inflation below target level, under the estimated monetary policy Taylor rule, requires the Fed to move the policy rate in opposite directions. The fact that the Fed actually lowers the policy rate means that the falling prices effect dominates, with other interest rates following the course of the federal funds rate (column 3, Figure F6). Declining interest rates boost real output even more, which in turn raises further the return on capital. As the positive impact of technological innovation dissipates, this higher return, through the future marginal costs channel, fuels inflationary expectations that ultimately translate into contemporaneous upward price pressures. The Fed reacts by increasing the policy rate, which explains the observed hump in the interest rate IRF. Given temporarily lower interest rates, households choose to hold, with some lag, relatively higher real money balances (from column 4, Figure F6, this applies more to M1S and the monetary base, and less to the M2S aggregate that comprises a hefty portion of interest-bearing time deposits). A part of the growing money demand comes endogenously from the elevated level of economic activity.

These results – both in terms of the magnitudes and shapes of responses – align fairly closely with findings in the regular DSGE literature (e.g., Smets and Wouters, 2007; Aruoba, Schorfheide, 2009; and DSSW 2007).

The responses of the non-core macroeconomic series (Figure F7) appear to enrich the story for core variables with additional details. Following a positive technology innovation, the subcomponents of real GDP (real consumption of durables, real residential investment) or the components of industrial production (e.g., production of business equipment) generally expand (although there is weaker agreement between the predictions of the DFM and the data-rich DSGE model). Measures of employment (e.g., employment in the services sector) increase. However, this stands in contrast to the results in Smets and Wouters (2003) and Adolfson, Laseén, Linde, Villani (2005), who find in European data that employment actually falls after a positive stationary TFP shock. As marginal costs fall, commodity price inflation (P_COM) and investment deflator inflation (PInv_GDP) follow the overall downward price pressures trend. The interest rate spreads (SFYGM6) shrink, in part reflecting the lower level of perceived risks, while credit conditions ease, leading to growth in business loans. Despite the interest rates being below average for a prolonged period of time, the dollar appreciates, but by less than after the monetary tightening. Finally, the real wage (RComp_Hour) increases, while average hours worked (Hours_AVG) decline. The rise in the real wage and the initial fall in hours worked are in line with evidence documented by Smets and Wouters (2007). However, the subsequent dynamics of hours are quite different: in Smets and Wouters the hours turn significantly positive after about two years. Here they stay below steady state for much longer. This may have something to do with a greater amount of persistence in the technology process in our model.

6 Conclusions

In this chapter, we have compared a data-rich DSGE model with a standard New Keynesian core to an empirical dynamic factor model by estimating both on a rich panel of U.S. macroeconomic and financial indicators compiled by Stock and Watson (2008). We have established that the spaces spanned by the empirical factors and by the data-rich DSGE model states are very closely aligned.

This key finding has several important implications. First, this finding implies that a DSGE model indeed captures the essential sources of co-movement in the data and that the differences in fit between a data-rich DSGE model and a DFM are potentially due to restricted factor loadings in the former. Second, it also implies a greater degree of comfort about the propagation of structural shocks to a wide array of macro and financial series. Third, the proximity of factor spaces facilitated economic interpretation of a dynamic factor model, since the empirical factors have become isomorphic to the DSGE model state variables with clear economic meaning.

Most important, the proximity of factor spaces in the two models has allowed us to propagate the monetary policy and technology innovations in an otherwise completely non-structural dynamic factor model to obtain predictions for many more series than just a handful of traditional macro variables, including measures of real activity, price indices, labor market indicators, interest rate spreads, money and credit stocks, and exchange rates. The responses of these non-core variables therefore provide a more complete and comprehensive picture of the effects of monetary policy and technology shocks and may serve as a check on the empirical plausibility of a DSGE model.

Appendix E. DFM: Gibbs Sampler: Drawing Transition Equation Matrix

We need to generate $\mathbf{G}$ from the conditional density $p(\mathbf{G} \mid \mathbf{Q}, \mathbf{\Lambda}, \mathbf{\Psi}, \mathbf{R}, F^T; X^T)$. Note, however, that the dependence of $\mathbf{G}$ on the other state-space matrices – except for $\mathbf{Q}$ – is exclusively through the factors. This is because given factors F_t, the transition equation (138) is a VAR(1):

$$F_t = \mathbf{G}F_{t-1} + \eta_t, \qquad \eta_t \sim iid\ N(\mathbf{0}, \mathbf{Q}), \qquad t = 1, ..., T . \tag{151}$$

Therefore, $p(\mathbf{G} \mid \mathbf{Q}, \mathbf{\Lambda}, \mathbf{\Psi}, \mathbf{R}, F^T; X^T) = p(\mathbf{G} \mid \mathbf{Q}, F^T)$.

Rewrite the VAR in matrix notation

$$Y = X\mathbf{G} + \eta \tag{152}$$

where Y, X and η are the $(T-1) \times N$ matrices with rows F_t', F_{t-1}' and η_t', respectively. To specify a prior distribution for the VAR parameters, we follow Lubik and Schorfheide (2005) and use a version of Minnesota Prior (Doan, Litterman, Sims 1994) implemented with T^* dummy observations Y^* and X^*. The likelihood function of dummy observations $p(Y^* \mid \mathbf{G}, \mathbf{Q})$ combined with the improper prior distribution $|\mathbf{Q}|^{-(N+1)/2} \times \mathbf{1}_{\mathbf{G}}$ induces the proper prior for the VAR parameters:

$$p(\mathbf{G}, \mathbf{Q}) \propto p(Y^* \mid \mathbf{G}, \mathbf{Q}) |\mathbf{Q}|^{-(N+1)/2} \times \mathbf{1}_{\mathbf{G}}, \tag{153}$$

where $\mathbf{1}_{\mathbf{G}}$ denotes an indicator function equal to 1 if all eigenvalues of $\mathbf{G}$ lie inside unit circle. In actual implementation of Minnesota Prior, we set the hyperparameters as follows $\tau = 5,\ d = 0.5,\ \iota = 1,\ \ w = 1,\ \lambda = 0,\ \mu = 0$ to generate Y^* and X^*. Essentially, our prior is tilting the transition equation (151) to a collection of the univariate random walks.

Combining this prior with the likelihood function $p(Y \mid \mathbf{G}, \mathbf{Q})$, we obtain the posterior density of the VAR parameters:

$$p(\mathbf{G}, \mathbf{Q} \mid Y) \propto p(Y \mid \mathbf{G}, \mathbf{Q})\, p(\mathbf{G}, \mathbf{Q}) = p(Y \mid \mathbf{G}, \mathbf{Q})\, p(Y^* \mid \mathbf{G}, \mathbf{Q}) \left| \mathbf{Q} \right|^{-(N+1)/2} \times \mathbf{1}_{\mathbf{G}}. \quad (154)$$

It can be shown (e.g. Del Negro, Schorfheide 2004) that our posterior density $p(\mathbf{G}, \mathbf{Q} \mid Y) = p(\mathbf{G}, \mathbf{Q} \mid F^T)$ is truncated Normal-Inverse-Wishart:

$$\mathbf{Q} \mid Y \quad \sim IW(\tilde{\mathbf{Q}}, (T + T^* - N)) \quad (155)$$

$$\mathbf{G} \mid \mathbf{Q}, Y \quad \sim N(\tilde{\mathbf{G}}, \mathbf{\Sigma}_G) \times \mathbf{1}_{\mathbf{G}} \quad (156)$$

where

$$\tilde{\mathbf{G}} = \left(X^{*\prime} X^* + X'X \right)^{-1} \left(X^{*\prime} Y^* + X'Y \right)$$

$$\tilde{\mathbf{Q}} = \left(Y^{*\prime} Y^* + Y'Y \right) - \left(X^{*\prime} Y^* + X'Y \right)' \left(X^{*\prime} X^* + X'X \right)^{-1} \left(X^{*\prime} Y^* + X'Y \right)$$

$$\mathbf{\Sigma}_G = \mathbf{Q} \otimes \left(X^{*\prime} X^* + X'X \right)^{-1}.$$

As discussed in Section 3.2, to fix the scale of factors F_t in estimation, we do not estimate $\mathbf{Q}$ and instead set $\mathbf{Q} = \mathbf{I}_N$. Given $\mathbf{Q}$, we then only draw $\mathbf{G}$ using the posterior distribution (156). Finally, we enforce the stationarity of factors by discarding those draws of matrix $\mathbf{G}$ that have at least one eigenvalue greater than or equal to one in absolute value (explosive eigenvalues).

Appendix F. Tables and Figures

Figure F1. DFM: Principal Components Analysis

Data set: DFM3.TXT (standardized)

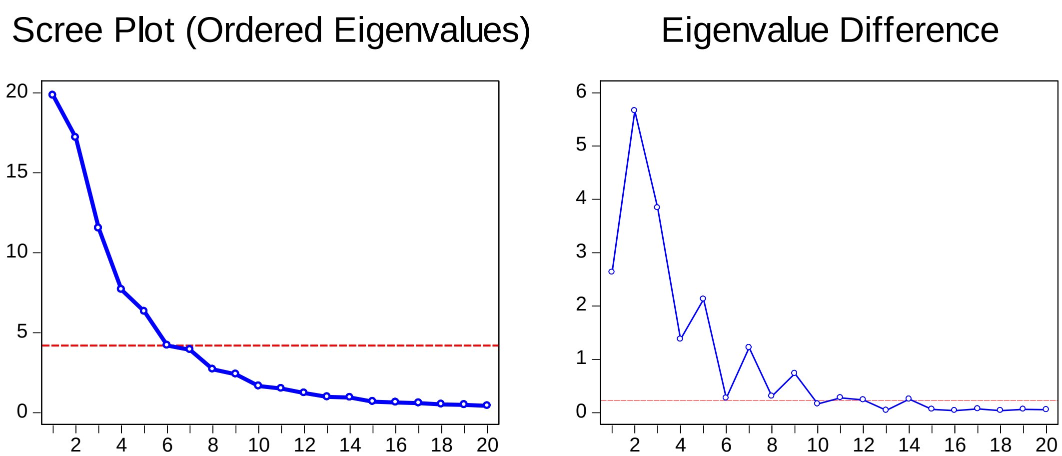


Table F1. DFM: Principal Components Analysis

Sample: 1984Q1 2005Q4
Included observations: 88
Computed using: Ordinary correlations
Extracting 20 of 89 possible components

Eigenvalues: (Sum = 89, Average = 1)

Number	Value	Difference	Proportion	Cumulative Value	Cumulative Proportion
1	19.82739	2.631345	0.2228	19.82739	0.2228
2	17.19605	5.659930	0.1932	37.02344	0.4160
3	11.53612	3.839474	0.1296	48.55955	0.5456
4	7.696642	1.375366	0.0865	56.25619	0.6321
5	6.321275	2.126480	0.0710	62.57747	0.7031
6	4.194795	0.270895	0.0471	66.77227	0.7503
7	3.923900	1.220256	0.0441	70.69617	0.7943
8	2.703644	0.305552	0.0304	73.39981	0.8247
9	2.398092	0.736125	0.0269	75.79790	0.8517
10	1.661967	0.160485	0.0187	77.45987	0.8703
11	1.501482	0.280114	0.0169	78.96135	0.8872
12	1.221368	0.238101	0.0137	80.18272	0.9009
13	0.983267	0.040017	0.0110	81.16598	0.9120
14	0.943250	0.252902	0.0106	82.10923	0.9226
15	0.690347	0.063015	0.0078	82.79958	0.9303
16	0.627333	0.038032	0.0070	83.42691	0.9374
17	0.589301	0.069497	0.0066	84.01621	0.9440
18	0.519803	0.038042	0.0058	84.53602	0.9498
19	0.481761	0.062722	0.0054	85.01778	0.9553
20	0.419039	0.054135	0.0047	85.43682	0.9600

Table F2. Pure DFM: Fraction of Unconditional Variance Captured by Factors

iid Measurement Errors; Dataset = DFM3.txt
<u>on average</u>, 100K draws, 20K burn-in

	All Factors	Error term
Core Variables	**0.948**	**0.052**
Real output	0.993	0.007
Inflation	0.896	0.104
Interest rates	0.990	0.010
Money velocities	0.914	0.086
Non-Core Variables	**0.941**	**0.059**
Output and components	0.982	0.018
Labor market	0.981	0.019
Investment, inventories, orders	0.986	0.014
Housing	0.970	0.030
Prices and wages	0.908	0.092
Financial variables	0.854	0.146
Other	0.973	0.027

Table F3. Data-Rich DSGE Model: Fraction of Unconditional Variance Captured by DSGE Model States

iid Measurement Errors; Dataset = DFM3.txt
<u>on average</u>, 20K draws, 4K burn-in

	GOV	CHI	MP	Z	All Shocks	Error term
Core Variables	**0.05**	**0.08**	**0.06**	**0.56**	**0.749**	**0.251**
Real output	0.14	0.21	0.03	0.48	0.852	0.148
Inflation	0.01	0.02	0.01	0.70	0.733	0.267
Interest rates	0.01	0.00	0.15	0.76	0.925	0.075
Money velocities	0.07	0.09	0.04	0.29	0.489	0.512
Non-Core Variables	**0.09**	**0.13**	**0.06**	**0.45**	**0.719**	**0.281**
Output and components	0.07	0.27	0.08	0.45	0.873	0.127
Labor market	0.19	0.14	0.06	0.46	0.848	0.152
Investment, inventories, orders	0.10	0.13	0.02	0.63	0.882	0.118
Housing	0.04	0.26	0.07	0.42	0.794	0.206
Prices and wages	0.03	0.05	0.04	0.45	0.568	0.432
Financial variables	0.06	0.03	0.05	0.32	0.451	0.549
Other	0.02	0.12	0.09	0.64	0.866	0.134

Table F4. Pure DFM: Unconditional Variance Captured by Factors

iid Measurement Errors; Dataset = DFM3.txt
on average, 100K draws, 20K burn-in

Algorithm: Jungbacker-Koopman
Identification: Scheme 2 - Block Diagonal

	F1	F2	F3	F4	F5	F6	All Factors	Measurement Error
Real GDP	0.119	0.142	0.301	0.160	0.115	0.148	0.984	0.016
IP_Total	0.137	0.105	0.343	0.135	0.113	0.164	0.996	0.004
IP_MFG	0.131	0.105	0.350	0.136	0.114	0.162	0.997	0.003
GDP Def inflation	0.147	0.173	0.166	0.169	0.110	0.142	0.907	0.094
PCE Def inflation	0.148	0.177	0.168	0.173	0.110	0.145	0.921	0.079
CPI ALL Inflation	0.130	0.167	0.159	0.166	0.102	0.138	0.862	0.138
FedFunds	0.135	0.169	0.185	0.169	0.186	0.148	0.993	0.008
3m T-Bill rate	0.136	0.166	0.185	0.168	0.189	0.148	0.991	0.009
AAA Bond yield	0.118	0.114	0.192	0.150	0.267	0.147	0.988	0.012
IVM_M1S_det	0.117	0.164	0.149	0.151	0.097	0.130	0.808	0.193
IVM_M2S	0.206	0.141	0.197	0.145	0.114	0.192	0.994	0.006
IVM_MBASE_bar	0.197	0.154	0.175	0.146	0.116	0.152	0.940	0.060
IP_CONS_DBLE	0.134	0.139	0.217	0.159	0.121	0.169	0.938	0.062
IP_CONS_NONDBLE	0.133	0.115	0.253	0.142	0.149	0.201	0.992	0.008
IP_BUS_EQPT	0.161	0.142	0.199	0.191	0.134	0.157	0.984	0.017
IP_DBLE_MATS	0.135	0.110	0.226	0.154	0.137	0.233	0.994	0.006
IP_NONDBLE_MATS	0.147	0.133	0.175	0.185	0.113	0.242	0.996	0.004
IP_FUELS	0.147	0.144	0.212	0.175	0.133	0.149	0.959	0.041
PMP	0.145	0.146	0.216	0.170	0.143	0.170	0.989	0.011
UTL11	0.141	0.181	0.184	0.183	0.143	0.165	0.997	0.003
RAHE_CONST	0.147	0.152	0.192	0.167	0.121	0.180	0.958	0.042
RAHE_MFG	0.166	0.137	0.184	0.149	0.120	0.228	0.983	0.017
EMP_MINING	0.130	0.118	0.211	0.210	0.123	0.169	0.960	0.040
EMP_CONST	0.153	0.141	0.193	0.166	0.112	0.234	0.998	0.002
EMP_DBLE_GDS	0.201	0.140	0.203	0.160	0.133	0.160	0.996	0.004
EMP_NONDBLES	0.158	0.120	0.183	0.183	0.116	0.236	0.995	0.005
EMP_SERVICES	0.164	0.155	0.211	0.141	0.126	0.201	0.997	0.003
EMP_TTU	0.140	0.159	0.184	0.173	0.139	0.176	0.971	0.029
EMP_WHOLESALE	0.144	0.167	0.168	0.142	0.114	0.145	0.879	0.121
EMP_RETAIL	0.162	0.157	0.177	0.163	0.143	0.164	0.967	0.033
EMP_FIRE	0.219	0.142	0.181	0.160	0.121	0.156	0.979	0.021
EMP_GOVT	0.150	0.135	0.266	0.137	0.152	0.155	0.996	0.004
URATE_ALL	0.124	0.175	0.255	0.157	0.141	0.141	0.993	0.007
U_DURATION	0.135	0.143	0.197	0.223	0.116	0.183	0.997	0.003
U_L5WKS	0.128	0.144	0.201	0.211	0.142	0.169	0.995	0.005
U_5_14WKS	0.145	0.143	0.195	0.167	0.154	0.163	0.966	0.034
U_M15WKS	0.132	0.153	0.198	0.218	0.121	0.177	0.998	0.002
U_15_26WKS	0.123	0.153	0.196	0.190	0.160	0.155	0.976	0.024
U_M27WKS	0.136	0.149	0.196	0.218	0.113	0.184	0.997	0.003
HOURS_AVG	0.151	0.147	0.207	0.163	0.145	0.178	0.991	0.009
HSTARTS_NE	0.132	0.135	0.193	0.173	0.154	0.175	0.962	0.038
HSTARTS_MW	0.118	0.121	0.240	0.163	0.155	0.145	0.942	0.058

HSTARTS_MW	0.118	0.121	0.240	0.163	0.155	0.145	0.942	0.058
HSTARTS_SOU	0.133	0.121	0.194	0.240	0.119	0.183	0.990	0.010
HSTARTS_WST	0.128	0.143	0.190	0.223	0.120	0.180	0.982	0.018
SFYGM6	0.138	0.143	0.201	0.167	0.152	0.168	0.970	0.030
SFYGT1	0.133	0.139	0.189	0.164	0.191	0.160	0.976	0.025
SFYGT10	0.150	0.197	0.182	0.160	0.132	0.153	0.974	0.026
SFYBAAC	0.151	0.188	0.178	0.170	0.129	0.171	0.988	0.012
BUS_LOANS	0.140	0.138	0.189	0.199	0.167	0.154	0.986	0.014
CONS_CREDIT	0.140	0.145	0.184	0.176	0.123	0.208	0.976	0.024
P_COM	0.139	0.133	0.189	0.151	0.112	0.150	0.874	0.126
P_OIL	0.117	0.121	0.181	0.139	0.104	0.130	0.792	0.208
P_NAPM_COM	0.138	0.128	0.197	0.147	0.125	0.148	0.882	0.118
DLOG_EXR_US	0.127	0.107	0.141	0.121	0.095	0.118	0.709	0.291
DLOG_EXR_CHF	0.107	0.100	0.135	0.112	0.090	0.111	0.655	0.345
DLOG_EXR_YEN	0.128	0.125	0.168	0.134	0.126	0.134	0.814	0.186
DLOG_EXR_GBP	0.098	0.095	0.129	0.111	0.088	0.105	0.626	0.374
DLOG_EXR_CAN	0.136	0.130	0.160	0.142	0.126	0.132	0.825	0.175
DLOG_SP500	0.133	0.136	0.171	0.138	0.111	0.137	0.827	0.173
DLOG_SP_IND	0.129	0.139	0.167	0.138	0.110	0.136	0.819	0.181
DLOG_DJIA	0.128	0.126	0.174	0.134	0.111	0.133	0.807	0.193
UMICH_CONS	0.142	0.121	0.246	0.142	0.130	0.167	0.949	0.051
NAPMI	0.144	0.149	0.219	0.173	0.140	0.170	0.994	0.006
NAPM_NEW_ORDRS	0.146	0.146	0.214	0.169	0.139	0.170	0.983	0.017
NAPM_VENDOR_DEL	0.142	0.147	0.222	0.170	0.137	0.168	0.985	0.015
NAPM_INVENTORIES	0.137	0.155	0.211	0.176	0.145	0.161	0.985	0.015
RCONS	0.172	0.144	0.187	0.175	0.127	0.177	0.982	0.018
RCONS_DUR	0.141	0.118	0.203	0.175	0.114	0.230	0.980	0.020
RCONS_SERV	0.139	0.134	0.186	0.202	0.115	0.214	0.990	0.010
RINV_GDP	0.153	0.125	0.225	0.155	0.145	0.192	0.995	0.005
RNONRESINV_STRUCT	0.165	0.138	0.187	0.153	0.118	0.224	0.984	0.016
RNONRESINV_BEQUIPT	0.141	0.168	0.185	0.198	0.128	0.156	0.976	0.024
RRESINV	0.176	0.155	0.182	0.186	0.128	0.150	0.977	0.023
REXPORTS	0.152	0.130	0.177	0.226	0.117	0.192	0.993	0.007
RIMPORTS	0.129	0.106	0.236	0.149	0.137	0.222	0.978	0.022
RGOV	0.203	0.133	0.207	0.141	0.138	0.171	0.994	0.006
LABOR_PROD	0.173	0.144	0.175	0.199	0.115	0.166	0.972	0.028
RCOMP_HOUR	0.183	0.161	0.190	0.153	0.123	0.177	0.987	0.014
ULC	0.134	0.151	0.187	0.225	0.122	0.170	0.989	0.011
PCED_DUR	0.135	0.133	0.178	0.174	0.181	0.150	0.950	0.050
PCED_NDUR	0.133	0.152	0.174	0.163	0.108	0.136	0.866	0.134
PCED_SERV	0.131	0.117	0.200	0.139	0.134	0.144	0.865	0.135
PINV_GDP	0.154	0.162	0.174	0.176	0.116	0.142	0.925	0.075
PINV_NRES_STRUCT	0.129	0.165	0.189	0.177	0.137	0.149	0.945	0.055
PINV_NRES_EQP	0.172	0.129	0.182	0.151	0.113	0.149	0.897	0.103
PINV_RES	0.121	0.135	0.191	0.173	0.110	0.140	0.870	0.130
PEXPORTS	0.164	0.147	0.204	0.170	0.123	0.155	0.963	0.037
PIMPORTS	0.149	0.142	0.192	0.162	0.117	0.144	0.906	0.094
PGOV	0.122	0.125	0.156	0.140	0.111	0.124	0.778	0.222

Notes: Please see Chapter 1, Appendix C. Data: Description and Transformations, p. 72 for the corresponding mnemonics of data indicators reported here.

Table F5. Data-Rich DSGE Model: Fraction of Unconditional Variance Captured by DSGE Model States

iid Measurement Errors; Dataset = DFM3.txt
on average, 20K draws, 4K burn-in

Algorithm: Jungbacker-Koopman

	GOV	CHI	MP	Z	All Shocks	Measurement Error
Real GDP	0.081	0.000	0.040	0.648	0.770	0.230
IP_Total	0.167	0.308	0.021	0.395	0.891	0.110
IP_MFG	0.166	0.317	0.020	0.392	0.894	0.106
GDP Def inflation	0.011	0.000	0.011	0.789	0.811	0.189
PCE Def inflation	0.004	0.035	0.003	0.703	0.745	0.255
CPI ALL Inflation	0.005	0.031	0.006	0.600	0.642	0.358
FedFunds	0.004	0.000	0.135	0.817	0.956	0.044
3m T-Bill rate	0.007	0.003	0.160	0.788	0.958	0.042
AAA Bond yield	0.013	0.008	0.168	0.672	0.861	0.139
IVM_M1S_det	0.055	0.174	0.016	0.404	0.648	0.352
IVM_M2S	0.042	0.063	0.003	0.071	0.178	0.822
IVM_MBASE_bar	0.099	0.031	0.104	0.406	0.639	0.361
IP_CONS_DBLE	0.051	0.090	0.018	0.650	0.810	0.190
IP_CONS_NONDBLE	0.151	0.551	0.025	0.109	0.836	0.164
IP_BUS_EQPT	0.259	0.103	0.106	0.407	0.874	0.126
IP_DBLE_MATS	0.069	0.677	0.024	0.131	0.901	0.099
IP_NONDBLE_MATS	0.060	0.229	0.028	0.645	0.962	0.038
IP_FUELS	0.081	0.136	0.044	0.457	0.718	0.282
PMP	0.085	0.046	0.014	0.702	0.848	0.153
UTL11	0.010	0.002	0.066	0.913	0.991	0.010
RAHE_CONST	0.131	0.010	0.035	0.566	0.742	0.258
RAHE_MFG	0.116	0.024	0.124	0.651	0.915	0.085
EMP_MINING	0.055	0.030	0.007	0.596	0.688	0.312
EMP_CONST	0.094	0.190	0.134	0.546	0.964	0.037
EMP_DBLE_GDS	0.137	0.272	0.177	0.381	0.967	0.034
EMP_NONDBLES	0.035	0.117	0.186	0.609	0.947	0.053
EMP_SERVICES	0.111	0.400	0.069	0.379	0.958	0.042
EMP_TTU	0.012	0.320	0.011	0.399	0.743	0.258
EMP_WHOLESALE	0.011	0.020	0.056	0.248	0.335	0.665
EMP_RETAIL	0.011	0.237	0.059	0.455	0.761	0.239
EMP_FIRE	0.022	0.150	0.111	0.501	0.784	0.216
EMP_GOVT	0.162	0.237	0.016	0.467	0.882	0.118
URATE_ALL	0.175	0.056	0.014	0.619	0.864	0.136
U_DURATION	0.656	0.149	0.015	0.147	0.967	0.033
U_L5WKS	0.384	0.051	0.031	0.463	0.928	0.072
U_5_14WKS	0.143	0.033	0.011	0.523	0.710	0.290
U_M15WKS	0.575	0.099	0.018	0.284	0.977	0.023
U_15_26WKS	0.096	0.006	0.043	0.715	0.859	0.141
U_M27WKS	0.664	0.160	0.014	0.135	0.973	0.027
HOURS_AVG	0.019	0.032	0.095	0.816	0.961	0.039

HSTARTS_NE	0.009	0.115	0.016	0.679	0.819	0.181
HSTARTS_MW	0.017	0.193	0.115	0.273	0.598	0.402
HSTARTS_SOU	0.058	0.601	0.059	0.152	0.870	0.130
HSTARTS_WST	0.019	0.328	0.075	0.404	0.826	0.174
SFYGM6	0.090	0.041	0.029	0.642	0.802	0.198
SFYGT1	0.067	0.024	0.054	0.698	0.843	0.157
SFYGT10	0.157	0.006	0.025	0.460	0.648	0.352
SFYBAAC	0.034	0.004	0.082	0.811	0.931	0.069
BUS_LOANS	0.279	0.032	0.230	0.251	0.791	0.209
CONS_CREDIT	0.064	0.212	0.065	0.275	0.616	0.384
P_COM	0.038	0.012	0.011	0.335	0.396	0.604
P_OIL	0.008	0.011	0.007	0.263	0.288	0.712
P_NAPM_COM	0.017	0.017	0.010	0.223	0.267	0.733
DLOG_EXR_US	0.008	0.016	0.039	0.118	0.180	0.820
DLOG_EXR_CHF	0.007	0.013	0.030	0.110	0.160	0.840
DLOG_EXR_YEN	0.011	0.010	0.010	0.116	0.147	0.853
DLOG_EXR_GBP	0.007	0.012	0.016	0.117	0.152	0.848
DLOG_EXR_CAN	0.010	0.029	0.058	0.184	0.280	0.720
DLOG_SP500	0.016	0.010	0.026	0.222	0.274	0.726
DLOG_SP_IND	0.016	0.009	0.024	0.259	0.308	0.692
DLOG_DJIA	0.010	0.010	0.017	0.147	0.183	0.817
UMICH_CONS	0.006	0.311	0.046	0.405	0.767	0.233
NAPMI	0.075	0.050	0.016	0.760	0.900	0.100
NAPM_NEW_ORDRS	0.093	0.047	0.010	0.652	0.802	0.198
NAPM_VENDOR_DEL	0.068	0.053	0.015	0.711	0.846	0.154
NAPM_INVENTORIES	0.047	0.046	0.023	0.804	0.919	0.081
RCONS	0.005	0.032	0.196	0.667	0.901	0.099
RCONS_DUR	0.044	0.319	0.144	0.353	0.859	0.141
RCONS_SERV	0.009	0.237	0.099	0.580	0.925	0.075
RINV_GDP	0.005	0.479	0.069	0.415	0.967	0.033
RNONRESINV_STRUCT	0.339	0.184	0.013	0.327	0.863	0.137
RNONRESINV_BEQUIPT	0.095	0.027	0.008	0.750	0.880	0.120
RRESINV	0.092	0.078	0.092	0.596	0.858	0.142
REXPORTS	0.018	0.093	0.196	0.635	0.942	0.058
RIMPORTS	0.055	0.615	0.025	0.119	0.813	0.186
RGOV	0.006	0.339	0.175	0.437	0.957	0.043
LABOR_PROD	0.033	0.044	0.161	0.602	0.839	0.161
RCOMP_HOUR	0.020	0.026	0.176	0.563	0.784	0.216
ULC	0.090	0.215	0.019	0.526	0.850	0.150
PCED_DUR	0.021	0.044	0.023	0.699	0.788	0.212
PCED_NDUR	0.009	0.023	0.006	0.438	0.474	0.526
PCED_SERV	0.007	0.088	0.005	0.457	0.557	0.443
PINV_GDP	0.015	0.036	0.045	0.544	0.639	0.361
PINV_NRES_STRUCT	0.019	0.048	0.023	0.397	0.486	0.514
PINV_NRES_EQP	0.008	0.118	0.023	0.447	0.596	0.404
PINV_RES	0.028	0.080	0.036	0.270	0.414	0.586
PEXPORTS	0.013	0.022	0.015	0.637	0.687	0.313
PIMPORTS	0.012	0.015	0.012	0.499	0.537	0.463
PGOV	0.009	0.019	0.029	0.177	0.233	0.767

Notes: Structural shocks are GOV – government spending, CHI – money demand, MP – monetary policy and Z – neutral technology. Please see Chapter 1, Appendix C. Data: Description and Transformations, p. 72 for the corresponding mnemonics of data indicators reported here.

Table F6. Regressing Data-Rich DSGE Model States on DFM Factors

Model Concept		R^2
Inflation	PI_t	0.984
Interest Rate	R_t	0.991
Real Consumption	X_t	0.998
Govt Spending shock	GOV_t	0.999
Money Demand shock	CHI_t	0.999
Technology shock	Z_t	0.990

Notes: Each line reports the R^2 from predictive linear regression:

$$S_{i,t}^{(pm)} = \alpha_{0,i} + \boldsymbol{\alpha}'_{1,i} F_t^{(pm)} + v_{i,t},$$

where $S_{i,t}^{(pm)}$ is the posterior mean of the i^{th} data-rich DSGE model state variable and $F_t^{(pm)}$ is the posterior mean of the empirical factors extracted by DFM.

Table F7. Regressing DFM Factors on Data-Rich DSGE Model States

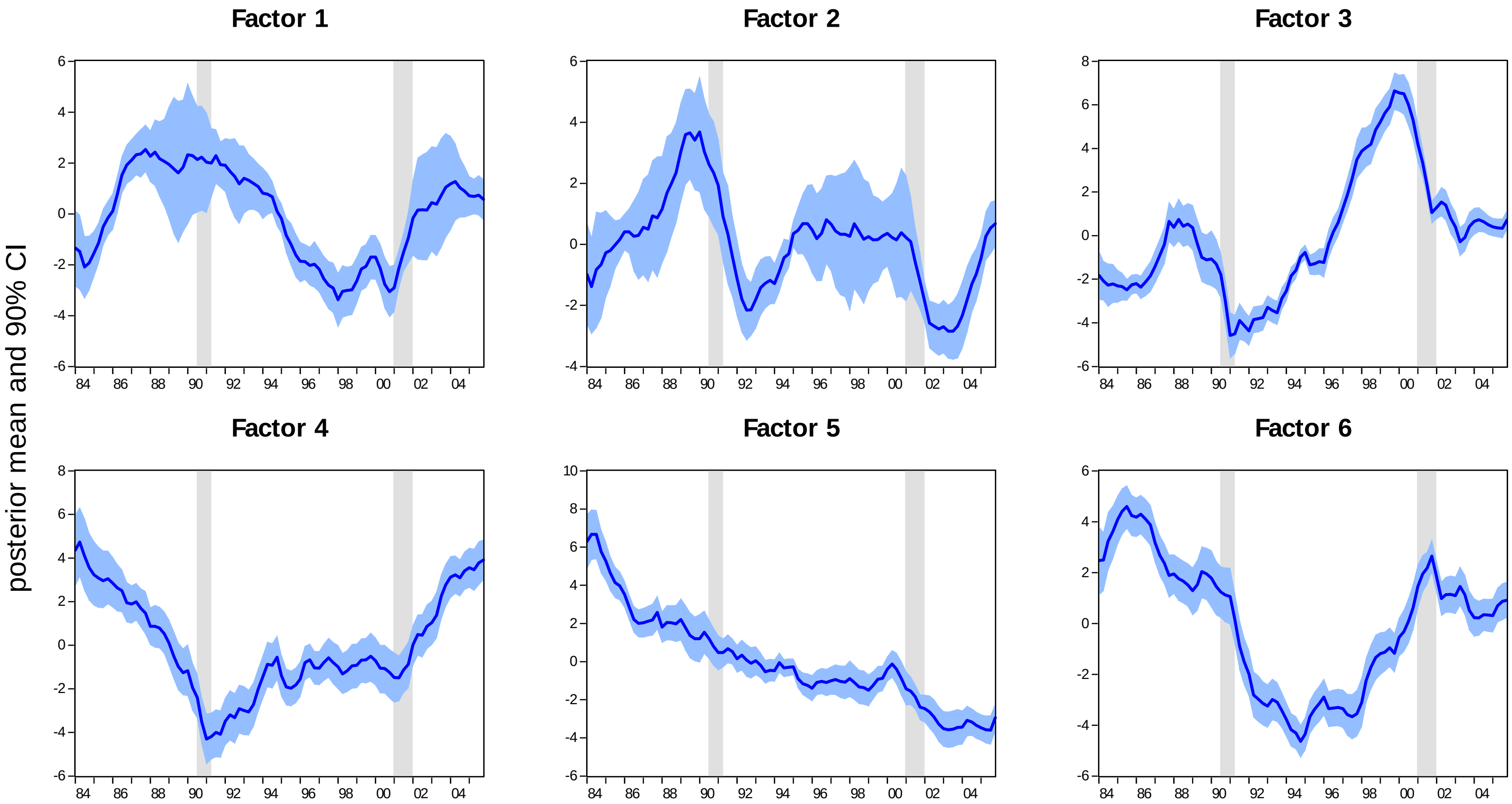
Factors	R^2
Factor 1	0.979
Factor 2	0.924
Factor 3	0.949
Factor 4	0.981
Factor 5	0.989
Factor 6	0.991

Notes: Each line reports the R^2 from predictive linear regression (see (147) in the main text):

$$F_{i,t}^{(pm)} = \beta_{0,i} + \boldsymbol{\beta}'_{1,i} S_t^{(pm)} + u_{i,t},$$

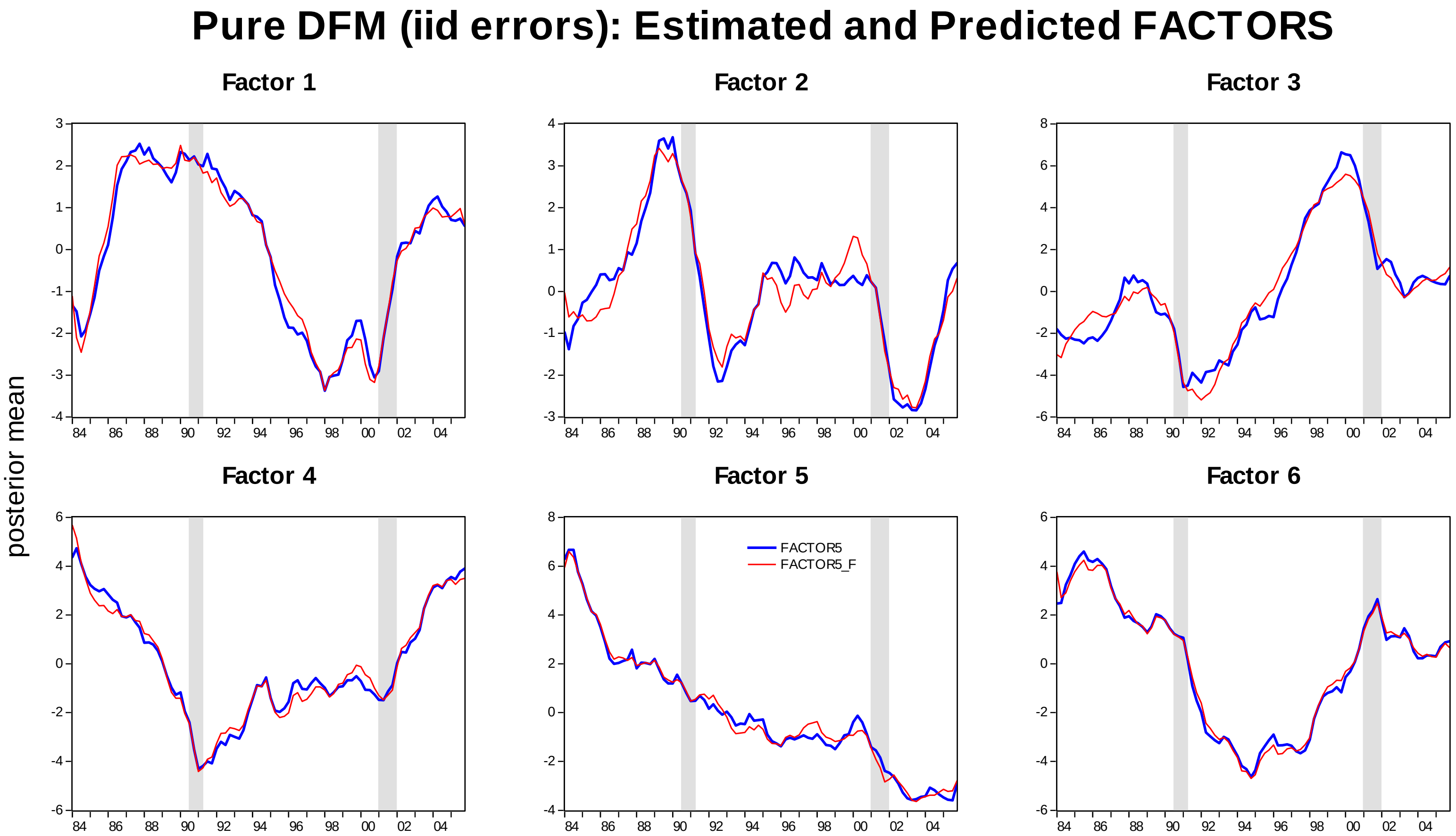
where $F_{i,t}^{(pm)}$ is the posterior mean of the i^{th} empirical factor extracted by DFM and $S_t^{(pm)}$ is the posterior mean of the data-rich DSGE model state variables.

Figure F2. Pure DFM (iid errors): Estimated Factors



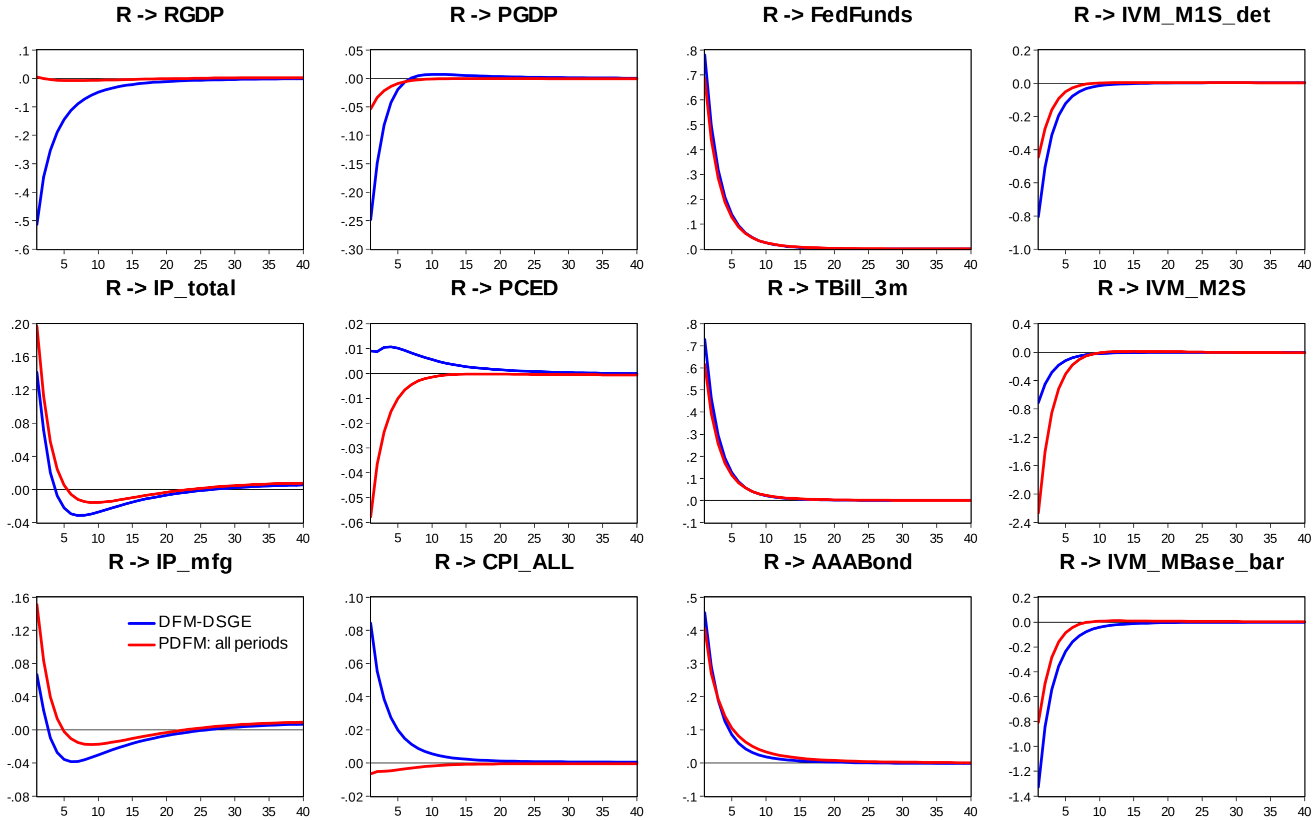
Notes: The figure plots the posterior means and 90% credible intervals of the latent empirical factors extracted by the empirical DFM (137)-(139). Normalization: block diagonal. Algorithm: Jungbacker-Koopman (2008).

Figure F3. Do Empirical Factors and DSGE Model State Variables Span the Same Space?



Notes: The figure plots the actual empirical factors extracted by the DFM (137)-(139) (blue line) and the empirical factors predicted by the data-rich DSGE model state variables using (148) in the main text (red line).

Figure F4. Impact of Monetary Policy Innovation on Core Macro Series

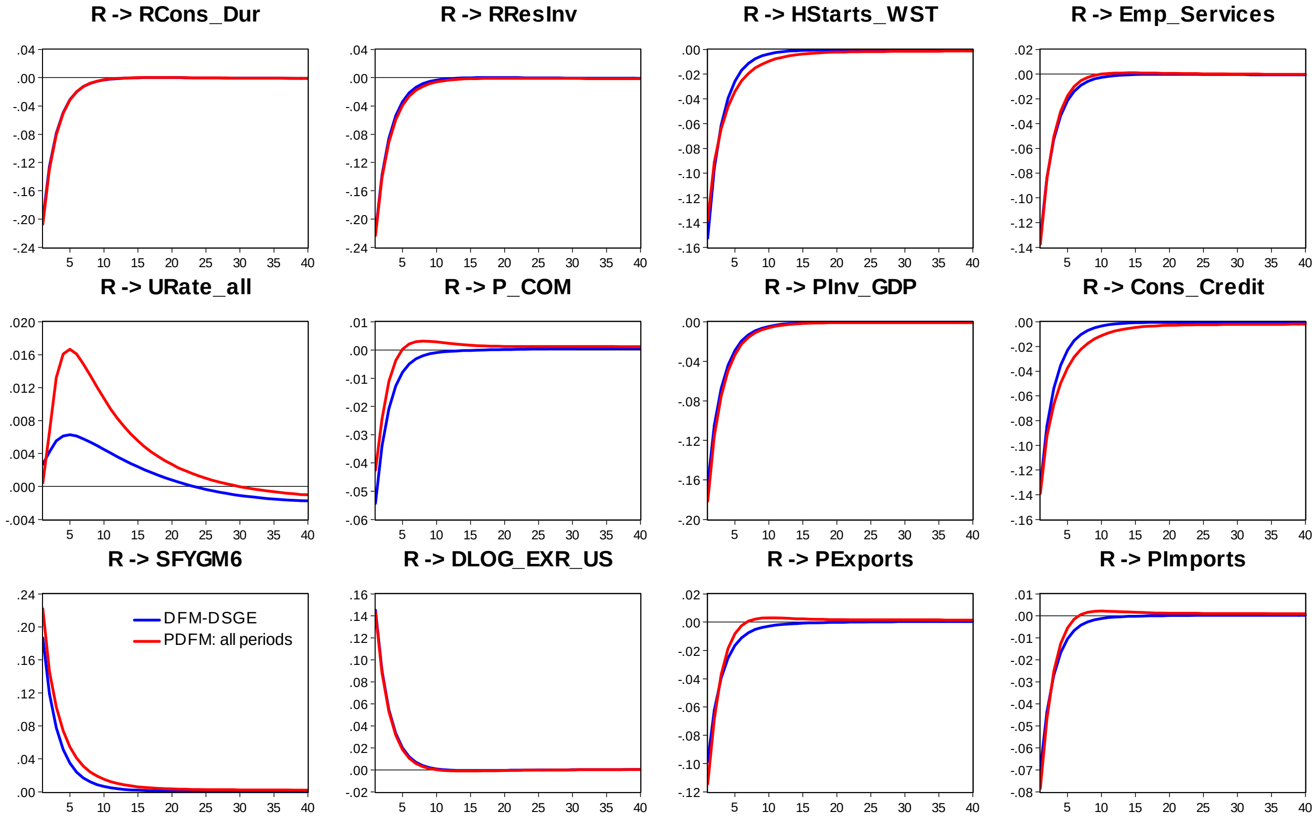


Notes: The figure plots the impulse responses of data indicators to a 1-standard-deviation **monetary policy** innovation $(\varepsilon_{R,t})$ computed in the data-rich DSGE model (blue line, "DFM-DSGE") and in empirical pure DFM (red line, "PDFM: all periods") according to (149) and (150), respectively.

The impact of structural shock is mapped from data-rich DSGE model into empirical DFM every period.

Data indicators are real GDP (RGDP), industrial production: total (IP_total), industrial production: manufacturing (IP_mfg), GDP deflator inflation (PGDP), PCE deflator inflation (PCED), CPI inflation (CPI_ALL), Federal Funds rate (FedFunds), 3-month T-Bill rate (TBill_3m), yield on AAA rated corporate bonds (AAABond), real money balances based on M1S aggregate (IVM_M1S_det), on M2S aggregate (IVM_M2S), and on adjusted monetary base (IVM_MBase_bar). See the corresponding mnemonics in Chapter 1, Appendix C. Data: Description and Transformations, p. 72.

Figure F5. Impact of Monetary Policy Innovation on Non-Core Macro Series

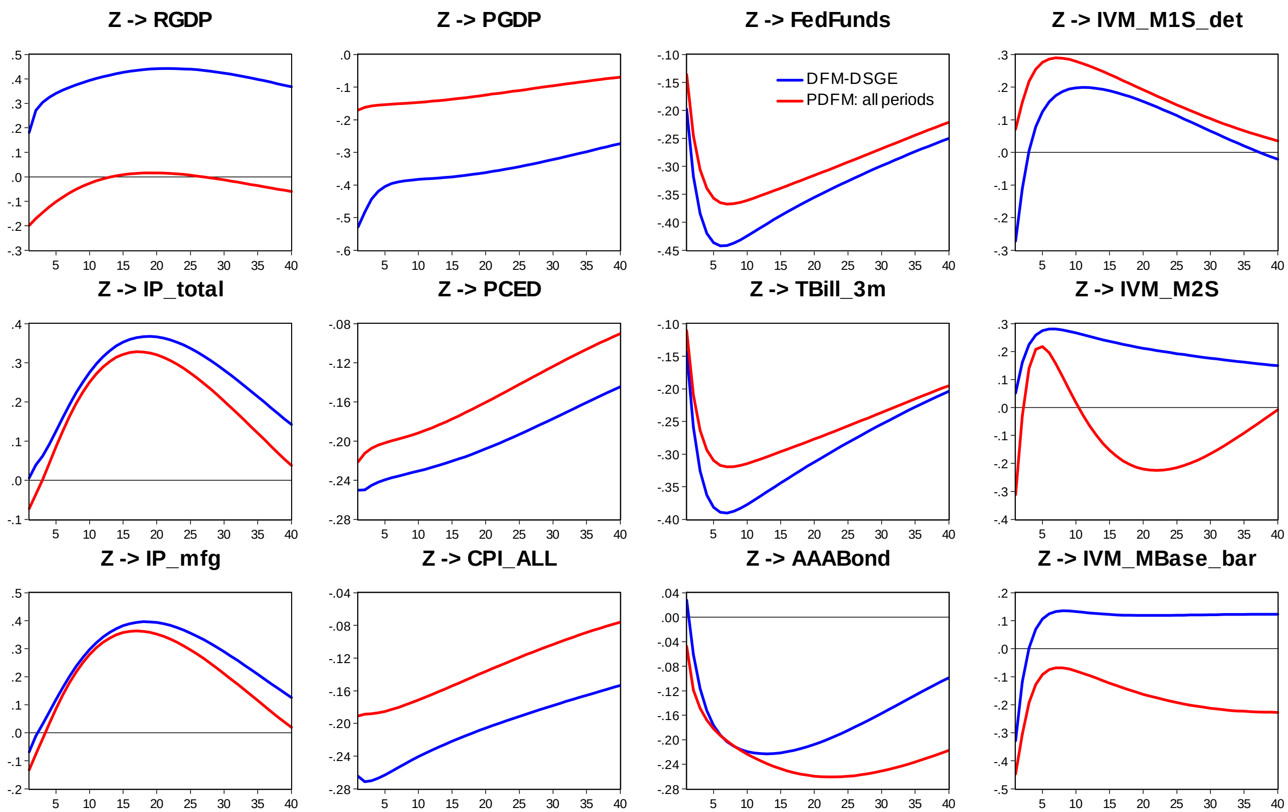


Notes: The figure plots the impulse responses of data indicators to a 1-standard-deviation **monetary policy** innovation $(\varepsilon_{R,t})$ computed in the data-rich DSGE model (blue line, "DFM-DSGE") and in empirical pure DFM (red line, "PDFM: all periods") according to (149) and (150), respectively.

The impact of structural shock is mapped from data-rich DSGE model into empirical DFM every period.

Data indicators are real consumption of durables (RCons_Dur), real residential investment (RResInv), housing starts: West (HStarts_WST), employment in services sector (Emp_Services), unemployment rate (URate_all), commodity price inflation (P_COM), investment deflator inflation (PInv_GDP), consumer credit outstanding (Cons_Credit), 6-month over 3-month T-Bill rate spread (SFYGM6), US effective exchange rate depreciation (DLOG_EXR_US), exports price index (PExports), imports price index (PImports). See the corresponding mnemonics in Chapter 1, Appendix C. Data: Description and Transformations, p. 72.

Figure F6. Impact of Technology Innovation on Core Macro Series

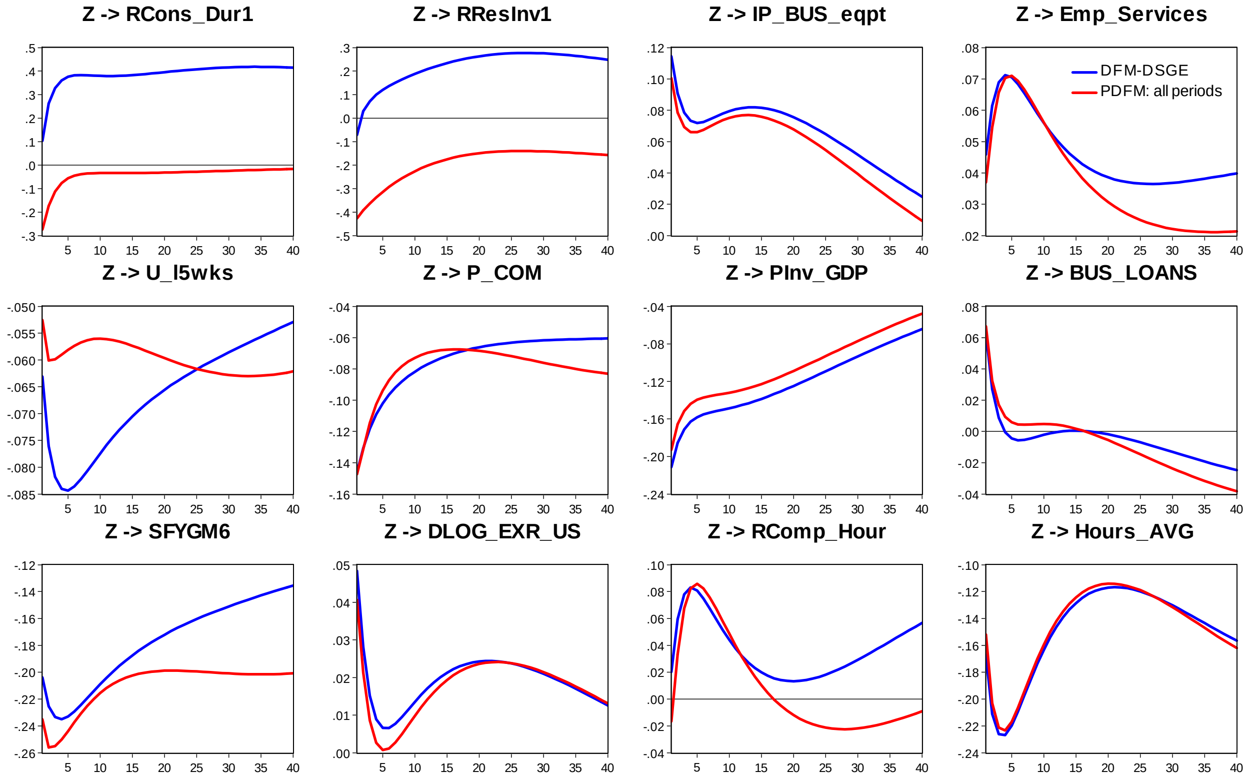


Notes: The figure plots the impulse responses of data indicators to a 1-standard-deviation **technology** innovation $(\varepsilon_{Z,t})$ computed in the data-rich DSGE model (blue line, "DFM-DSGE") and in empirical pure DFM (red line, "PDFM: all periods") according to (149) and (150), respectively.

The impact of structural shock is mapped from data-rich DSGE model into empirical DFM every period.

Data indicators are real GDP (RGDP), industrial production: total (IP_total), industrial production: manufacturing (IP_mfg), GDP deflator inflation (PGDP), PCE deflator inflation (PCED), CPI inflation (CPI_ALL), Federal Funds rate (FedFunds), 3-month T-Bill rate (TBill_3m), yield on AAA rated corporate bonds (AAABond), real money balances based on M1S aggregate (IVM_M1S_det), on M2S aggregate (IVM_M2S), and on adjusted monetary base (IVM_MBase_bar). See the corresponding mnemonics in Chapter 1, Appendix C. Data: Description and Transformations, p. 72.

Figure F7. Impact of Technology Innovation on Non-Core Macro Series



Notes: The figure plots the impulse responses of data indicators to a 1-standard-deviation **technology** innovation $(\varepsilon_{Z,t})$ computed in the data-rich DSGE model (blue line, "DFM-DSGE") and in empirical pure DFM (red line, "PDFM: all periods") according to (149) and (150), respectively.

The impact of structural shock is mapped from data-rich DSGE model into empirical DFM every period.

Data indicators are real consumption of durables (RCons_Dur1), real residential investment (RResInv1), industrial production: business equipment (IP_BUS_eqpt), employment in services sector (Emp_Services), persons unemployed less than 5 weeks (U_l5wks), commodity price inflation (P_COM), investment deflator inflation (PInv_GDP), commercial and industrial loans (BUS_LOANS), 6-month over 3-month T-Bill rate spread (SFYGM6), US effective exchange rate depreciation (DLOG_EXR_US), real compensation per hour (RComp_Hour), average weekly hours worked (Hours_AVG). See the corresponding mnemonics in Chapter 1, Appendix C. Data: Description and Transformations, p. 72.

CHAPTER 3. DSGE MODEL BASED FORECASTING OF NON-MODELED VARIABLES

Joint work with Frank Schorfheide and Keith Sill

1 Introduction

Dynamic stochastic general equilibrium (DSGE) models estimated using Bayesian methods are increasingly being used by central banks around the world as tools for projections and policy analysis. Examples of such models include the small open economy model developed by the Sveriges Riksbank (Adolfson, Laseen, Linde, & Villani, 2007, 2008; Adolfson, Andersson, Linde, Villani, & Vredin, 2007) the new area-wide model developed at the European Central Bank (Coenen, McAdam, & Straub, 2008) and the Federal Reserve Board's new estimated, dynamic, optimization-based model (Edge, Kiley, & Laforte, 2009). These models extend the specifications studied by Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2003) to open economy and multisector settings. A common feature of these models is that the decision rules of economic agents are derived from assumptions about preferences and technologies by solving intertemporal optimization problems.

Compared to previous generations of macroeconometric models, the DSGE paradigm delivers empirical models with a strong degree of theoretical coherence. The costs associated with this theoretical coherence are two-fold. First, tight cross-equation restrictions could potentially introduce misspecification problems that manifest themselves through an inferior fit compared to less-restrictive time series models (Del Negro, Schorfheide, Smets, & Wouters, 2007 henceforth DSSW). Second, it is more cumbersome to incorporate variables other than a core set of macroeconomic aggregates such as real gross domestic product (GDP), consumption, investment, wages, hours, inflation, and interest rates than in a traditional system-of-equations approach. Nonetheless, in practical work at central banks it might be important to also generate forecasts for economic variables that do not explicitly appear in medium-scale DSGE models. This chapter focuses on this second problem.

In principle there are two options for generating forecasts for additional variables. First, one could enlarge the structural model to incorporate these variables explicitly. The advantage of a larger model is its ability to deliver a coherent narrative that can accompany the forecasts. The disadvantages are that identification problems are often exacerbated in large-scale models, the numerical analysis (e.g., estimation procedures that utilize numerical optimization or posterior simulation routines) becomes more tenuous, and the maintenance of the model requires more staff resources. The second option is to develop a hybrid empirical model that augments a medium-scale core DSGE model with auxiliary equations that create a link between explicitly modelled variables and non-modelled variables. For the sake of brevity we will refer to the latter as non-core variables. One could interpret these auxiliary equations as log-linear approximations of agents' decision rules in a larger DSGE model. This chapter explores the second approach.

Recently, Boivin and Giannoni (2006, henceforth BG) integrated a medium-scale DSGE model into a dynamic factor model for a large cross section of macroeconomic indicators, thereby linking non-core variables to a DSGE model. We will refer to this hybrid model as DSGE-DFM. The authors jointly estimated the DSGE model parameters and the factor loadings for the non-core variables. Compared to the estimation of a "non-structural" dynamic factor model, the BG approach leads to factor estimates that have clear economic interpretation. The joint estimation is conceptually very appealing, in part because it exploits information that is contained in the non-core variables when making inferences about the state of the economy.[19] The downside of the joint estimation is its computational complexity, which currently makes it impractical for real time forecasting applications.

This chapter proposes a simpler two-step estimation approach for an empirical model that consists of a medium-scale DSGE model for a set of core macroeconomic variables and a collection of measurement equations or auxiliary regressions that link the state variables of the DSGE model with the non-core variables of interest to the analyst. In the first step we estimate the DSGE model using the core variables as measurements. Based on the DSGE model parameter estimates, we use the Kalman filter to obtain estimates of the latent state variables given the most recent information set. We then use

[19] Formally, when the term "state of the economy" is used, we mean information about the latent state variables that appear in the DSGE model.

the filtered state variables as regressors to estimate simple linear measurement equations with serially correlated idiosyncratic errors.

There are three advantages of our procedure. First, since the DSGE model estimation is fairly tedious and delicate, in real time applications the DSGE model could be re-estimated infrequently (for instance, once a year). Second, the estimation of the measurement equations is quick and can easily be repeated in real time as new information arrives or interest arises in additional non-core variables. The estimated auxiliary regressions can then be used to generate forecasts of the non-core variables. Third, our empirical model links the non-core variables to the fundamental shocks that are believed to drive business cycle fluctuations. In particular, the model allows monetary policy shocks and other structural shocks to propagate through to non-core variables. This allows us to study the effect of unanticipated changes in monetary policy on a broad set of economic variables.[20]

The remainder of the chapter is organized as follows. The DSGE model used for the empirical analysis is described in Section 2; we are using a variant of the Christiano et al. (2005) and Smets and Wouters (2003) model, which is described in detail by DSSW. Our econometric framework is presented in Section 3. Section 4 summarizes the results of our empirical analysis. We estimate the DSGE model recursively based on US quarterly data, starting with a sample from 1984:I to 2000:IV, and generate estimates of the latent states as well as pseudo-out-of-sample forecasts for a set of core variables (the

[20] The goal of our analysis is distinctly different from that of recent work by Giannone, Monti, and Reichlin (2008) and Monti (2008), who develop state space models that allow the analyst to use high frequency data or professional forecasts to update or improve the DSGE-model based forecasts of the core variables.

growth rates of output, consumption, investment, nominal wages, the GDP deflator, as well as the levels of interest rates and hours worked). We then estimate measurement equations for four additional variables: personal consumption expenditures (PCE) inflation, core PCE inflation, the unemployment rate, and housing starts. We provide pseudo-out-of-sample forecast error statistics for both the core and non-core variables using our empirical model and compare them to simple AR(1) forecasts. Finally, we study the propagation of monetary policy shocks to auxiliary variables, as well as features of the joint predictive distribution. Section 5 concludes and discusses future research. Details of the Bayesian computations are relegated to the Appendix.

2 The DSGE Model

We use a medium-scale New Keynesian model with price and wage rigidities, capital accumulation, investment adjustment costs, variable capital utilization, and habit formation. The model is based on the work of Smets and Wouters (2003) and Christiano et al. (2005), and this specific version is taken from DSSW. For the sake of brevity we present only the log-linearized equilibrium conditions, and refer the reader to the above-referenced papers for the derivation of these conditions from assumptions about preferences and technologies.

The economy is populated by a continuum of firms that combine capital and labor to produce differentiated intermediate goods. These firms all have access to the same Cobb–Douglas production function with capital elasticity α and total factor productivity A_t. The total factor productivity is assumed to be non-stationary. We denote its growth rate by $a_t = \ln(A_t / A_{t-1})$, which is assumed to have a mean of γ. Output, consumption, investment, capital, and the real wage can be detrended by A_t. In terms of the detrended variables, the model has a well-defined steady state. All variables that appear subsequently are expressed as log-deviations from this steady state.

The intermediate goods producers hire labor and rent capital in competitive markets, and face identical real wages, w_t, and rental rates for capital, r_t^k. Cost minimization implies that all firms produce with the same capital–labor ratio

$$k_t - L_t = w_t - r_t^k \tag{157}$$

and have marginal costs

$$mc_t = (1-\alpha)w_t + \alpha r_t^k \tag{158}$$

The intermediate goods producers sell their output to perfectly competitive final good producers, which aggregate the inputs according to a CES function. Profit maximization of the final good producers implies that

$$\hat{y}_t(j) - \hat{y}_t = -\left(1 + \frac{1}{\lambda_f e^{\tilde{\lambda}_{f,t}}}\right)(p_t(j) - p_t). \tag{159}$$

Here $\hat{y}_t(j) - \hat{y}_t$ and $p_t(j) - p_t$ are the quantity and price for the good *j* relative to the quantity and price of the final good. The price p_t of the final good is determined from a zero-profit condition for the final good producers.

We assume that the price elasticity of the intermediate goods is time-varying. Since this price elasticity affects the mark-up that intermediate goods producers can charge over marginal costs, we refer to $\tilde{\lambda}_{f,t}$ as the mark-up shock. Following Calvo (1983), we assume that a certain fraction of the intermediate goods producers ζ_p is unable to re-optimize their prices in each period. These firms adjust their prices mechanically according to steady state inflation π_*. All other firms choose their prices to maximize the expected discounted sum of future profits, which leads to the following equilibrium relationship, known as the New Keynesian Phillips curve:

$$\pi_t = \beta E_t[\pi_{t+1}] + \frac{(1-\zeta_p\beta)(1-\zeta_p)}{\zeta_p} mc_t + \frac{1}{\zeta_p}\lambda_{f,t} \tag{160}$$

where π_t is inflation and β is the discount rate.[21] Our assumption on the behavior of firms which are unable to re-optimize their prices implies the absence of price dispersion in the steady state. As a consequence, we obtain a log-linearized aggregate production function of the form

$$\hat{y}_t = (1-\alpha)L_t + \alpha k_t \tag{161}$$

Eqs. (158), (157) and (161) imply that the labor share lsh_t equals the marginal costs in terms of log-deviations: $lsh_t = mc_t$.

There is a continuum of households with identical preferences, which are separable in consumption, leisure, and real money balances. Households' preferences display a degree of (internal) habit formation in consumption, captured by the parameter *h*. The period *t* utility is a function of $\ln(C_t - hC_{t-1})$. Households supply monopolistically

[21] We used the following re-parameterization: $\lambda_{f,t} = [(1-\zeta_p\beta)(1-\zeta_p)\lambda_f/(1+\lambda_f)]\tilde{\lambda}_{f,t}$, where λ_f is the steady state of $\tilde{\lambda}_{f,t}$.

differentiated labor services. These services are aggregated according to a CES function that leads to a demand elasticity $1+1/\lambda_w$. The composite labor services are then supplied to the intermediate goods producers at a real wage w_t. To introduce nominal wage rigidity, we assume that in each period, a certain fraction ζ_w of households is unable to re-optimize their wages. These households adjust their nominal wage by the steady state wage growth $e^{(\pi_*+\gamma)}$. All other households re-optimize their wages. The first-order conditions imply that

$$\begin{aligned}\tilde{w}_t &= \zeta_w \beta E_t[\tilde{w}_{t+1} + \Delta w_{t+1} + \pi_{t+1} + a_{t+1}] + \\ &+ \frac{1-\zeta_w\beta}{1+\nu_l(1+\lambda_w)/\lambda_w} \times \left(\nu_l L_t - w_t - \xi_t + \tilde{b}_t + \frac{1}{1-\zeta_w\beta}\phi_t \right)\end{aligned} \tag{162}$$

where $\tilde{w}_t$ is the optimal real wage relative to the real wage for aggregate labor services, w_t, and ν_l is the inverse Frisch labor supply elasticity in a model without wage rigidity ($\zeta_w = 0$) and differentiated labor. Moreover, $\tilde{b}_t$ is a shock to the household's discount factor[22] and ϕ_t is a preference shock that affects the household's intratemporal substitution between consumption and leisure. The real wage paid by intermediate goods producers evolves according to

$$w_t = w_{t-1} - \pi_t - a_t + \frac{1-\zeta_w}{\zeta_w}\tilde{w}_t \tag{163}$$

[22] For the estimation we re-parameterize the shock as follows: $b_t = e^{\gamma}(e^{\gamma} - h)/(e^{2\gamma} + \beta h^2)\tilde{b}_t$

Households are able to insure against the idiosyncratic wage adjustment shocks with state contingent claims. As a consequence, they all share the same marginal utility of consumption ξ_t, which is given by the expression:

$$(e^{\gamma}-h\beta)(e^{\gamma}-h)\xi_t=-(e^{2\gamma}+\beta h^2)c_t+\beta he^{\gamma}E_t[c_{t+1}+a_{t+1}]+ \\ +he^{\gamma}(c_{t-1}-a_t)+e^{\gamma}(e^{\gamma}-h)\tilde{b}_t-\beta h(e^{\gamma}-h)E_t[\tilde{b}_{t+1}] \tag{164}$$

where c_t is consumption. In addition to state-contingent claims, households accumulate three types of assets: one-period nominal bonds that yield the return R_t, capital $\bar{k}_t$, and real money balances. Since the preferences for real money balances are assumed to be additively separable and monetary policy is conducted through a nominal interest rate feedback rule, money is block exogenous and we will not use the households' money demand equation in our empirical analysis.

The first order condition with respect to bond holdings delivers the standard Euler equation:

$$\xi_t=E_t[\xi_{t+1}]+R_t-E_t[\pi_{t+1}]-E_t[a_{t+1}]. \tag{165}$$

Capital accumulates according to the following law of motion:

$$\bar{k}_t=(2-e^{\gamma}-\delta)[\bar{k}_{t-1}-a_t]+(e^{\gamma}+\delta-1)[i_t+(1+\beta)S''e^{2\gamma}\mu_t], \tag{166}$$

where i_t is investment, δ is the depreciation rate of capital, and μ_t can be interpreted as an investment-specific technology shock. Investment in our model is subject to adjustment costs, and S'' denotes the second derivative of the investment adjustment cost function at the steady state. The optimal investment satisfies the following first-order condition:

$$i_t = \frac{1}{1+\beta}[i_{t-1} - a_t] + \frac{\beta}{1+\beta}E_t[i_{t+1} + a_{t+1}] + \frac{1}{(1+\beta)S''e^{2\gamma}}(\xi_t^k - \xi_t) + \mu_t \tag{167}$$

where ξ_t^k is the value of the installed capital, which evolves according to:

$$\xi_t^k - \xi_t = \beta e^{-\gamma}(1-\delta)E_t\left[\xi_{t+1}^k - \xi_{t+1}\right] + E_t\left[(1-(1-\delta)\beta e^{-\gamma})r_{t+1}^k - (R_t - \pi_{t+1})\right]. \tag{168}$$

The capital utilization u_t in our model is variable, and r_t^k in all previous equations represents the rental rate of effective capital $k_t = u_t + \bar{k}_{t-1}$. The optimal degree of utilization is determined by

$$u_t = \frac{r_*^k}{a''}r_t^k. \tag{169}$$

Here a'' is the derivative of the per-unit-of-capital cost function $a(u_t)$, evaluated at the steady state utilization rate. The central bank follows a standard feedback rule:

$$R_t = \rho_R R_{t-1} + (1-\rho_R)(\psi_1 \pi_t + \psi_2 \hat{y}_t) + \sigma_R \varepsilon_{R,t}, \tag{170}$$

where $\varepsilon_{R,t}$ represents monetary policy shocks. The aggregate resource constraint is given by:

$$\hat{y}_t = (1+g_*)\left[\frac{c_*}{y_*}c_t + \frac{i_*}{y_*}\left(i_t + \frac{r_*^k}{e^\gamma - 1 + \delta}u_t\right)\right] + g_t. \tag{171}$$

Here c_*/y_* and i_*/y_* are the steady state consumption-output and investment-output ratios, respectively, and $g_*/(1+g_*)$ corresponds to the government share of the aggregate output. The process g_t can be interpreted as the exogenous government spending shock. It is assumed that fiscal policy is passive, in the sense that the government uses lump-sum taxes to satisfy its period budget constraint.

There are seven exogenous disturbances in the model, and six of them are assumed to follow AR(1) processes:

$$\begin{aligned} a_t &= \rho_a a_{t-1} + (1-\rho_a)\gamma + \sigma_a \varepsilon_{a,t} \\ \mu_t &= \rho_\mu \mu_{t-1} + \sigma_\mu \varepsilon_{\mu,t} \\ \lambda_{f,t} &= \rho_{\lambda_f} \lambda_{f,t-1} + \sigma_{\lambda_f} \varepsilon_{\lambda_f,t} \\ g_t &= \rho_g g_{t-1} + \sigma_g \varepsilon_{g,t} \\ b_t &= \rho_b b_{t-1} + \sigma_b \varepsilon_{b,t} \\ \phi_t &= \rho_\phi \phi_{t-1} + \sigma_\phi \varepsilon_{\phi,t}. \end{aligned} \tag{172}$$

We assume that the innovations of these exogenous processes, as well as the monetary policy shock $\varepsilon_{R,t}$, are independent standard normal random variates, and collect them in the vector ε_t. We stack all of the DSGE model parameters in the vector θ. The equations presented in this section form a linear rational expectations system that can be solved numerically, for instance using the method described by Sims (2002).

3 Econometric Methodology

Our econometric analysis proceeds in three steps. First, we use Bayesian methods to estimate the linearized DSGE model described in Section 2 on seven core macroeconomic time series. Second, we estimate so-called auxiliary regression equations that link the state-variables associated with the DSGE model to various other macroeconomic variables which are of interest to the analyst but are not explicitly included in the structural DSGE model (non-core variables). Finally, we use the estimated DSGE model to forecast its state variables, and then map these state forecasts into predictions for the core and non-core variables.

3.1 DSGE Model Estimation

The solution of the linear rational expectations system given in Section 2 can be expressed as a vector autoregressive law of motion for a vector of non-redundant state variables s_t:

$$s_t = \Phi_1(\theta)s_{t-1} + \Phi_\varepsilon(\theta)\varepsilon_t \tag{173}$$

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

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

The variables c_t, i_t, $\bar{k}_t$, R_t, and w_t are endogenous state variables, whereas the remaining elements of s_t are exogenous state variables. When estimating the DSGE model based on a sequence of observations $Y^T = [y_1, \ldots, y_T]$, it is convenient to construct a state-space model by specifying a system of measurement equations that link the observables y_t to the states s_t.

The vector y_t used in our empirical analysis consists of quarter-to-quarter growth rates (measured in percentages) of real GDP, consumption, investment and nominal wages, as well as a measure of the number of hours worked, GDP deflator inflation, and the federal funds rate. Since some of our observables include growth rates, we augment the set of model states s_t by lagged values of output, consumption, investment, and real wages. More specifically, notice that lagged consumption, investment, and real wages are elements of the vector s_{t-1}. Moreover, according to the DSGE model solution, the lagged output, $\hat{y}_{t-1}$, can be expressed as a linear function of the elements of s_{t-1}. Thus, we can write

$$[\hat{y}_{t-1}, c_{t-1}, i_{t-1}, w_{t-1}]' = M_S(\theta)s_{t-1}$$

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

$$\varsigma_t = [s_t', s_{t-1}', M_S'(\theta)]'. \tag{174}$$

This allows us to express the set of measurement equations as

$$y_t = A_0(\theta) + A_1(\theta)\varsigma_t. \tag{175}$$

The state space representation of the DSGE model is comprised of Eqs. (173)-(175).

Under the assumption that the innovations ε_t are normally distributed, the likelihood function for the DSGE model, denoted by $p(Y^T \mid \theta)$, can be evaluated using the Kalman filter. The Kalman filter also generates a sequence of estimates of the state vector ς_t:

$$\varsigma_{t|t}(\theta) = E[\varsigma_t \mid \theta, Y^t], \tag{176}$$

where $Y^t = [y_1, \ldots, y_t]$. Our Bayesian estimation of the DSGE model combines a prior $p(\theta)$ with the likelihood function $p(Y^T \mid \theta)$ in order to obtain a joint probability density function for data and parameters. The posterior distribution is given by

$$p(\theta \mid Y^T) = \frac{p(Y^T \mid \theta)p(\theta)}{p(Y^T)} \tag{177}$$

where $p(Y^T) = \int p(Y^T \mid \theta)p(\theta)d\theta$.

We employ the Markov chain Monte Carlo (MCMC) methods described in detail by An and Schorfheide (2007) to implement the Bayesian inference. More specifically, a random-walk Metropolis algorithm is used to generate draws from the posterior distribution $p(\theta \mid Y^T)$, and averages of these draws (and suitable transformations) serve as approximations for the posterior moments of interest.

3.2 Linking Model States to Non-Core Variables

Due to the general equilibrium structure, the variables that are included in state-of-the-art DSGE models are limited to a set of core macroeconomic indicators. However, in practice an analyst might be interested in forecasting a broader set of time series. For instance, the DSGE model described in Section 2 generates predictions for the numbers of hours worked, but does not include unemployment as one of the model variables. We use z_t to denote a particular variable that is not included in the DSGE model but is of interest to the forecaster nonetheless. We will express z_t as a function of the DSGE model state variables s_t. According to Eq. (174), one can easily recover s_t from the larger vector ς_t using a selection matrix M with the property $s_t = M\varsigma_t$. As was discussed in the previous subsection, the Kalman filter delivers a sequence $\varsigma_{t|t}(\theta)$, $t = 1, \ldots, T$. We use $\hat{\varsigma}_{t|t}$ to denote an estimate of $\varsigma_{t|t}(\theta)$ that is obtained by replacing θ with the posterior mean estimate $\hat{\theta}_T$. Define $\hat{s}_{t|t} = M\hat{\varsigma}_{t|t}$, and let[23]

$$z_t = \alpha_0 + \hat{s}'_{t|t}\alpha_1 + \xi_t, \qquad \xi_t = \rho\xi_{t-1} + \eta_t, \qquad \eta_t \sim N(0, \sigma_\eta^2). \qquad (178)$$

[23] At this point it is important to ensure that the state vector does not contain redundant elements, since if it did, the auxiliary regression (Eq. (178)) would suffer from perfect collinearity.

Moreover, ξ_t is a variable-specific noise process. The parameters of this auxiliary regression are collected in the vector $\psi = [\alpha_0, \alpha_1', \rho, \sigma_\eta]'$. As for the estimation of the DSGE model, we use Bayesian methods for the estimation of the auxiliary regression (Eq. (178)).

A few remarks about our setup are in order. First, Eqs. (173)–(175), and (178) can be interpreted as a factor model. The factors are given by the state variables of the DSGE model, the measurement equation associated with the DSGE model describes how our core macroeconomic variables load on the factors, and auxiliary regressions of the form of Eq. (178) describe how additional (non-core) macroeconomic variables load on the factors. The random variable ξ_t in Eq. (178) plays the role of an idiosyncratic error term.

Second, our setup can be viewed as a simplified version of BG's framework. Unlike BG, we do not attempt to estimate the DSGE model and the auxiliary equations simultaneously. While we are therefore ignoring any information about s_t which is contained in the z_t variables, our analysis reduces the computational burden considerably and can be used for real time forecasting more easily. The BG approach is computationally cumbersome. A Markov chain Monte Carlo algorithm has to iterate over the conditional distributions of θ, ψ, and the sequence of states $S^T = [s_1, \ldots, s_T]$. Drawing from the posterior of S^T is computationally costly because it requires forward and backward iterations of the Kalman filter. Drawing from the distribution of θ requires a Metropolis–Hastings step, and, unlike in a standalone estimation of the DSGE model, the proposal density needs to be tailored as a function of ψ. In turn, it is more difficult to ensure that the resulting Markov chain mixes properly and converges to its ergodic distribution at a sufficiently fast rate. Our framework de-couples the estimation of the DSGE model and the analysis of the auxiliary regressions. If necessary, additional non-core variables can easily be analyzed without the DSGE model having to be re-estimated. We view this as a useful feature for real-time applications.

Third, in addition to ignoring the information in the z_t s about the latent states, we take one more shortcut. Rather than using estimates of $s_{t|t}$ that depend on θ, we condition on the posterior mean of θ in our construction of $\hat{s}_{t|t}$. As a consequence, our posterior draws of the DSGE and auxiliary model parameters are uncorrelated, and we potentially understate the posterior uncertainty about ψ. However, in practice we have found that there are few gains from using a more elaborate sampling procedure.

We proceed by re-writing Eq. (178) in a quasi-differenced form as

$$\begin{aligned} z_1 &= \alpha_0 + \hat{s}'_{1|1}\alpha_1 + \xi_1 \\ z_t &= \alpha_0 + \alpha_0(1-\rho) + [\hat{s}'_{t|t} - \hat{s}'_{t-1|t-1}]\alpha_1 + \eta_t, \qquad t = 2,\ldots,T \end{aligned} \tag{179}$$

Instead of linking the distribution of ξ_1 to the parameters ρ and σ_η^2, we assume that $\xi_1 \sim N(0,\tau^2)$ and discuss the choice of τ below. A particular advantage of the Bayesian framework is that we can use the DSGE model to derive a prior distribution for the α s for any variables z_t that are conceptually related to variables that appear in the DSGE model. Let $\alpha = [\alpha_0, \alpha_1']'$. Our prior takes the form

$$\alpha \sim N(\mu_{\alpha,0}, V_{\alpha,0}), \qquad \rho \sim U(-1,1), \qquad \sigma_\eta \sim IG(\nu,\tau), \tag{180}$$

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

We choose the prior mean $\mu_{\alpha,0}$ based on the DSGE model's implied factor loadings for a model variable, say $z_t^\dagger$, that is conceptually similar to z_t. For concreteness, suppose that z_t corresponds to PCE inflation. Since there is only one type of final good, our DSGE model does not distinguish between, say, the GDP deflator and a price index of consumption expenditures. Hence, a natural candidate for $z_t^\dagger$ is final good inflation. Let $E_\theta^D[.]$ denote an expectation taken under the probability distribution generated by the DSGE model, conditional on the parameter vector θ. We construct $\mu_{\alpha,0}$ using a population regression of the form

$$\mu_{\alpha,0} = \left(E_\theta^D[\tilde{s}_t\tilde{s}_t']\right)^{-1} E_\theta^D[\tilde{s}_t z_t^\dagger], \tag{181}$$

where $\tilde{s}_t = [1, s_t']'$ and in practice θ is replaced by its posterior mean $\hat{\theta}_T$. If $z_t^\dagger$ is among the observables, then this procedure recovers the corresponding rows of $A_0(\theta)$ and $A_1(\theta)$ in the measurement equation (175). Details on the choice of $z_t^\dagger$ are provided in the empirical section. Our prior covariance matrix is diagonal with the following elements

$$diag(V_{\alpha,0}) = \left[\lambda_0, \frac{\lambda_1}{\omega_1}, \ldots, \frac{\lambda_1}{\omega_J}\right]. \tag{182}$$

Here 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. We scale the diagonal elements of $V_{\alpha,0}$ by ω_j^{-1}, $j = 1, \ldots, J$, where ω_j denotes the DSGE model's implied variance of the jth element of $\hat{s}_{t|t}$ (evaluated at the posterior mean of θ).[24] Draws from the posterior distribution can easily be obtained using the Gibbs sampler described in Appendix.

3.3 Forecasting

Suppose that the forecast origin coincides with the end of the estimation sample, denoted by T. Forecasts from the DSGE model are generated by sampling from the posterior predictive distribution of y_{T+h}. For each posterior draw $\theta^{(i)}$ we start from $\hat{\varsigma}_{T|T}(\theta^{(i)})$ and draw a random sequence $\left\{\varepsilon_{T+1}^{(i)}, \ldots, \varepsilon_{T+h}^{(i)}\right\}$. We then iterate the state transition equation forward to construct

$$\begin{aligned} s_{T+h|T}^{(i)} &= \Phi_1(\theta^{(i)}) s_{T+h-1|T}^{(i)} + \Phi_\varepsilon(\theta^{(i)}) \varepsilon_{T+h}^{(i)}, \qquad h = 1, \ldots, H \\ \varsigma_{T+h|T}^{(i)} &= \left[s_{T+h|T}^{(i)}{}', s_{T+h-1|T}^{(i)}{}' M_S'(\theta^{(i)}) \right]'. \end{aligned} \tag{183}$$

Finally, we use the measurement equation to compute

$$y_{T+h|T}^{(i)} = A_0(\theta^{(i)}) + A_1(\theta^{(i)}) \varsigma_{T+h|T}^{(i)}. \tag{184}$$

---

[24] Instead of assuming that the elements of α are independent, one could use the inverse of the covariance matrix of $\hat{s}_{t|t}$ to construct a non-diagonal prior covariance matrix for α. To the extent that some of the elements of s_t are highly correlated, such a prior will be highly non-informative in the corresponding directions of the α parameter space. We found this feature unattractive and decided to proceed with a diagonal $V_{\alpha,0}$.

The posterior mean forecast $\hat{y}_{T+h|T}$ is obtained by averaging the $y^{(i)}_{T+h|T}$ s.

A draw from the posterior predictive distribution of a non-core variable z_{T+h} is obtained as follows. Using the sequence $s^{(i)}_{T+1|T},\ldots,s^{(i)}_{T+H|T}$ constructed in Eq. (183), we iterate the quasi-differenced version (Eq. (179)) of the auxiliary regression forward:

$$z^{(i)}_{T+h|T} = \rho^{(i)} z^{(i)}_{T+h-1} + \alpha^{(i)}_0 (1-\rho^{(i)}) + \left(s^{(i)}_{T+h|T}{}' - s^{(i)}_{T+h-1|T}{}' \rho^{(i)} \right) \alpha^{(i)}_1 + \eta^{(i)}_{T+h},$$

where the superscript i for the parameters of Eq. (178) refers to the ith draw from the posterior distribution of ψ, and $\eta^{(i)}_{T+h}$ is a draw from a $N(0,\sigma^{2(i)}_\eta)$. The point forecast $\hat{z}_{T+h|T}$ is obtained by averaging the $z^{(i)}_{T+h|T}$ s. While our draws from the posterior distribution of θ and ψ are independent, we still maintain much of the correlation in the joint predictive distribution of y_{T+h} and z_{T+h}, because the ith draw is computed from the same realization of the state vector $s^{(i)}_{T+h|T}$.

4 Empirical Application

We use post-1983 US data to recursively estimate the DSGE model and the auxiliary regression equations and to generate pseudo-out-of-sample forecasts. We begin with a description of our data set and the prior distribution for the DSGE model parameters. Next, we discuss the estimates of the DSGE model parameters and its forecast performance for the core variables. Third, we estimate the auxiliary regressions and examine their forecasts of PCE inflation, core PCE inflation, the unemployment rate, and housing starts. Finally, we explore the multivariate aspects of the predictive distribution generated by our model. We report conditional forecast error statistics and illustrate the joint predictive distribution, as well as the propagation of a monetary policy shock to the core and non-core variables.

4.1 Data and Priors

Seven series are included in the vector of core variables y_t that is used for the estimation of the DSGE model: the growth rates of output, consumption, investment, and nominal wages, as well as the levels of hours worked, inflation, and the nominal interest rate. These series are obtained from Haver Analytics (Haver mnemonics are in italics). Real output is computed by dividing the nominal series (*GDP*) by the population 16 years and older (*LN16N*) as well as the chained-price GDP deflator (*JGDP*). Consumption is defined as nominal personal consumption expenditures (*C*) less the consumption of durables (*CD*). We divide by *LN16N* and deflate using *JGDP*. Investment is defined as *CD* plus the nominal gross private domestic investment (*I*). It is converted to real per-capita terms similarly. We compute quarter-to-quarter growth rates as the log difference of the real per capita variables and multiply the growth rates by 100 to convert them into percentages.

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

Observations for the non-core variables were also obtained from Haver Analytics. We consider PCE inflation, core PCE inflation, the unemployment rate, and housing starts as candidates for z_t in this chapter. We extract quarterly data on the chain price index for personal consumption expenditures (*JC*) and personal consumption expenditures less food and energy (*JCXF*). Inflation rates are calculated as 100 times the log difference of the series. The unemployment rate measure is the civilian unemployment rate for ages 16 years and older (*LR*). Finally, housing starts are defined as millions of new privately owned housing units started (*HST*).We use quarterly averages of seasonally adjusted monthly data, converted to an annual rate.

Our choice of the prior distribution for the DSGE model parameters follows DSSW and the specification of what is called a "standard" prior by Del Negro and Schorfheide (2008). The prior is summarized in the first four columns of Table 1 and Table 2. To make this chapter self-contained we briefly review some of the details of the prior elicitation.

Table 1. Prior and Posterior of DSGE Model Parameters: Part 1

Parameter	Prior			Posterior	
	Density	Para 1	Para 2	Mean	90% Interval
Household					
h	Beta	0.70	0.05	0.65	[0.58, 0.72]
a''	Gamma	0.20	0.10	0.30	[0.13, 0.47]
ν_l	Gamma	2.00	0.75	2.29	[1.33, 3.28]
ζ_w	Beta	0.60	0.20	0.25	[0.15, 0.35]
$400(1/\beta-1)$	Gamma	2.00	1.00	1.034	[0.45, 1.60]
Firms					
α	Beta	0.33	0.10	0.20	[0.15, 0.24]
ζ_p	Beta	0.60	0.20	0.66	[0.53, 0.84]
S''	Gamma	4.00	1.50	2.29	[0.84, 3.91]
λ_f	Gamma	0.15	0.10	0.14	[0.01, 0.26]
Monetary policy					
$400\pi_*$	Normal	3.00	1.50	2.94	[2.08, 3.78]
ψ_1	Gamma	1.50	0.40	3.05	[2.43, 3.68]
ψ_2	Gamma	0.20	0.10	0.06	[0.03, 0.10]
ρ_R	Beta	0.50	0.20	0.86	[0.83, 0.89]

Para 1 and Para 2 list the means and standard deviations for the Beta, Gamma, and Normal distributions; the upper and lower bound of the support for the Uniform distribution; and s and ν for the Inverse Gamma (InvGamma) distribution, where $p_{IG}(\sigma \mid \nu, s) \propto \sigma^{-(\nu+1)} e^{-\nu s^2/2\sigma^2}$. The joint prior distribution is obtained as a product of the marginal distributions tabulated in the table, with this product truncated 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: 1984:I to 2007:III.

The priors for parameters that affect the steady state relationships, e.g. the capital share α in the Cobb–Douglas production function or the capital depreciation rate, are chosen to be commensurate with pre-sample (1955 to 1983) averages in the US data. The priors for the parameters of the exogenous shock processes are chosen such that the implied variance and persistence of the endogenous model variables is broadly consistent with the corresponding pre-sample moments. Our priors for the Calvo parameters that control the degree of nominal rigidity are fairly agnostic and span values that imply flexible as well as rigid prices and wages.

Table 2. Prior and Posterior of DSGE Model Parameters: Part 2

Parameter	Prior			Posterior	
	Density	Para 1	Para 2	Mean	90% Interval
Shocks					
400γ	Gamma	2.00	1.00	1.57	[1.13, 2.02]
g_*	Gamma	0.30	0.10	0.29	[0.13, 0.43]
ρ_a	Beta	0.20	0.10	0.19	[0.10, 0.29]
ρ_μ	Beta	0.80	0.05	0.80	[0.74, 0.87]
ρ_{λ_f}	Beta	0.60	0.20	0.67	[0.30, 0.94]
ρ_g	Beta	0.80	0.05	0.96	[0.95, 0.98]
ρ_b	Beta	0.60	0.20	0.85	[0.78, 0.93]
ρ_ϕ	Beta	0.60	0.20	0.98	[0.96, 0.99]
σ_a	InvGamma	0.75	2.00	0.62	[0.54, 0.69]
σ_μ	InvGamma	0.75	2.00	0.53	[0.38, 0.68]
σ_{λ_f}	InvGamma	0.75	2.00	0.18	[0.15, 0.21]
σ_g	InvGamma	0.75	2.00	0.33	[0.29, 0.37]
σ_b	InvGamma	0.75	2.00	0.36	[0.28, 0.45]
σ_ϕ	InvGamma	4.00	2.00	2.90	[1.99, 3.80]
σ_R	InvGamma	0.20	2.00	0.14	[0.12, 0.16]

Notes: see Table 1, p. 153

Our prior for the central bank's response to inflation and output movements is roughly centered at Taylor's (1993) values. The prior for the interest rate smoothing parameeer ρ_R is almost uniform on the unit interval. The 90% interval for the prior distribution on ν_l implies that the Frisch labor supply elasticity lies between 0.3 and 1.3, reflecting the micro-level estimates at the lower end, and the estimates of Chang and Kim

(2006) and Kimball and Shapiro (2008) at the upper end. The density for the adjustment cost parameter S'' spans the values that Christiano et al. (2005) find when matching DSGE and vector autoregression (VAR) impulse response functions. The density for the habit persistence parameter h is centered at 0.7, which is the value used by Boldrin, Christiano, and Fisher (2001). They find that $h = 0.7$ enhances the ability of a standard DSGE model to account for key asset market statistics. The density for a'' implies that utilization rates rise by 0.1%–0.3% in response to a 1% increase in the return to capital.

4.2 DSGE Model Estimaton and Forecasting of Core Variables

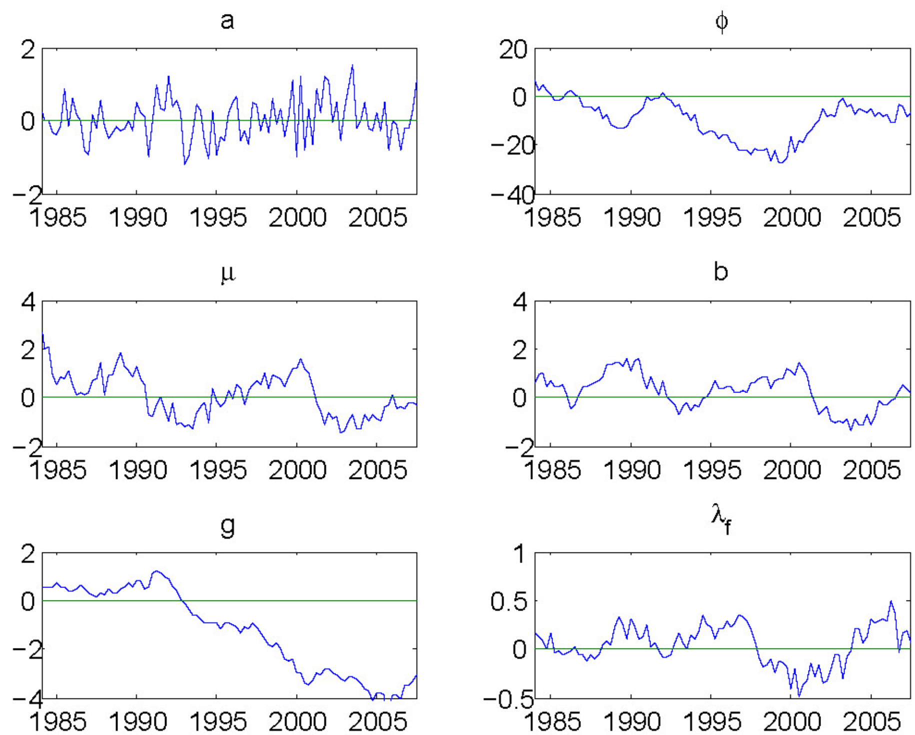
The first step of our empirical analysis is to estimate the DSGE model. While we estimate the model recursively, starting with the sample 1984:I to 2000:IV and ending with the sample 1984:I to 2007:III, we will focus our discussion of the parameter estimates on the final estimation sample. Summary statistics for the posterior distribution (means and 90% probability intervals) are provided in Table 1 and Table 2. For long horizon forecasts, the most important parameters are γ, π_* and β. Our estimate of the average technology growth rate implies that output, consumption, and investment all grow at an annualized rate of 1.6%. According to our estimates of π_* and β, the target inflation rate is 2.9% and the long-run nominal interest rate is 5.5%. The cross-equation restrictions of our model generate a nominal wage growth of about 4.5%.

Our policy rule estimates imply a strong response of the central bank to inflation $\hat{\psi}_1 = 3.05$ and a tempered reaction to deviations of output from its long-run growth path $\hat{\psi}_2 = 0.06$. As was discussed by Del Negro and Schorfheide (2008), estimates of wage and price stickiness based on aggregate price and wage inflation data tend to be somewhat fragile. We obtain $\hat{\zeta}_p = 0.66$ and $\hat{\zeta}_w = 0.25$, which means that wages are nearly flexible and the price stickiness is moderate. According to the estimated Calvo parameter, firms re-optimize their prices every three quarters.

The technology growth shocks have very little serial correlation, and the estimated innovation standard deviation is about 0.6%. These estimates are consistent with direct calculations based on Solow residuals. At an annualized rate, the monetary policy shock has a standard deviation of 56 basis points. Both the government spending shock g_t and the labor supply shock ϕ_t have estimated autocorrelations near unity. The labor supply shock captures much of the persistence in the hours series.

We proceed by plotting estimates of the exogenous shocks in Figure 1. These shocks are included in the vector $s_t = M\varsigma_t$ that is used as regressor in the auxiliary model (178). Formally, we depict filtered latent variables, $\hat{s}_{j,t|t}$, conditional on the posterior mean $\hat{\theta}_T$ for the period 1984:I to 2007:III. In line with the parameter estimates reported in Table 1 and Table 2, the filtered technology growth process appears to be essentially *iid*. The processes g_t and ϕ_t exhibit long-lived deviations from zero, and partially capture low frequency movements of the exogenous demand components and hours worked, respectively. μ_t is the investment-specific technology shock. Its low frequency movements capture trend differentials in output, consumption, and investment.

Figure 1. Latent State Variables of the DSGE Model



Notes: The six panels of the figure depict time series of the elements of $\hat{s}_{t|t}$.
Estimation sample: 1984:I to 2007:III.

At this point a comparison between our estimates of the latent shock processes and the estimates reported by BG is instructive. By construction, our filtered state variables $\hat{s}_{t|t}$ are moving averages of the observables y_t. In contrast, BG's estimates of the latent states are functions not only of y_t (in our notation), but also of all of the other observables included in their measurement equations, namely numerous measures of inflation as well as 25 principal components constructed from about 70 macroeconomic time series. Due to differences in the model specification and data definitions, it is difficult to directly compare our estimates of the latent states with those reported by BG.

However, BG overlay smoothed states obtained from the direct estimation of their DSGE model with estimates obtained from their DSGE-DFM. The main difference between the estimated DSGE and DSGE-DFM states is that some of the latter, namely productivity, preferences, and government spending, are a lot smoother. The most likely reason for this is that the DSGE-DFM measurement equations for the seven core variables contain autoregressive measurement errors, which absorb some of the low frequency movements in these series.

Table 3 summarizes pseudo-out-of-sample root mean squared error (RMSE) statistics for the seven core variables that are used to estimate the DSGE model: the growth rates of output, consumption, investment and nominal wages, as well as log hours worked, GDP deflator inflation, and the federal funds rate. We report RMSEs for horizons h = 1; 2; 4 and 12, and compare the DSGE model forecasts to those from an AR(1) model, which is recursively estimated by OLS.[25] The h-step-ahead growth rate (inflation) forecasts refer to percentage changes between periods $T+h-1$ and $T+h$. Boldface entries in the table indicate that the RMSE of the DSGE model is lower than that of the AR(1) model. We used the Harvey, Leybourne, and Newbold (1998) version of the Diebold and Mariano (1995) test for equal forecast accuracy of the DSGE and the AR(1) model, employing a quadratic loss function. However, due to the fairly short forecast period, most of the loss differentials are insignificant.

[25] The h-step-ahead forecast is generated by iterating one-step-ahead predictions forward, ignoring parameter uncertainty: $\hat{y}_{i,T+h|T} = \hat{\beta}_{0,OLS} + \hat{\beta}_{1,OLS}\hat{y}_{i,T+h-1|T}$, where the OLS estimators are obtained from the regression $y_{i,t} = \beta_0 + \beta_1 y_{i,t-1} + u_{i,t}$.

Table 3. RMSE Comparison: DSGE Model versus AR(1)

Series	Model	$h = 1$	$h = 2$	$h = 4$	$h = 12$
Output Growth (Q%)	DSGE	0.51	0.50	**0.41**	**0.36**
	AR(1)	0.50	0.49	0.44	0.37
Consumption Growth (Q%)	DSGE	0.39	0.38	0.39	0.39
	AR(1)	0.37	0.37	0.34	0.31
Investment Growth (Q%)	DSGE	**1.44**	**1.56**	**1.47****	**1.52**
	AR(1)	1.56	1.67	1.60	1.60
Nominal Wage Growth (Q%)	DSGE	0.67	0.70	0.66	**0.56**
	AR(1)	0.59	0.59	0.59	0.56
100 x log Hours	DSGE	**0.52****	**0.88****	**1.44****	**2.07****
	AR(1)	0.66	1.20	2.08	3.40
Inflation (Q%)	DSGE	**0.22**	**0.23**	**0.19****	0.24
	AR(1)	0.22	0.23	0.22	0.23
Interest Rate (A%)	DSGE	0.71	1.34	2.13	**2.25**
	AR(1)	0.54**	1.00**	1.73	2.93

We report RMSEs for the DSGE and AR(1) models. Numbers in boldface indicate a lower RMSE of the DSGE model. The RMSEs are computed based on recursive estimates, starting with the sample 1984:I to 2000:IV and ending with the samples 1984:I to 2007:III ($h = 1$), 1984:I to 2007:II ($h = 2$), 1984:I to 2006:III ($h = 4$), and 1984:I to 2004:III ($h = 12$), respectively. The h-step-ahead growth (inflation) rate forecasts refer to percentage changes between the periods $T + h - 1$ and $T + h$.
* (**) indicates significance of the two-sided modified Diebold–Mariano test of equal predictive accuracy under quadratic loss at the 10% (5%) level.

The RMSEs for one-quarter-ahead forecasts of output and consumption obtained from the estimated DSGE model are only slightly larger than those associated with the AR(1) forecasts. The DSGE model generates lower RMSEs for forecasts of investment and hours worked, while the RMSEs for inflation rates are essentially identical across the two models. The AR(1) model performs better than the DSGE model for forecasting nominal wage growth and interest rates. The accuracy of long-run forecasts is sensitive to mean growth estimates, which are restricted to be equal for output, consumption, and investment. Moreover, the DSGE model implies that the nominal wage growth equals output plus inflation growth in the long-run.

In Table 4 we are comparing the pseudo-out-of-sample RMSEs obtained using our estimated DSGE model to those reported in three other studies, namely those of (i) DSSW, (ii) Edge et al. (EKL, 2009), and (iii) Smets and Wouters (2007). Since the studies all differ with respect to the forecast period, we report sample standard deviations over the respective forecast periods, computed from our data set. Unlike the other three studies, EKL use real time data.

Table 4. One-Step-Ahead Forecast Performance of DSGE Models

Study	Forecast Period	Output Growth Q%	Inflation Q%	Interest Rate A%
Shorfheide, Sill, Kryshko	2001:I to 2007:IV	0.51	0.22	0.71
		(0.47)	(0.22)	(1.68)
DSSW	1985:IV to 2000:I	0.73	0.27	0.87
		(0.52)	(0.25)	(1.72)
Edge et al. (2009)	1996:III to 2004:IV	0.45	0.29	0.83
		(0.57)	(0.20)	(1.96)
Smets, Wouters (2007)	1990:I to 2004:IV	0.57	0.24	0.43
		(0.57)	(0.22)	(1.97)

Schorfheide, Sill, Kryshko: RMSEs, the DSGE model is estimated recursively with data starting in 1984:I. DSSW (2007, Table 2): RMSEs, VAR approximation of the DSGE model estimated based on rolling samples of 120 observations. Edge et al. (2009, Table 5) RMSEs, the DSGE model is estimated recursively using real time data starting in 1984:II. Smets andWouters (2007, Table 3): RMSEs, the DSGE model is estimated recursively, starting with data from 1966:I. The numbers in parentheses are sample standard deviations for the forecast period, computed from the Schorfheide, Sill, Kryshko data set. Q% is the quarter-to-quarter percentage change, and A% is an annualized rate.

Overall, the RMSEs reported by DSSW are slightly worse than those in the other three studies. This might be due to the fact that DSSW use a rolling window of 120 observations to estimate their DSGE model and start forecasting in the mid 1980s, whereas the other papers let the estimation sample increase and start forecasting in the 1990s. Only EKL are able to attain an RMSE for output growth that is lower than the sample standard deviation. The RMSEs for the inflation forecasts range from 0.22 to 0.29 and are very similar across studies. They are only slightly larger than the sample standard deviations. Finally, the interest rate RMSEs are substantially lower than the sample standard deviations, because the forecasts are able to exploit the high persistence of the interest rate series.

4.3 Forecasting Non-Core Variables with Auxiliary Regressions

We now turn to the estimation of the auxiliary regressions for PCE inflation, core PCE inflation, the unemployment rate, and housing starts. The following elements are included in the vector s_t, which appears as regressor in Eq. (178):

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

To construct a prior mean for α_1, we link each z_t with a conceptually related DSGE model variable $z_t^\dagger$ and use Eq. (181). More specifically, we link the two measures of PCE inflation to the final good inflation π_t, the unemployment rate to a scaled version of log hours worked L_t, and housing starts to scaled percentage deviations i_t of investment from its trend path; see Table 5 below. Our DSGE model has only one final good, which is domestically produced and used for both consumption and investment. Hence, using the same measurement equation for both inflation in consumption expenditures and GDP seems reasonable. Linking the unemployment rate with the hours worked can be justified by the observation that most of the variation in the hours worked over the business cycle is due to changes in employment rather than variation along the intensive margin. Finally, housing starts can be viewed as a measure of investment, namely investment in residential structures. Since the housing starts series has no apparent trend, we link it to investment deviations from trend.

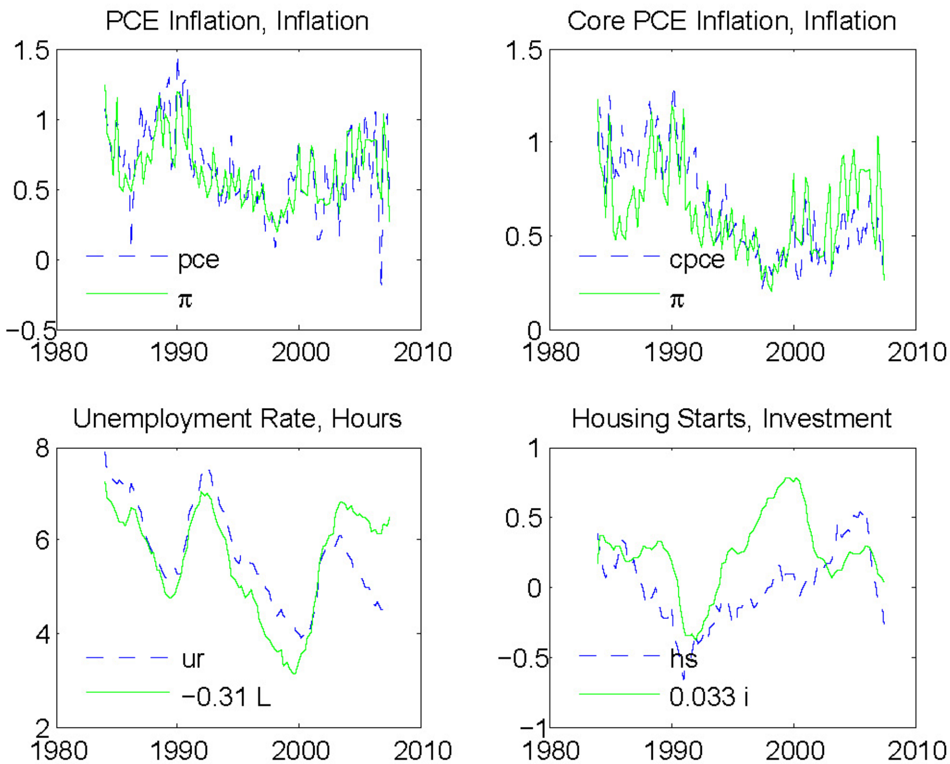
Table 5. Non-Modelled and Related DSGE Model Variables

Non-Modelled Variable	DSGE Model Variable	Transformation
PCE Inflation	Final good inflation π_t	None
Core PCE Inflation	Final good inflation π_t	None
Unemployment Rate	Hours worked L_t	$-0.31L_t$
Housing Starts	Investment i_t	$0.033i_t$

Here, π_t, L_t and i_t are the DSGE model variables that appear in the DSGE model description in Section 2 of this chapter.

The four panels of Figure 2 depict the sample paths of the non-core variables z_t and the related DSGE model variables $z_t^\dagger$. The GDP deflator and hours worked are directly observable, while the investment series i_t is latent and obtained from $\hat{s}_{t|t}$. The inflation measures are highly correlated. PCE inflation is more and core PCE inflation less volatile than GDP deflator inflation. In the bottom left panel we re-scale and re-center log hours such that it is commensurate with the unemployment rate. These two series are also highly correlated. The bottom right panel shows that the investment series implied by the DSGE model is somewhat smoother than the housing starts series. However, except for the period from 2000 to 2002, the low frequency movements of the two series are at least qualitatively similar.

Figure 2. Non-Core Variables and Related DSGE Model Variables



Notes: The top two panels show quarter-to-quarter inflation rates. In the bottom panels we add constants to the scaled log hours worked and investment deviations from the trend to match the means of the unemployment rate and housing starts over the period 1984:I to 2007:III.

To proceed with the Bayesian estimation of Eq. (179) we have to specify the hyperparameters. In our framework, τ can be interpreted as the prior standard deviation of the idiosyncratic error ξ_1. We set τ equal to 0.12 (PCE inflation), 0.11 (core PCE inflation), 0.40 (unemployment rate), and 0.10 (housing starts). These values imply that the prior variance of ξ_1 is about 15% to 20% of the sample variance of z_1. We set the degrees of freedom parameter ν of the inverted gamma prior for σ_η equal to 2, and restrict $\lambda_0 = \lambda_1 = \lambda$ to one of three values: 1.00, 0.10, and 10^{-5}. The value $\lambda = 10^{-5}$

corresponds to a dogmatic prior, under which the posterior estimate and prior mean essentially coincide. As we increase λ, we allow the factor loading coefficients α to differ from the prior mean.

The estimates of the auxiliary regressions are summarized in Table 6. Rather than providing numerical values for the entire α vector, we focus on the persistence and the standard deviation of the innovation to the idiosyncratic component. By construction, $\hat{s}'_{t|t}\mu_{\alpha_1,0}$, where $\mu_{\alpha_1,0}$ is the prior mean of α_1, reproduces the time paths of the GDP deflator inflation, log hours worked, and investment deviations from trend. Thus, for $\lambda = 10^{-5}$ the idiosyncratic error term ξ_t essentially picks up the discrepancies between non-core variables and the related DSGE model variables depicted in Figure 2. For the two inflation series, the estimate of σ_η falls as we increase the hyperparameter. The larger the value of λ, the more of the variation in the variable is explained by $\hat{s}'_{t|t}\hat{\alpha}'_1$, where $\hat{\alpha}_1$ is the posterior mean of α_1. For instance, the variability of the core PCE inflation captured by the factors is five times as large as the variability due to the idiosyncratic disturbance ξ_t if λ is equal to one. This factor drops to 1.4 if the prior is tightened. For PCE inflation the idiosyncratic disturbance is virtually serially uncorrelated, whereas for core PCE inflation the serial correlation ranges from 0.2 ($\lambda = 1$) to 0.5 ($\lambda = 10^{-5}$).

For unemployment, setting $\lambda = 10^{-5}$ implies that the prior and posterior means of the factor loadings α are essentially identical. Unemployment loads on c_t, i_t, $\bar{k}_t$, μ_t, and g_t. The intuition is that output in our model can be obtained from consumption, investment, and government spending (see Eq. (171)), while the hours worked can be determined from the production function as a function of output and capital (see Eq. (161)). If the hyperparameter is raised to 0.1 or 1.0, then unemployment also loads on the interest rate, wages, and the shocks a_t and b_t. However, in general we find it difficult to interpret the estimates of particular elements of α_1, because some of the variables contained in the vector s_t are endogenous equilibrium objects which themselves respond to the exogenous state variables in turn. Hence, we will focus below on the estimate of $\hat{s}'_{t|t}\alpha_1$ and the response of z_t to structural shocks. The most striking feature of the unemployment estimates is the high persistence of ξ_t, with ρ_ξ estimates around 0.98.

Table 6. Auxiliary Regression Estimates

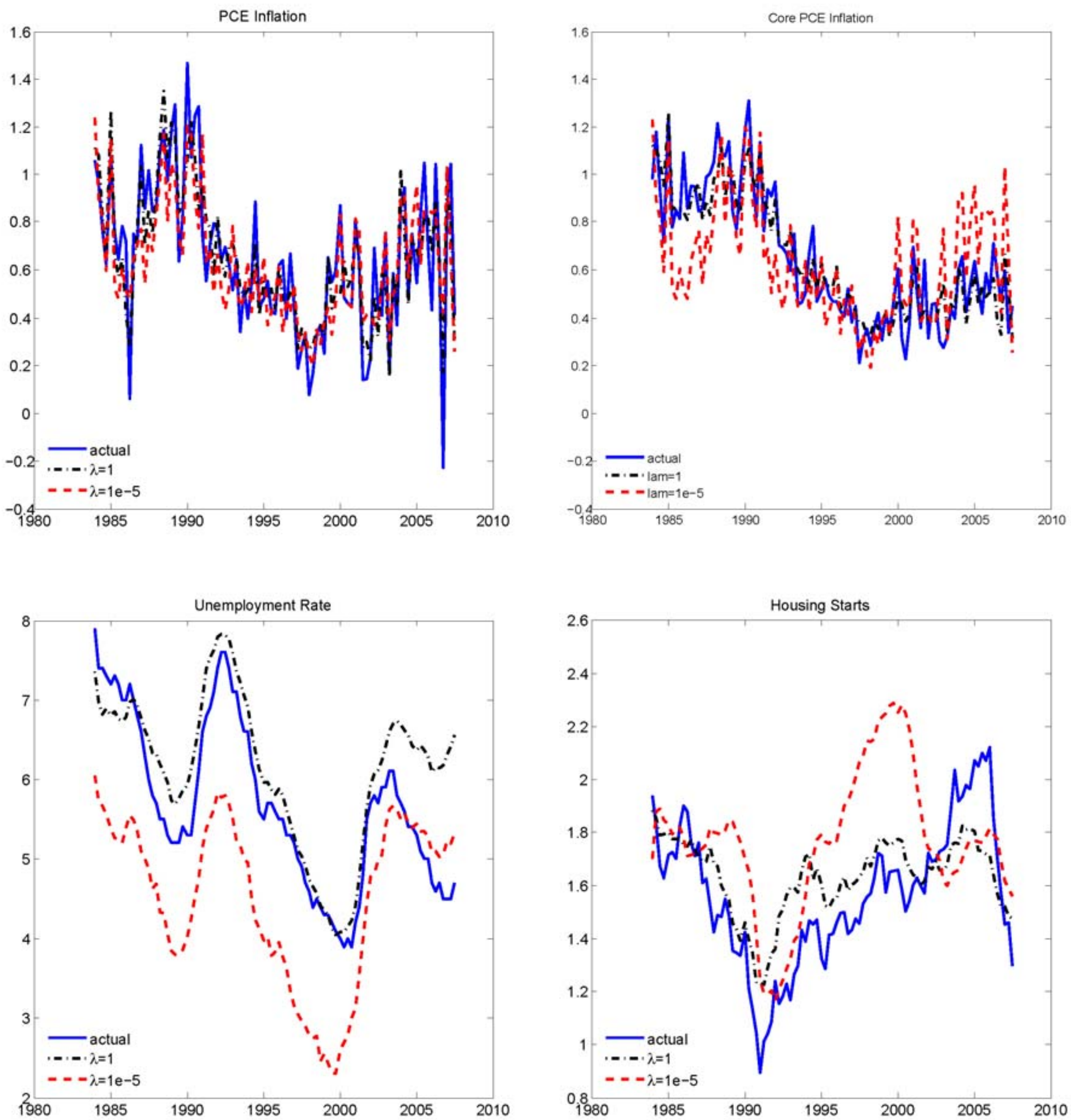
Series	λ	ρ Mean	ρ 90% interval	σ_η Mean	σ_η 90% interval	Signal/Noise $\frac{\widehat{\text{var}}(\hat{s}'_{t\|t}\hat{a}_1)}{\widehat{\text{var}}(\hat{\xi}_t)}$	$\ln p_\lambda(Z^T)$
PCE inflation	1.0	0.05	[-0.14, 0.26]	0.03	[0.02, 0.03]	3.15	-0.03
	0.1	0.05	[-0.16, 0.25]	0.03	[0.02, 0.04]	2.62	4.82
	10^{-5}	0.07	[-0.11, 0.24]	0.04	[0.02, 0.05]	1.47	12.27
Core PCE inflation	1.0	0.23	[0.03, 0.45]	0.01	[0.01, 0.02]	4.99	29.53
	0.1	0.21	[-0.02, 0.41]	0.01	[0.01, 0.02]	4.88	39.12
	10^{-5}	0.53	[0.38, 0.68]	0.03	[0.02, 0.04]	1.35	22.58
Unemployment rate	1.0	0.98	[0.96, 1.00]	0.019	[0.01, 0.02]	3.45	17.71
	0.1	0.97	[0.95, 1.00]	0.019	[0.01, 0.02]	3.67	23.68
	10^{-5}	0.98	[0.97, 1.00]	0.025	[0.02, 0.03]	1.91	22.78
Housing starts	1.0	0.89	[0.76, 1.00]	0.007	[0.00, 0.01]	0.74	68.21
	0.1	0.88	[0.74, 1.00]	0.007	[0.01, 0.01]	0.95	80.81
	10^{-5}	0.96	[0.92, 1.00]	0.009	[0.01, 0.01]	0.88	82.64

The posterior summary statistics are computed based on the output of Gibbs sampler. The sample variance ratios are computed using the posterior mean estimate of α_1. Estimation sample: 1984:I to 2007:III.

For housing starts, the measurement error process is slightly less persistent than for unemployment, but the signal-to-noise ratio is generally low, which is not surprising in view of the fairly large discrepancy between housing starts and i_t shown in the bottom right panel of Figure 2. Unlike for the other three non-core series, the lowest signal-to-noise ratio for housing starts is obtained for $\lambda = 1$. An increase in λ from 10^{-5} to 1 decreases the variability of $\hat{s}'_{t|t}\hat{a}_1$ by more than the variability of the measurement error process, as is evident from the bottom right panel of Figure 3.

Figure 3 displays the time path of $\hat{\alpha}_0 + \hat{s}'_{t|t}\hat{\alpha}_1$ for different choices of the hyperparameter. Consider the two inflation series. For $\lambda = 10^{-5}$, the factor predicted path for the two inflation rates is essentially identical and reproduces the GDP deflator inflation. As λ is increased to one, they follow the two PCE inflation measures more closely, which is consistent with the estimates of ρ and σ_η reported in Table 6. The predicted paths for the unemployment rate behave in a markedly different manner. If we set $\lambda = 1$, then the predicted path resembles the actual path fairly closely except at the end of the sample. Hence, the implied ξ_t series stays close to zero until about 2002, and then drops to about –2% between 2002 and 2006. As we decrease λ to 10^{-5}, the predicted path shifts downward. The estimate of ξ_1 is roughly 2%, and subsequently ξ_t approximately follows a random walk process that captures the gap between the path predicted by the factors and the actual unemployment series.

Figure 3. Non-Core Variables and Factors



Notes: The figure depicts the actual (blue, solid) path of the non-core variables, as well as the factor predictions $\hat{\alpha}_0 + \hat{s}_{t|t}'\hat{\alpha}_{1,T}$ for $\lambda = 10^{-5}$ (red, dashed) and $\lambda = 1$ (black, dotted).

The last column of Table 6 contains log marginal likelihood values $\ln p_{\lambda}(Z^T)$ for the four auxiliary regression models as a function of the hyperparameter λ. These values can be used to make a data-driven hyperparameter choice that trades off in-sample fit against the complexity of the regression model.[26] According to the marginal likelihoods, the preferred choice for λ is 0.1 for core PCE inflation and the unemployment rate and 10^{-5} for PCE inflation and housing starts. The log marginal data density can also be interpreted as a one-step-ahead predictive score:

$$\ln p_{\lambda}(Z^T) = \sum_{t=0}^{T-1} \int p(z_{t+1} \mid \psi, Z^t) p_{\lambda}(\psi \mid Z^t) d\psi. \tag{185}$$

Thus, we would expect the λ rankings obtained from one-step-ahead pseudo-out-of-sample forecast error statistics to be comparable to the rankings obtained from the marginal likelihoods.

Forecast error statistics for the non-modelled variables are provided in Table 7. We compare the RMSEs of the forecasts generated by our auxiliary models to two alternative models. First, as in Section 4.2, we consider an AR(1) model for z_t which is estimated by OLS and from which we generate *h*-step-ahead forecasts by iterating one-step-ahead predictions forward. Second, we consider multi-step least squares regressions of the form

$$z_t = \beta_0 + y'_{t-h}\beta_1 + z_{t-h}\beta_2 + u_t \tag{186}$$

estimated for horizons $h = 1$, 2, 4 and 12. Recall that the filtered states $\hat{s}_{t|t}$ are essentially moving averages of y_t and its lags. Hence, both Eqs. (179) and (186) generate predictions of z_{t+h} as a function of z_t as well as y_t and its lags. However, the restrictions

[26] A detailed discussion of hyperparameter selection based on marginal likelihoods is given, for instance, by DSSW.

imposed on the parameters of the implied prediction functions are very different. While our least squares estimation of Eq. (186) leaves the coefficient vector essentially unrestricted and excludes additional lags of y_t, the auxiliary regression model (179) tilts the estimates of α_1 toward loadings derived from the DSGE model, and additional lags of y_t implicitly enter the prediction through the filtered state vector.

Table 7. Root Mean Squared Errors for Auxiliary Regressions

Non-Core Series and Models	λ	$h = 1$	$h = 2$	$h = 4$	$h = 12$
PCE Inflation (Q%)					
Auxiliary model	1.00	**0.34**	0.37	0.34	**0.32**
Auxiliary model	0.10	**0.33**	**0.35**	**0.32**	0.35
Auxiliary model	10^{-5}	**0.32**	**0.34**	**0.30**	0.33
Regression		**0.33**	**0.35**	**0.32**	0.49
AR(1)		0.36	0.35	0.33	0.32
Core PCE Inflation (Q%)					
Auxiliary model	1.00	0.18	0.19	**0.16**	**0.12**
Auxiliary model	0.10	0.18	0.18	**0.15**	**0.11**
Auxiliary model	10^{-5}	0.16	0.20	**0.18**	**0.15**
Regression		**0.14**	**0.14**	**0.17**	0.35
AR(1)		0.16	0.16	0.18	0.17
Unemployment Rate (%)					
Auxiliary model	1.00	**0.16****	**0.27**	**0.43**	**1.02**
Auxiliary model	0.10	**0.15****	**0.24**	**0.39**	**0.97**
Auxiliary model	10^{-5}	**0.15****	**0.23***	**0.37**	**0.74**
Regression		**0.20**	0.37	0.72	1.39
AR(1)		0.21	0.37	0.63	1.01
Housing Starts (4 million/Q)					
Auxiliary model	1.00	0.11	0.18	0.31	0.50
Auxiliary model	0.10	0.11	0.17	0.29	0.48
Auxiliary model	10^{-5}	**0.10**	**0.16**	0.27	0.45
Regression		**0.10**	**0.16**	**0.26**	**0.43**
AR(1)		0.10	0.16	0.27	0.43

We report RMSEs for the DSGE, AR(1) and regression models. Numbers in boldface indicate that DSGE model or a regression model (186) attains a lower RMSE than AR(1) model. The RMSEs are computed based on recursive estimates, starting with the sample 1984:I to 2000:IV and ending with the samples 1984:I to 2007:III (h = 1), 1984:I to 2007:II (h = 2), 1984:I to 2006:III (h = 4), and 1984:I to

2004:III (h = 12), respectively. The h-step-ahead growth (inflation) rate forecasts refer to percentage changes between the periods $T + h - 1$ and $T + h$.
* (**) indicates significance of the two-sided modified Diebold–Mariano test of equal predictive accuracy under quadratic loss at the 10% (5%) level.

Over short horizons, our auxiliary regression models attain a lower RMSE than the AR(1) benchmark for PCE inflation, the unemployment rate, and housing starts. The improvements in the unemployment forecasts are significant. For one-step-ahead forecasts, the preferred choice of λ is 10^{-5}. For PCE inflation and housing starts, the value of λ that yields the highest marginal likelihood also generates the lowest RMSE. For the unemployment rate, the marginal likelihoods for $\lambda = 0.1$ and 10^{-5} are very similar, and so are the RMSE statistics. The only discrepancy between RMSEs and the marginal likelihood ranking arises for core PCE inflation. We conjecture that the different rankings could be due in part to the persistent deviations of core PCE inflation from $\hat{s}'_{t|t}\hat{\alpha}_1$ at the beginning of the sample, as is evident from the top right panel of Figure 3. According to Eq. (185), the predictive accuracy at the beginning of the sample affects the marginal likelihood, but it does not enter our RMSE statistics, which are computed from 2001 onward. Over longer horizons, core PCE and unemployment forecasts from our auxiliary regressions dominate the AR(1) forecasts, whereas the PCE inflation and housing starts forecasts are slightly less precise. Except for short- to medium-term core PCE inflation forecasts, our auxiliary regressions with $\lambda = 10^{-5}$ are slightly better than the forecasts obtained from the simple predictive regression (Eq. (186)).

4.4 Multivariate Considerations

So far the analysis has focused on univariate measures of forecast accuracy. A conservative interpretation of our findings and those reported elsewhere, e.g. Adolfson et al. (2007, 2008) and Edge et al. (2009), is that by and large the univariate forecast performance of DSGE models is not worse than that of competitive benchmark models, such as simple AR(1) specifications or more sophisticated Bayesian VARs. The key advantage of DSGE models, and the reason why central banks are considering them for projections and policy analysis, is that these models use modern macroeconomic theory to explain and predict the comovements of aggregate time series over the business cycle. Historical observations can be decomposed into the contributions of the underlying exogenous disturbances, such as technology, preference, government spending, or monetary policy shocks. Future paths of the endogenous variables can be constructed conditional on particular realizations of the monetary policy shocks that reflect potential future nominal interest rate paths. While it is difficult to quantify some of these desirable attributes of DSGE model forecasts and trade them off against forecast accuracy in an RMSE sense, we will focus on three multivariate aspects. First, we conduct posterior predictive checks for the correlation between core and non-core variables captured by our framework. Second, we present impulse response functions to a monetary policy shock and document the way in which the shock is transmitted to the non-core variables through our auxiliary regression equations. Third, we examine some features of the predictive density that our empirical model generates for the core and non-core variables.

Posterior predictive checks for correlations between the non-core and core variables are summarized in Table 8 for $\lambda = 10^{-5}$, which is the value of λ that leads to the lowest one-step-ahead forecast RMSE. Using the posterior draws for the DSGE and auxiliary model parameters, we simulate a trajectory of 100 z_t and y_t observations and compute sample correlations of interest. The posterior predictive distribution of these sample correlations is then summarized by 90% credible intervals. Moreover, we report sample correlations computed from US data. The empirical model captures the correlations between non-core and core variables well, provided that the actual sample correlations do not lie too far in the tails of the corresponding posterior predictive distribution. With the exception of the correlations between output growth and the unemployment rate, all of the correlations computed from US data lie inside the corresponding 90% credible sets.

Table 8. Posterior Predictive Check: Cross-Correlations

		Output Growth	Inflation	Interest Rate
PCE Inflation, $\lambda = 10^{-5}$	90% CI	[-0.46, 0.01]	[0.50, 0.91]	[0.11, 0.63]
	Data	-0.07	0.75	0.42
Core PCE Inflation, $\lambda = 10^{-5}$	90% CI	[-0.47, 0.03]	[0.50, 0.91]	[0.07, 0.63]
	Data	0.01	0.68	0.61
Unemployment Rate, $\lambda = 10^{-5}$	90% CI	[-0.32, 0.09]	[-0.26, 0.36]	[-0.24, 0.63]
	Data	0.15	0.17	0.12
Housing Starts, $\lambda = 10^{-5}$	90% CI	[-0.11, 0.33]	[-0.26, 0.33]	[-0.47, 0.43]
	Data	0.23	0.05	-0.22

We report 90% credible intervals of the posterior predictive distribution for the sample correlations of non-modelled variables with core variables. The data entries refer to sample correlations calculated from US data.

An important aspect of monetary policy making is assessing the effect of changes in the federal funds rate. In the DSGE model we represent these changes – unanticipated deviations from the policy rule – as monetary policy shocks. An attractive feature of our framework is that it generates a link between the structural shocks that drive the DSGE model and other non-modeled variables through the auxiliary regressions. We can then compute the impulse response function of z_t to a monetary policy shock as follows:

$$\frac{\partial z_{t+h}}{\partial \varepsilon_{R,t}} = \frac{\partial s'_{t+h}}{\partial \varepsilon_{R,t}} \alpha_1$$

where $\partial s'_{t+h} / \partial \varepsilon_{R,t}$ is obtained from the DSGE model.

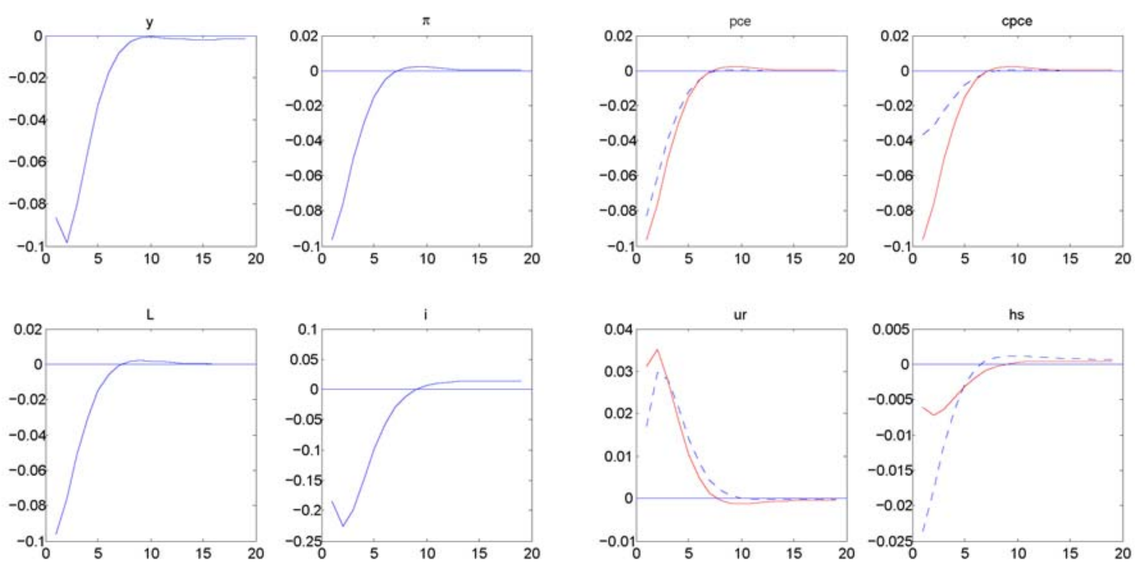
In Figure 4 we plot impulse responses of the four non-core variables (right panels) and the four related DSGE model variables (left panels: output, inflation, investment, and hours) to a one standard deviation monetary policy shock. The one standard deviation increase to the monetary policy shock translates into a 40 basis point increase in the funds rate, measured at an annual rate. The estimated DSGE model predicts that output and hours worked will drop by 10 basis points in the first quarter and return to their trend paths after seven quarters. Investment is more volatile, and drops by about 19 basis points. Quarter-to-quarter inflation falls by 10 basis points and returns to its steady state within two years. Regardless of the choice of hyperparameter, the PCE inflation responses closely resemble the GDP deflator inflation responses both qualitatively and quantitatively. The core PCE inflation, unemployment, and housing starts responses are more sensitive to the choice of hyperparameter. If λ is equal to 10^{-5} and we force the factor loadings to match those of hours worked, the unemployment rises by about 3.5

basis points one period after impact. As we relax the hyperparameter, which worsens the RMSE of the unemployment forecast, the initial effect of the monetary policy shock on unemployment is dampened. Likewise, the core PCE response drops from 10 basis points to about 4 basis points. The annualized number of housing starts drops by about 6000 units for $\lambda = 10^{-5}$ and by 22,000 units if $\lambda = 1$. Unlike for core PCE inflation, housing starts respond more strongly to a monetary policy shock if the restrictions on the factor loadings are relaxed.

Figure 4. Impulse Response to a Contractionary Monetary Policy Shock

Core variables: output, GDP deflator inflation, hours, investment

Non-core variables: PCE inflation, core PCE inflation, unemployment rate, housing starts



Notes: (i) Core variables: we depict log-level responses for output, hours and investment. (ii) Non-core variables: we overlay two responses, corresponding to the auxiliary regressions estimated with $\lambda = 10^{-5}$ (red, solid) and $\lambda = 1$ (blue, dashed). Estimation sample: 1984:I to 2007:III.

Our empirical model generates a joint density forecast for the core and non-core variables, which reflects the uncertainty about both the parameters and future realizations of shocks. A number of different methods for evaluating multivariate predictive densities exist. To assess whether the probability density forecasts are well calibrated, that is, are consistent with empirical frequencies, one can construct the multivariate analog of a probability integral transformation of the actual observations and test whether these transformations are uniformly distributed and serially uncorrelated. A formalization of this idea is provided by Diebold, Hahn, and Tay (1999).

From now on we will focus on log predictive scores (Good, 1952). To fix ideas, consider the following simple example. Let $x_t = [x_{1,t}, x_{2,t}]'$ be a 2×1 vector and consider the following two forecast models

$$M_1: \quad x_t \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right),$$
$$M_2: \quad x_t \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \right).$$

Under a quadratic loss function, the two models deliver identical univariate forecasts for each linear combination of the elements of x_t. Nonetheless, the predictive distributions are still distinguishable. Let Σ_i be the covariance matrix of the predictive distribution associated with the model M_i. The log predictive score is defined as the log predictive density evaluated at a sequence of realizations of x_t, $t = 1, \ldots, T$:

$$LPSC(M_i) = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln|\Sigma_i| - \frac{1}{2}\sum_{t=1}^{T} x_t' \Sigma_i^{-1} x_t.$$

Roughly speaking, if the actual x_t was deemed unlikely by M_i and falls in a low density region (e.g., the tails) of the predictive distribution, then the score is low. Let Σ_{11}, Σ_{12}, and Σ_{22} denote partitions of Σ that conform with the partitions of x_t. If we factorize the joint predictive density of x_t into a marginal and a conditional density, we can rewrite the predictive score as

$$LPSC(M_i) = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln\left|\Sigma_{i,11}\right| - \frac{1}{2\Sigma_{i,11}}\sum_{t=1}^{T} x_{1.t}^2 - \frac{T}{2}\ln\left|\Sigma_{i,22|11}\right| - \frac{1}{2\Sigma_{i,22|11}} \times \sum_{t=1}^{T}\left(x_{2,t} - \Sigma_{i,21}\Sigma_{i,11}^{-1}x_{1,t}\right)^2, \tag{187}$$

where

$$\Sigma_{i,22|11} = \Sigma_{i,22} - \Sigma_{i,21}\Sigma_{i,11}^{-1}\Sigma_{i,12}.$$

We can express the difference between log predictive scores of models M_1 and M_2 as

$$LPSC(M_1) - LPSC(M_2) = \frac{T}{2}\ln\left|1-\rho^2\right| - \frac{1}{2}\sum_{t=1}^{T} x_{2,t}^2 + \frac{1}{2(1-\rho^2)}\sum_{t=1}^{T}\left(x_{2,t} - \rho x_{1,t}\right)^2.$$

Here, the contribution of the marginal distribution of $x_{1,t}$ to the predictive scores cancels out, because it is the same for M_1 and M_2. It is straightforward to verify that the predictive score will be negative for large values of T if in fact the x_t s are generated from M_2. In fact, the log score differential has properties similar to those of a log likelihood ratio, and is widely used in the prequential theory discussed by Dawid (1992). Moreover, notice that $\frac{1}{T}\sum_{t=1}^{T}(x_{2,t} - \rho x_{1,t})^2$ can be interpreted as the mean squared error of a forecast of $x_{2,t}$ conditional on the realization of $x_{1,t}$. If $x_{1,t}$ and $x_{2,t}$ have a non-zero correlation, the conditioning improves the accuracy of the $x_{2,t}$ forecast. We will exploit this insight below.

Figure 5 depicts bivariate scatter plots generated from the joint predictive distribution of core and non-core variables. The predictive distribution captures both parameter uncertainty and shock uncertainty. We focus on one-step-ahead predictions for 2001:IV and 2006:III. We use filled circles to indicate the actual values (small, light blue), the unconditional mean predictions (medium, yellow), and the conditional means of ouput growth, PCE inflation, and unemployment, given the actual realization of the nominal interest rate. We approximate the predictive distributions using Student t distributions with mean μ, variance Σ, and ν degrees of freedom.[27] We replace μ and Σ with the sample means and covariance matrices computed from the draws from the predictive distributions. Regardless of the degrees of freedom ν, the conditional mean of x_2 given the realization of x_1 is given by:

$$\hat{x}_{2|1} = \mu_2 + \Sigma_{21}\Sigma_{11}^{-1}(x_1 - \mu_1). \tag{188}$$

In Figure 5 the nominal interest rate plays the role of the conditioning variable x_1.

First, consider the predictive distribution for output growth and interest rates in 2001:IV. The predictive distribution is centered at an interest rate of 4% and an output growth of about 0%. The actual interest rate turned out to be 2% and output grew at about 20 basis points over the quarter. Since the predictive distribution exhibits a negative correlation between interest rates and output growth, conditioning on the actual realization of the interest rate leads to an upward revision of the output growth forecast to

[27] Under this parameterization, the density of an m-variate Student t distribution is proportional to $[1 + (\nu - 2)(x - \mu)'\Sigma^{-1}(x - \mu)]^{-(\nu+m)/2}$.

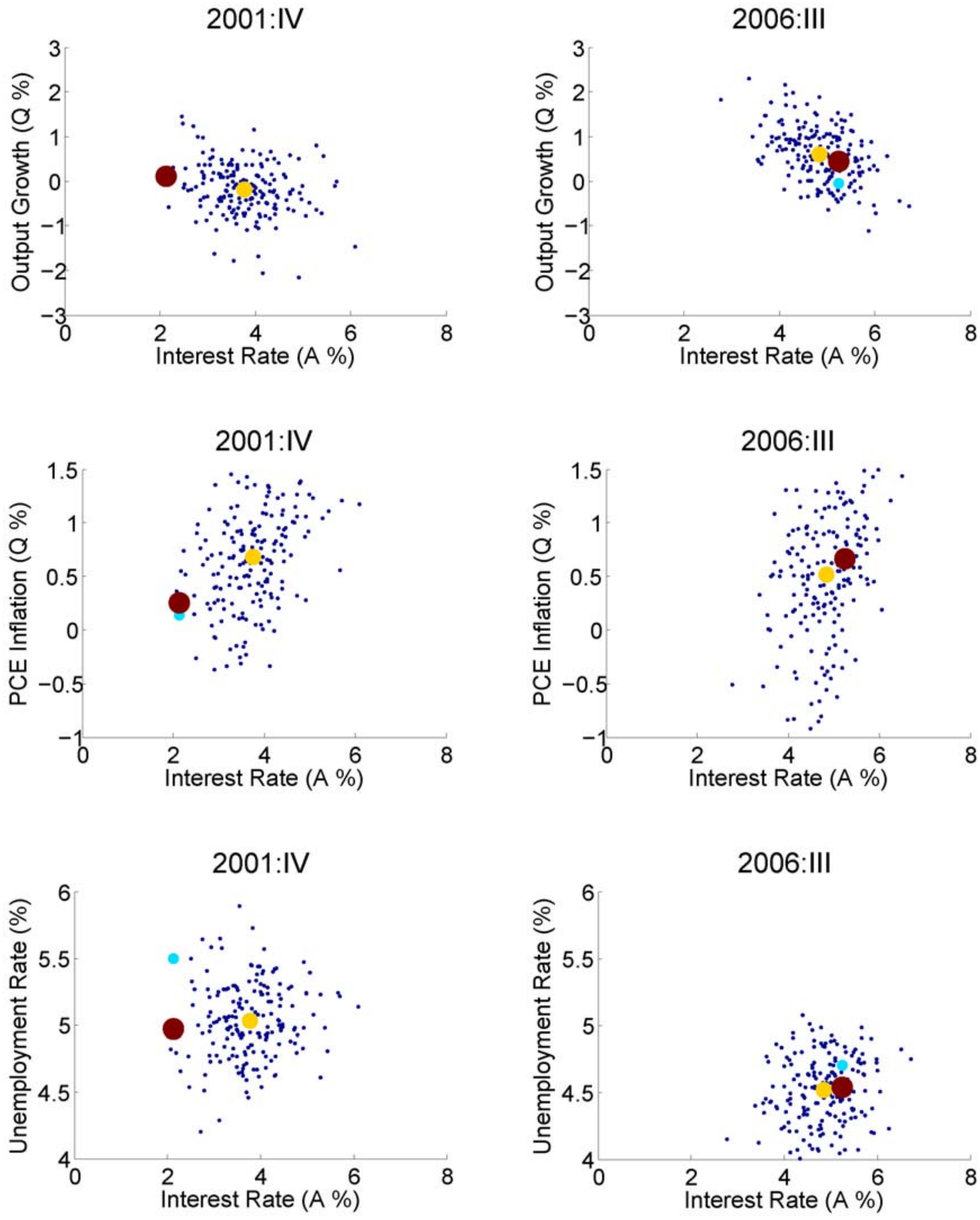
about 22 basis points. In 2006:III the actual interest rate exceeds the mean of the predictive distribution, and hence conditioning reduces the output growth forecast.

PCE inflation ($\lambda = 10^{-5}$) and the interest rate are strongly positively correlated, and the conditioning leads to a downward revision of the inflation forecast in 2001:IV and an upward revision in 2006:III. Our estimation procedure is set up in such a way as to leave the coefficients of the auxiliary regression uncorrelated with the DSGE model parameters. Hence, all of the correlation in the predictive distribution is generated by shock uncertainty and the fact that the auxiliary regression links the non-core variable to the DSGE model states.

Finally, we turn to the joint predictive distribution of unemployment ($\lambda = 10^{-5}$) and interest rates. Since, according to our estimates, the idiosyncratic shock ξ_t plays an important role in the unemployment dynamics and is assumed to be independent of the DSGE model shocks, the predictive distribution exhibits very little correlation. In this case, conditioning hardly affects the unemployment forecast.

Figure 5 focuses on two particular time periods. More generally, if the family of *t*-distributions provides a good approximation to the predictive distribution, and our model captures the comovements between interest rates and the other variables, then we should be able to reduce the RMSEs of the output, unemployment, and inflation forecasts by conditioning on the interest rate.

Figure 5. Bivariate One-Step-Ahead Predictive Distributions



Notes: The panels depict a scatter plot of draws from the one-step-ahead predictive distribution. The three filled circles denote the actual value (small, light blue), the unconditional mean predictor (medium, yellow), and the conditional mean mean predictor (large, brown). We set $\lambda = 10^{-5}$.

Table 9 and Table 10 provide RMSE ratios of conditional and unconditional forecasts. To put these numbers into perspective, we also report the ratio of the conditional versus the unconditional variance computed from a *t*-distribution with $\nu = 5$ degrees of freedom and a normal distribution ($\nu = \infty$). Using the subscript j to index the pseudo-out-of-sample forecasts, we define the average theoretical RMSE ratio as given below:

$$R(\nu) = \sqrt{\frac{\frac{1}{J}\sum_{j=1}^{J} \frac{\nu-2}{\nu}\left(1+\frac{1}{\nu-2}(x_{1,j}-\mu_{1,j})'\Sigma_{11,j}^{-1}(x_{1,j}-\mu_{1,j})\right)\left(\Sigma_{22,j}-\Sigma_{21,j}\Sigma_{11,j}^{-1}\Sigma_{12,j}\right)}{\frac{1}{J}\sum_{j=1}^{J}\Sigma_{22,j}}} \qquad (189)$$

The results obtained when conditioning on the interest rate, reported in Table 9, are somewhat disappointing. Although the bivariate correlations between the interest rate and the other variables are non-zero and would imply a potential RMSE reduction of up to 20% (except for housing starts), the RMSE obtained from the conditional forecasts exceeds that from the unconditional forecasts.[28] If we condition on the realization of the GDP deflator inflation (Table 10), then the results improve and we observe an RMSE reduction, at least for output growth and PCE inflation, although not as large as that predicted by $R(\nu)$.

[28] 2001:IV and 2006:III are not representative, since conditioning in these periods leads to a reduction of the forecast error.

Table 9. RMSE Ratios: Conditional (on Interest Rates) vs. Unconditional Forecasts

Series		$h = 1$	$h = 2$	$h = 4$	$h = 12$
Output Growth (Q%)	Actual	1.08	1.18	1.22	1.17
	(Theory)	(0.93, 0.96)	(0.92, 1.03)	(0.91, 1.09)	(0.92, 0.97)
100 x log Hours	Actual	1.23	1.42	1.57	2.05
	(Theory)	(0.96, 1.00)	(0.96, 1.06)	(0.95, 1.13)	(0.92, 0.95)
Inflation (Q%)	Actual	1.14	1.18	1.86	2.02
	(Theory)	(0.80, 0.82)	(0.83, 0.91)	(0.85, 0.98)	(0.82, 0.86)
PCE Inflation (Q%)	Actual	0.96	1.00	1.40	1.68
$\lambda = 10^{-5}$	(Theory)	(0.90, 0.91)	(0.90, 1.00)	(0.90, 1.05)	(0.90, 0.93)
Core PCE Inflation (Q%)	Actual	0.99	1.05	1.91	3.26
$\lambda = 10^{-5}$	(Theory)	(0.88, 0.88)	(0.89, 0.99)	(0.90, 1.05)	(0.90, 0.94)
Unemployment Rate (%)	Actual	1.16	1.43	1.60	1.45
$\lambda = 10^{-5}$	(Theory)	(0.98, 1.00)	(0.97, 1.08)	(0.96, 1.13)	(0.93, 0.96)
Housing Starts (mln/Q)	Actual	1.01	1.00	0.99	1.00
$\lambda = 10^{-5}$	(Theory)	(1.00, 1.06)	(1.00, 1.17)	(1.00, 1.20)	(1.00, 0.99)

Using the draws from the posterior predictive distribution of two variables x_1 and x_2, we construct conditional mean forecasts for x_2 given x_1, assuming that the predictive distribution is Student-*t* with $\nu = 5$ or $\nu = \infty$ degrees of freedom. We report RMSE ratios for conditional and unconditional recursive *h*-step-ahead pseudo-out-of-sample forecasts, with the theoretical reductions $R(\infty)$ and $R(5)$ in parentheses (see (189) for a definition).

Table 10. RMSE Ratios: Conditional (on GDP Deflator Inflation) vs. Unconditional Forecasts

Series		$h = 1$	$h = 2$	$h = 4$	$h = 12$
Output Growth (Q%)	Actual	0.94	0.91	0.94	1.04
	(Theory)	(0.94, 0.88)	(0.74, 0.70)	(0.75, 0.68)	(0.98, 0.90)
100 x log Hours	Actual	1.01	1.03	1.06	0.92
	(Theory)	(0.98, 0.92)	(0.74, 0.70)	(0.73, 0.65)	(0.98, 0.90)
PCE Inflation (Q%)	Actual	0.71	0.68	0.83	0.83
$\lambda = 10^{-5}$	(Theory)	(0.69, 0.65)	(0.67, 0.63)	(0.66, 0.60)	(0.67, 0.60)
Core PCE Inflation (Q%)	Actual	1.07	0.98	1.26	2.11
$\lambda = 10^{-5}$	(Theory)	(0.58, 0.54)	(0.62, 0.58)	(0.66, 0.59)	(0.68, 0.62)
Unemployment Rate (%)	Actual	1.06	1.08	1.09	1.10
$\lambda = 10^{-5}$	(Theory)	(0.99, 0.92)	(0.99, 0.93)	(0.99, 0.89)	(0.95, 0.86)
Housing Starts (mln/Q)	Actual	1.00	1.00	1.00	1.00
$\lambda = 10^{-5}$	(Theory)	(1.00, 0.93)	(1.00, 0.93)	(1.00, 0.90)	(1.00, 0.91)

See notes for Table 9.

These last results have to be interpreted carefully. It is important to keep in mind that we are examining particular dimensions of the joint predictive density generated by our model. While in the past researchers have reported log predictive scores and predictive likelihood ratios for DSGE model predictions, these summary statistics make it difficult to disentangle which dimensions the predictive distributions are well calibrated in. We decided to focus on bivariate distributions, in an attempt to assess whether the DSGE model and the auxiliary regressions capture the comovements of, say, interest rates with output growth, inflation, and unemployment. Our results were mixed: bivariate distributions that involved the interest rate were not well calibrated in view of the actual realizations, while bivariate distributions that involved the GDP deflator were somewhat more successful capturing the uncertainty about future pairwise realizations. An examination of the sequences of predictive densities and realizations – several of which are displayed in Figure 5 – suggested to us that the high RMSEs of the conditional forecasts were often caused by a small number of outliers, that is, actual observations that fall far in the tails of the predictive distribution. This suggests that more elaborate distributions for the structural DSGE model shocks might provide a remedy.

5 Conclusion

This chapter has developed a framework for generating DSGE model-based forecasts for economic variables which are not explicitly modelled but are of interest to the forecaster. Our framework can be viewed as a simplified version of the DSGE model-based factor model proposed by BG. We initially estimate the DSGE model on a set of core variables, extract the latent state variables, and estimate auxiliary regressions that relate non-modelled variables to the model-implied state variables. We compare the forecast performance of our model with those of a collection of AR(1) models based on pseudo-out-of-sample RMSEs. While our approach does not lead to a dramatic reduction in the forecast errors, by and large the forecasts are competitive with those of the statistical benchmark model. We also examined bivariate predictive distributions generated from our empirical model. Our framework inherits the two key advantages of DSGE model based forecasting: it delivers an interpretation of the predicted trajectories in light of modern macroeconomic theory and it enables the forecaster to conduct a coherent policy analysis.

Appendix. MCMC Implementation

DSGE model coefficients. The posterior sampler for the DSGE model is described by An and Schorfheide (2007).

Gibbs sampler for the coefficients that appear in measurement equations. We will in turn derive the conditional distributions for a Gibbs sampler that iterates over the conditional posteriors of α , ρ , and σ_η^2. We will start from the quasi-differenced form (Eq. (179)) of the auxiliary regression. τ , λ_0 , and λ_1 are treated as hyperparameters and considered as fixed in the description of the Gibbs sampler. Let L denote the lag operator.

Conditional posterior of α : The posterior density is of the form

$$p(\alpha \mid \rho, \sigma_\eta^2, Z^T, S^T) \propto p(Z^T \mid S^T, \alpha, \rho, \sigma_\eta^2) p(\alpha). \tag{190}$$

Define

$$\begin{aligned} y_1 &= \frac{\sigma_\eta}{\tau} z_1, & x_1' &= \frac{\sigma_\eta}{\tau}[1, \hat{s}_{1|1}'] & & \\ y_t &= (1-\rho L) z_t, & x_t' &= [1-\rho, (1-\rho L)\hat{s}_{t|t}'], & & t = 2, \ldots, T \end{aligned}$$

which implies that Eq. (179) can be expressed as a linear regression

$$y_t = x_t' \alpha + \eta_t. \tag{191}$$

If we let Y be a $T \times 1$ matrix with rows y_t and X be a $T \times k$ matrix with rows x_t' , then we can rewrite the regression in matrix form

$$Y = X\alpha + E.$$

We deduce

$$p(\alpha \mid \rho, \sigma_\eta^2, Z^T, S^T) \propto \exp\left\{-\frac{1}{2\sigma_\eta^2}(\alpha-\hat{\alpha})'\left(X'X\right)(\alpha-\hat{\alpha})\right\}\times \exp\left\{-\frac{1}{2}(\alpha-\mu_{\alpha,0})'V_{\alpha,0}^{-1}(\alpha-\mu_{\alpha,0})\right\}, \tag{192}$$

where

$$\hat{\alpha} = \left(X'X\right)^{-1} X'Y.$$

Thus, the conditional posterior of α is $N(\mu_{\alpha,T}, V_{\alpha,T})$ with

$$\mu_{\alpha,T} = V_{\alpha,T}\left[V_{\alpha,0}^{-1}\mu_{\alpha,0} + \frac{1}{\sigma_\eta^2}\left(X'X\right)\hat{\alpha}\right]$$
$$V_{\alpha,T} = \left(V_{\alpha,0}^{-1} + \frac{1}{\sigma_\eta^2}\left(X'X\right)\right)^{-1}.$$

Conditional posterior of ρ: Given the $U(-1;1)$ prior for ρ, the posterior density is of the form

$$p(\rho \mid \alpha, \sigma_\eta^2, Z^T, S^T) \propto p(Z^T \mid S^T, \alpha, \rho, \sigma_\eta^2) I\{|\rho| < 1\}. \tag{193}$$

We now define

$$y_t = z_t - \alpha_0 - \hat{s}'_{t|t}\alpha_1,$$
$$x_t = z_{t-1} - \alpha_0 - \hat{s}'_{t-1|t-1}\alpha_1.$$

Again, we can express Eq. (179) as a linear regression model

$$y_t = x_t\rho + \eta_t \tag{194}$$

Using the same arguments as before, we deduce that

$$p(\rho \mid \alpha, \sigma_\eta^2, Z^T, S^T) \propto \exp\left\{-\frac{1}{2\sigma_\eta^2}(\rho-\hat{\rho})'\left(X'X\right)(\rho-\hat{\rho})\right\}\times I\{|\rho|<1\} \tag{195}$$

with

$$\hat{\rho} = \left(X'X \right)^{-1} X'Y.$$

Thus, the conditional posterior is a truncated normal: $N(\mu_{\rho,T}, V_{\rho,T}) I\{|\rho| < 1\}$, with

$$\mu_{\rho,T} = \hat{\rho}, \qquad V_{\rho,T} = \sigma_\eta^2 \left(X'X \right)^{-1}.$$

Conditional posterior of σ_η^2: The posterior density is of the form

$$p(\sigma_\eta^2 \mid \alpha, \rho, Z^T, S^T) \propto p(Z^T \mid S^T, \alpha, \rho, \sigma_\eta^2) p(\sigma_\eta^2). \tag{196}$$

Solve Eq. (179) for η_t:

$$\eta_t = (1-\rho L) z_t - (1-\rho)\alpha_0 - (1-\rho L)\hat{s}'_{t|t}\alpha_1 \tag{197}$$

Now, notice that

$$p(\sigma_\eta^2 \mid \alpha, \rho, Z^T, S^T) \propto \left(\sigma_\eta^2 \right)^{-\frac{T+2}{2}} \exp\left\{ -\frac{1}{2\sigma_\eta^2} \sum_{t=1}^{T} \eta_t^2 \right\}. \tag{198}$$

This implies that the conditional posterior of σ_η^2 is inverted Gamma with T degrees of freedom and location parameter $s^2 = \sum_{t=1}^{T} \eta_t^2$. To sample a σ_η^2 from this distribution, generate T random draws $Z_1, \ldots, Z_T$ from a $N(0, 1/s^2)$ and let

$$\tilde{\sigma}_\eta^2 = \left[\sum_{j=1}^{T} Z_j^2 \right]^{-1}.$$

Marginal data density: Can be approximated using Chib's (1995) method. Let $\hat{\alpha}$, $\hat{\rho}$ and $\hat{\sigma}_\eta^2$ be the posterior mean estimates computed from the output of the Gibbs sampler. According to Bayes' Theorem,

$$p(Y) = \frac{p(Y \mid \hat{\alpha}, \hat{\rho}, \hat{\sigma}_\eta^2) p(\hat{\alpha}) p(\hat{\rho}) p(\hat{\sigma}_\eta^2)}{p(\hat{\alpha} \mid \hat{\rho}, \hat{\sigma}_\eta^2, Y) p(\hat{\rho} \mid \hat{\sigma}_\eta^2, Y) p(\hat{\sigma}_\eta^2 \mid Y)}. \tag{199}$$

All but the following two terms are straightforward to evaluate. First, let $\alpha_{(i)}$ and $\rho_{(i)}$ denote the i th draw from the Gibbs sampler. Then we can use the approximation:

$$\hat{p}(\hat{\sigma}_\eta^2 \mid Y) = \frac{1}{n} \sum_{i=1}^{n} p(\hat{\sigma}_\eta^2 \mid \alpha_{(i)}, \rho_{(i)}, Y). \tag{200}$$

Now consider a "reduced" run of the Gibbs sampler, in which we fix $\sigma_\eta^2 = \hat{\sigma}_\eta^2$ and iterate over $p(\alpha \mid \rho, \hat{\sigma}_\eta^2, Y)$ and $p(\rho \mid \alpha, \hat{\sigma}_\eta^2, Y)$ using the conditional densities in Eqs. (192) and (195). Denote the output of this Gibbs sampler by $\alpha_{(s)}$ and $\rho_{(s)}$. Then,

$$\hat{p}(\hat{\rho} \mid \hat{\sigma}_\eta^2, Y) = \frac{1}{m} \sum_{s=1}^{m} p(\hat{\rho} \mid \alpha_{(s)}, \hat{\sigma}_\eta^2, Y). \tag{201}$$

Generalization to AR(p): Let $\rho(L) = 1 - \sum_{j=1}^{p} \rho_p L^p$, where L is the lag operator; we can then express the auxiliary model as:

$$\begin{aligned} z_t &= \alpha_0 + \hat{s}'_{t|t}\alpha_1 + \xi_t, & t &= 1, \ldots, p \\ \rho(L) z_t &= \rho(1)\alpha_0 + \rho(L)\hat{s}'_{t|t}\alpha_1 + \eta_t, & t &= p+1, \ldots, T, \end{aligned}$$

where $[\xi_1, \ldots, \xi_p]' \sim N(0, \tau^2 \Omega(\rho(L)))$ and $\Omega(\rho(L))$ is the correlation matrix associated with the stationary AR(p) specification of ξ_t. The conditional posteriors of α and σ_η^2 are obtained from a straightforward generalization of Eqs. (192) and (198). The conditional posterior distribution of $\rho_1, \ldots, \rho_p$ is now non-normal and requires a Metropolis step. A generalization of Eq. (195) can serve as the proposal density. To conveniently enforce the stationarity of the autoregressive measurement error process, it could be re-parameterized in terms of partial autocorrelations as done by Barndorff-Nielsen and Schou (1973).

Bibliography

Adolfson, Malin, Laseén, Stefan, Linde, Jesper and Mattias Villani (2005), "Bayesian Estimation of an Open Economy DSGE Model with Incomplete Pass-Through," *Sveriges Riksbank Working Paper Series*, #179, March

Adolfson, Malin, Laseén, Stefan, Linde, Jesper and Mattias Villani (2007), "Forecasting Performance of an Open Economy Dynamic Stochastic General Equilibrium Model," *Econometric Reviews*, 26, pp. 289-328

Adolfson, Malin, Andersson, M., Linde, Jesper, Villani, Matias, and Vredin, A. (2007), "Modern Forecasting Models in Action: Improving Macroeconomic Analyses at Central Banks," *International Journal of Central Banking*, 3, pp. 111–144

Adolfson, Malin, Laseén, Stefan, Linde, Jesper and Mattias Villani (2008), "Evaluating an Estimated New Keynesian Small Open Economy Model," *Journal of Economic Dynamics and Control*, 32, pp. 2690-2721

Altissimo, F., Bassanetti, A., Cristadoro R., Forni, Mario, Hallin, Marc, Lippi, Marco, and Lucrezia Reichlin (2001), "EuroCOIN: A Real Time Coincident Indicator of the Euro Area Business Cycle," *Center for Economic Policy Research Discussion Paper 3108*

Altug, Sumru (1989), "Time-to-Build and Aggregate Fluctuations: Some New Evidence," *International Economic Review*, 30 (4), pp. 889-920

Amengual, D. and Mark Watson (2007), "Consistent Estimation of the Number of Dynamic Factors in a Large N and T Panel," *Journal of Business and Economic Statistics*, pp. 91-96

An, Sungbae and Frank Schorfheide (2007), "Bayesian Analysis of DSGE Models," *Econometric Reviews*, 26, pp. 113-172

Aruoba, S. Borağan, Diebold, Francis and Chiara Scotti (2009), "Real-Time Measurement of Business Conditions," *Journal of Business and Economic Statistics*, 24 (4), pp. 417-427 (October)

Aruoba, S. Borağan and Frank Schorfheide (2009), "Sticky Prices versus Monetary Frictions: An Estimation of Policy Trade-offs," *NBER Working Paper No 14870*

Bai, Jushan and Serena Ng (2002), "Determining the Number of Factors in Approximate Factor Models," *Econometrica*, 70 (1), pp. 191-221

Bai, Jushan and Serena Ng (2007), "Determining the Number of Primitive Shocks in Factor Models," *Journal of Business and Economic Statistics*, 25 (1), pp. 52-60

Barndorff-Nielsen, O. and G. Schou (1973), "On the Parameterization of Autoregressive Models by Partial Autocorrelations," *Journal of Multivariate Analysis*, 3, pp. 408–419

Bauwens, Luc, Lubrano, Michel, and Jean-François Richard (1999), *Bayesian Inference in Dynamic Econometric Models*, New York, NY: Oxford University Press

Bäurle, Gregor (2008), "Priors from DSGE Models for Dynamic Factor Analysis," *University of Bern Discussion Paper*, No 08-03, August

Bernanke, Ben S. (2004), "The Great Moderation," Federal Reserve Board, Remarks by Governor Ben S. Bernanke, February 20, 2004 (available online at http://www.federalreserve.gov/BOARDDOCS/SPEECHES/2004/20040220/default.htm)

Bernanke, Ben, Boivin, Jean and Piotr Eliasz (2005), "Measuring the Effects of Monetary Policy: A Factor Augmented Vector Autoregressive (FAVAR) Approach," *Quarterly Journal of Economics*, 120, pp. 387-422

Bernanke, Ben, Gertler, Mark and Simon Gilchrist (1999), "The Financial Accelerator in a Quantitative Business Cycle Framework," in J. Taylor, M. Woodford (eds.), *Handbook of Macroeconomics*, Vol. 1, Chapter 21, pp. 1341-1393

Blanchard, Olivier and Charles Kahn (1980), "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, 48 (5), pp. 1305-1311

Boivin, Jean and Marc Giannoni (2006), "DSGE Models in a Data-Rich Environment," *Manuscript*, HEC Montreal and Columbia University

Boivin, Jean and Serena Ng (2005), "Understanding and Comparing Factor Based Macroeconomic Forecasts," *International Journal of Central Banking*, 1, pp. 117-152

Boldrin, M., Christiano, L. and Jonas Fisher (2001), "Habit Persistence, Asset Returns, and the Business Cycle," *American Economic Review*, 91, pp. 149–166

Calvo, Guillermo (1983), "Staggered Prices in a Utility-Maximizing Framework," *Journal of Monetary Economics*, 12 (3), pp. 383-398

Carter, Christopher and Robert Kohn (1994), "On Gibbs Sampling for State Space Models," *Biometrika*, 81 (3), pp. 541-553

Chamberlain, Gary and M. Rothschild (1983), "Arbitrage, Factor Structure and Mean-Variance Analysis in Large Asset Markets," *Econometrica*, 70, pp. 191-221

Chang, Y. and S.-B. Kim (2006), "From Individual to Aggregate Labor Supply: A Quantitative Analysis Based on a Heterogeneous Agent Macroeconomy," *International Economic Review*, 47, pp. 1–27

Chib, Siddhartha (1995), "Marginal Likelihood from the Gibbs Output," *Journal of the American Statistical Association*, 90, pp. 1313-1321

Chib, Siddhartha (2001), "Markov Chain Monte Carlo Methods: Computation and Inference", in J.J. Heckman and E. Leamer (eds.), *Handbook of Econometrics*. Vol. 5, Elsevier Science BV, Chapter 57

Chib, Siddhartha and Edward Greenberg (1994), "Bayes Inference in Regression Models with ARMA (p, q) Errors," *Journal of Econometrics*, 64, pp. 183-206

Christiano, Lawrence, Eichenbaum, Martin and Charles Evans (2005), "Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy," *Journal of Political Economy*, 113 (1), pp. 1-45

Coenen, Günter, McAdam, Peter and Roland Straub (2008), "Tax Reform and Labour-Market Performance in the Euro Area: A Simulation-Based Analysis Using the New Area-Wide Model," *Journal of Economic Dynamics and Control*, 32, pp. 2543-2583

Cynamon, Barry, Dutkowsky, Donald and Barry Jones (2006), "Redefining the Monetary Aggregates: A Clean Sweep," *Eastern Economist Journal*, 32 (4), pp. 661-672

Dawid, A. P. (1992), "Prequential Analysis, Stochastic Complexity and Bayesian Inference," in J. M. Bernado, et al. (Eds.), *Bayesian Statistics*, Vol. 4 (pp. 109–125), Oxford University Press

Del Negro, Marco and Christopher Otrok (2008), "Dynamic Factor Models with Time-Varying Parameters. Measuring Changes in International Business Cycles," *Federal Reserve Bank of New York Staff Report*, #325

Del Negro, Marco and Frank Schorfheide (2004), "Priors from General Equilibrium Models for VARs," *International Economic Review*, 45 (2), pp. 643- 673

Del Negro, Marco and Frank Schorfheide (2008), "Forming Priors for DSGE Models (And How It Affects the Assessment of Nominal Rigidities)," *Journal of Monetary Economics*, 55, pp. 1191-1208

Del Negro, Marco, Schorfheide, Frank, Smets, Frank and Raf Wouters (2007), "On the Fit of New Keynesian Models," *Journal of Business and Economic Statistics*, 25, pp. 123-162

Diebold, Francis, Hahn, J., and A. Tay (1999), "Multivariate Density Forecast Evaluation and Calibration in Financial Risk Management: High Frequency Returns on Foreign Exchange," *Review of Economics and Statistics*, 81, pp. 661–673

Diebold, Francis and Roberto Mariano (1995), "Comparing Predictive Accuracy," *Journal of Business and Economic Statistics*, 13, pp. 253–263

Doan, Thomas, Litterman, Robert and Christopher Sims (1984), "Forecasting and Conditional Projections Using Realistic Prior Distributions," *Econometric Reviews*, 3 (4), pp. 1-100

Edge, Rochelle, Kiley, Michael and Jean-Philippe Laforte (2009), "A Comparison of Forecast Performance between Federal Reserve Staff Forecasts, Simple Reduced-Form Models, and a DSGE Model," *Finance and Economics Discussion Series 2009-2010*, Federal Reserve Board of Governors

Erceg, C., Henderson, D. and Andrew Levin (2000), "Optimal Monetary Policy with Staggered Wage and Price Contracts," *Journal of Monetary Economics*, 46 (2), pp. 281-313

Forni, Mario, Giannone, Domenico, Lippi, Marco and Lucrezia Reichlin (2007), "Opening the Black Box: Structural Factor Models with Large Cross-Sections," *Manuscript*, Center for Economic and Policy Research (CEPR)

Forni, Mario, Hallin, Marc, Lippi, Marco, and Lucrezia Reichlin (2003), "Do Financial Variables Help Forecasting Inflation and Real Activity in the Euro Area?" *Journal of Monetary Economics*, 50, pp. 1243-1255

Geweke, John (1977), "The Dynamic Factor Analysis of Economic Time Series," in D.J. Aigner and A.S. Goldberger (eds.), *Latent Variables in Socio-Economic Models*, Amsterdam, North-Holland Publishing, Ch. 19

Geweke, John (2005), *Contemporary Bayesian Econometrics and Statistics*, Hoboken, New Jersey: John Wiley and Sons, Inc

Geweke, John and Guofu Zhu (1996), "Measuring the Pricing Error of the Arbitrage Pricing Theory," *Review of Financial Studies*, 9 (2), pp. 557-587

Giannone, Domenico, Monti, Francesca, and Lucrezia Reichlin (2008), "Incorporating Conjunctural Analysis in Structural Macroeconometric Models," *Manuscript*, ECARES and London Business School

Giannone, Domenico, Reichlin, Lucrezia, and Luca Sala (2004), "Monetary Policy in Real Time," *NBER Macroeconomics Annual 2004*, pp. 161-200

Giannone, Domenico, Reichlin, Lucrezia, and D. Small (2008), "Nowcasting: The Real-Time Informational Content of Macroeconomic Data," *Journal of Monetary Economics*, 55, pp. 665-676

Good, I. J. (1952), "Rational Decisions," *Journal of the Royal Statistical Society*, Series B, 14, pp. 107–114

Greenwood, Jeremy, Hercowitz, Zvi, and Per Krusell (1998), "Long-Run Implications of Investment-Specific Technological Change," *American Economic Review*, 87 (3), pp. 342-362

Hallin, Marc and Roman Liška (2007), "Determining the Number of Factors in General Dynamic Factor Model," *Journal of the American Statistical Association*, 102, pp. 603-617

Harvey, Andrew (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge: Cambridge University Press

Harvey, D., Leybourne, S. and P. Newbold (1998), "Tests for Forecast Encompassing," *Journal of Business and Economic Statistics*, 16, pp. 254–259

Ireland, Peter (2004), "A Method for Taking Models to the Data," *Journal of Economic Dynamics and Control*, 28 (6), pp. 1205-1226

Jungbacker, Borus and Siem Jan Koopman (2008), "Likelihood-Based Analysis for Dynamic Factor Models," *Manuscript*, Department of Econometrics, VU University Amsterdam

Kim, Chang-Jin and Charles Nelson (1999a), "Has the U.S. Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle," *The Review of Economics and Statistics,* 81, pp. 608 – 616

Kim, Chang-Jin and Charles Nelson (1999b), *State-Space Models with Regime Switching*, Cambridge, MA: MIT Press

Kimball, M. and M. Shapiro (2008), "Labor Supply: Are the Income and Substitution Effects Both Large or Both Small?" *NBER Working Paper No 14208*

King, Robert and Mark Watson (2002), "System Reduction and Solution Algorithms for Solving Linear Difference Systems under Rational Expectations," *Computational Economics*, 20, pp. 57-86

Klein, Paul (2000), "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model," *Journal of Economic Dynamics and Control*, 24, pp. 1405-1423

Kose, M. Ayhan, Otrok, Christopher, and Charles Whiteman (2003), "International Business Cycles: World, Region and Country Specific Factors," *American Economic Review*, 93 (4), pp. 1216-1239

Kose, M. Ayhan, Otrok, Christopher, and Charles Whiteman (2008), "Understanding the Evolution of World Business Cycles," *Journal of International Economics*, 75, pp. 110-130

Leeper, Eric and Christopher Sims (1994), "Toward a Modern Macroeconomic Model Usable for Policy Analysis," *NBER Macroeconomics Annual 1994*, pp. 81-117

Levin, Andrew, Onatsky, Alexei, Williams, John and Noah Williams (2005), "Monetary Policy under Uncertainty in Micro-Founded Macroeconometric Models," *NBER Macroeconomic Annual 2005*, pp. 229-287

Lubik, Thomas and Frank Schorfheide (2004), "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic Review*, 94 (1), pp. 190-217

Lubik, Thomas and Frank Schorfheide (2005), "A Bayesian Look at New Open Economy Macroeconomics," *NBER Macroeconomics Annual 2005*, pp. 313–366

McCallum, Bennett T. (1983), "On Non-Uniqueness in Rational Expectations Models: An Attempt at Perspective," *Journal of Monetary Economics*, 11 (March), pp. 139-168

McCallum, Bennett T. (1999), "Role of the Minimal State Variable Criterion in Rational Expectations Models," *International Tax and Public Finance*, 6 (November), pp. 621-639

McCallum, Bennett T. (2003), "The Unique Minimum State Variable RE Solution is e-Stable in All Well Formulated Linear Models," *NBER Working Paper No 9960*, September

McConnell, Margaret and Gabriel Perez-Quiros (2000), "Output Fluctuations in the United States: What has Changed since the Early 1980's," *American Economic Review*, Vol. 90, No. 5, pp. 1464-1476

Moench, Emanuel, Ng, Serena and Simon Potter (2008), "Dynamic Hierarchical Factor Models," *Manuscript*, Federal Reserve Bank of New York and Columbia University

Monti, Francesca (2008), "Forecast with Judgement and Models," *National Bank of Belgium Working Paper*, #153, December

Sargent, Thomas (1989), "Two Models of Measurements and the Investment Accelerator," *Journal of Political Economy*, 97, pp. 251-287

Sargent, Thomas and Christopher Sims (1977), "Business Cycle Modeling Without Pretending to Have too Much A-priori Economic Theory," in C.Sims et al. (eds.), *New Methods in Business Cycle Research*, Federal Reserve Bank of Minneapolis, Minneapolis

Schorfheide, Frank, Sill, Keith and Maxym Kryshko (2010), "DSGE Model Based Forecasting of Non-Modeled Variables," *International Journal of Forecasting* (forthcoming, July)

Sidrauski, Miguel (1967), "Rational Choice and Patterns of Growth in a Monetary Economy," *American Economic Review*, 57 (2), pp. 534-544

Sims, Christopher (1992), "Interpreting the Macroeconomic Time Series Facts: The Effects of Monetary Policy," *European Economic Review*, 36 (5), pp. 975-1000

Sims, Christopher (2002), "Solving Rational Expectations Models," *Computational Economics*, 20, pp. 1-20

Smets, Frank and Raf Wouters (2003), "An Estimated Stochastic Dynamic General Equilibrium Model for the Euro Area," *Journal of the European Economic Association*, 1, pp. 1123-1175

Smets, Frank and Raf Wouters (2007), "Shocks and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97, pp. 586-606

Stock, James and Mark Watson (1989), "New Indices of Coincident and Leading Economic Indicators," *NBER Macroeconomics Annual 1989*, pp. 351-394

Stock, James and Mark Watson (1999), "Forecasting Inflation," *Journal of Monetary Economics*, 44, pp. 293-335

Stock, James and Mark Watson (2002a), "Macroeconomic Forecasting Using Diffusion Indexes," *Journal of Business and Economic Statistics*, 20, pp. 147-162

Stock, James and Mark Watson (2002b), "Forecasting Using Principal Components from a Large Number of Predictors," *Journal of the American Statistical Association*, 97, pp. 1167-1179

Stock, James and Mark Watson (2002c), "Has the Business Cycle Changed and Why?" *NBER Working Paper No 9127*, September

Stock, James and Mark Watson (2005), "Implications of Dynamic Factor Models for VAR Analysis," *NBER Working Paper No 11467*

Stock, James and Mark Watson (2008), "Forecasting in Dynamic Factor Models Subject to Structural Instability," in *The Methodology and Practice of Econometrics, A Festschrift in Honour of Professor David F. Hendry,* Jennifer Castle and Neil Shephard (eds), Oxford: Oxford University Press (http://www.princeton.edu/~mwatson/papers/hendryfestschrift_stockwatson_April28 2008.pdf)

Taylor, John (1993), "Discretion versus Policy Rules in Practice," *Carnegie-Rochester Conference Series on Public Policy*, 39, pp. 195-214

Tierney, Luke (1994), "Markov Chains for Exploring Posterior Distributions," *Annals of Statistics*, 22, pp. 1701-1762

Uhlig, Harald (1999), "A Toolkit for Analyzing Non-Linear Dynamic Stochastic Models Easily," in Ramon Marimon and Andrew Scott (eds), *Computational Methods for the Study of Dynamic Economies,* Oxford University Press, New York, pp. 30-61

Walsh, Carl (2003), *Monetary Theory and Policy*, 2^{nd} edition, Cambridge, MA: MIT Press

Woodford, Michael (2003), *Interest and Prices: Foundations of a Theory of Monetary Policy*, Princeton: Princeton University Press