Correlated Disturbances and U.S. Business Cycles*

Vasco Cúrdia — Federal Reserve Bank of New York

Ricardo Reis — Columbia University

February 2011

Abstract

The dynamic stochastic general equilibrium (DSGE) models that are used to study business cycles typically assume that exogenous disturbances are independent autoregressions of order one. This paper relaxes this tight and arbitrary restriction, by allowing for disturbances that have a rich contemporaneous and dynamic correlation structure. Our first contribution is a new Bayesian econometric method that uses conjugate conditionals and Gibbs sampling to make the estimation of DSGE models with correlated disturbances feasible. This provides a useful check for model misspecification in the search for models with structural disturbances. Our second contribution is a reexamination of U.S. business cycles. We find that allowing for correlated disturbances resolves some conflicts between estimates from DSGE models and those from vector autoregressions, and that treating government spending as exogenous in spite of its clear countercyclicality in the data is the main source of misspecification. According to our estimates, government spending and technology disturbances play a larger role in the business cycle than previously ascribed, while changes in markups are less important. (*JEL* E30, E10)

*We are especially grateful to Marco Del Negro, Juan Rubio-Ramirez and Frank Schorfheide for helping us significantly improve our method and to Craig Burnside, Lawrence Christiano, Martin Eichenbaum, Alejandro Justiniano, Jonathan Parker, Giorgio Primiceri, Chris Sims, Bent Sorensen, and Mike Woodford for useful comments and discussions. The views expressed in this paper are those of the authors and do not necessarily reflect positions of the Federal Reserve Bank of New York or the Federal Reserve System.
Contacts: vasco.curdia@ny.frb.org and rreis@columbia.edu.

1 Introduction

A typical macroeconomic model takes as given some exogenous disturbances, proposes a model for the behavior of economic agents, and makes predictions for some endogenous variables. Because the disturbances are exogenous to the theory, by definition they are unexplained and must be taken as given, so it would be desirable to impose on them as few arbitrary restrictions as possible. However, the common practice in dynamic stochastic general-equilibrium (DSGE) models is the opposite, with very strict assumptions on the processes driving disturbances. This paper develops new estimation techniques for models with a rich correlation structure for the disturbance vector and applies them to study U.S. business cycles.

Our first contribution is a new Bayesian econometric technique to estimate dynamic macroeconomic models with potentially rich processes for the disturbances. We show that the economic structure of the models implies that key conditional posterior distributions belong either exactly or approximately to the family of conjugate distributions with known analytical form. We propose a new *conjugate-conditionals algorithm* that exploits this knowledge to efficiently characterize the posterior. When applied to DSGE models that assume that disturbances are independent first-order autoregressions, AR(1)s, our method significantly speeds up estimation. Because the parameters associated with the disturbances are part of the conjugate conditional distributions, the efficiency gains are even larger with correlated disturbances. Our method makes feasible the estimation of DSGE models that were previously prohibitively numerically costly by breaking a curse of dimensionality that plagues existing algorithms.

Our second contribution is methodological. In the simultaneous-equation reduced-form macroeconomic model tradition, there has long been a careful treatment of disturbances. Researchers routinely allow for rich dynamic cross and auto-correlations across disturbances, sometimes estimated non-parametrically. This literature has convincingly established that arbitrary restrictions on the disturbances can severely bias the estimates of key parameters and impulse responses and lead researchers astray in attempts to endogenize incorrectlyidentified disturbances.[1] However, DSGE macroeconometric models routinely assume that disturbances are independent AR(1)s, which is arbitrary and potentially dangerous for inference. Moreover, from a Bayesian perspective, since researchers are typically uncertain about the source and nature of the disturbances, generalizing the disturbance process ensures that this uncertainty is reflected in the posterior distribution.

We envision three possible uses for correlated disturbances. First, allowing for more flexible specifications than the independent AR(1) should robustify inferences in DSGE models, in the same way that good practice adjusts standard errors in linear regressions to allow for unknown heteroskedasticity and autocorrelation in the disturbances (Stock, 2010). It is even more important to be careful with the disturbances in the non-linear DSGE models than in linear regressions, because correlations will lead to not just inefficient but also biased estimates.

Second, allowing for correlated disturbances lets the data speak freely on the dimensions along which the model is inadequate. Therefore, it provides a check or test for model misspecification, as previously argued by Del Negro and Schorfheide (2009). We suggest that, after estimating a model with independent disturbances, a researcher should check whether allowing for correlated disturbances significantly affects the inferences. Ideally, one would like to be confident that the model captures the relevant co-movements among macroeconomic variables endogenously without needing to rely on exogenous correlated disturbances.

Third, if there is a strong correlation between different elements of the disturbance vector, the pattern of the estimated correlations should provide useful information on where is the model failing to endogenously match the data. Checking for correlated disturbances can therefore suggest the path to improving the model in endogenize these correlations, towards the ultimate aim of a model with truly "structural" disturbances that are uncorrelated with each other.

To be clear, we are not proposing the use of models with correlated disturbances as an end in it self. Rather, we see our method as (i) providing a useful specification check, (ii)

[1] See Cochrane and Orcutt (1949), Zellner (1962), and Newey and West (1987) for the evolution on dealing with disturbances, and Fair (2004) for a recent careful application.

allowing a researcher to robustify inferences against the possibility that disturbances are correlated, and (iii) highlighting directions for improving the endogenous propagation of the model.

The third contribution of this paper is to study U.S. business cycles. Not only is this an important field to which DSGEs have been applied, but also the assumption of uncorrelated AR(1) disturbances is clearly incredible in medium-scale business-cycle models. Whenever economists have measured disturbances directly, whether to total factor productivity (Solow, 1957), to government spending (Rotemberg and Woodford, 1992), to labor supply (Parkin, 1988, Hall, 1997), or to investment productivity (Jorgenson, 1966, Greenwood, Hercowitz and Krusell, 1997), they have almost always found that these measures of disturbances are cross and dynamically correlated in ways that are inconsistent with independent AR(1)s. Two striking examples were provided by Evans (1992) and Chari, Kehoe and McGrattan (2007). Evans (1992) estimated vector autoregressions using military spending to measure government-spending disturbance and using Solow residuals to measure productivity disturbances, and found that government spending Granger-causes productivity. Chari, Kehoe and McGrattan (2007) estimated a first-order vector autoregression, VAR(1), for the disturbances of a business-cycle model and found that most cross-correlations are large and statistically significant.

After a brief literature review and discussion of some issues, the paper is organized as follow. Section 2 introduces a simple real business-cycle model and uses it to present the conjugate-conditionals estimation method. The estimates of the model in the U.S. data show that disturbances are correlated in a particular way: government spending tends to strongly increase after a fall in productivity. This explains four long-standing puzzles for full-information estimates of this model: why hours tend to fall after an increase in productivity, why changes in productivity have a delayed and persistent effect on output, why productivity accounts for a large part of the business cycle, and why the intertemporal elasticity of substitution is small.

Section 3 presents the estimation method more generally. We show that the conjugate conditionals arise in a broad class of equilibrium macroeconomic models, and discuss how to exploit the knowledge of this known slice of the posterior distribution in making inferences.

The model in section 2 is quite simple and its limitation at fitting the data are well-known. Section 4 therefore focuses on a richer business-cycle model, from Smets and Wouters (2007), that provides a reasonably good fit to the data. Allowing for correlated disturbances does not significantly improve the fit of the model, nor does it affect its main qualitative predictions on the impact of policy changes in the economy. However, with correlated disturbances, wage markups are now less important sources of business cycles, being replaced by productivity and government spending as key drivers. Moreover, the data suggest that endogenizing the changes in investment-specific productivity and in risk premia, perhaps through financial frictions, is a promising way to improve the empirical performance of the model.

Section 5 concludes with a brief review of the main results.

1.1 Literature review

The closest paper to this one is Ireland (2004). He adds measurement errors to the reduced-form equations of a DSGE model and allows them to follow a VAR(1), proceeding to estimate the model by maximum likelihood and to statistically test for structural stability. We differ in several respects. First, our focus is on the exogenous disturbances of the model, not on measurement error (which we will even abstract from). A key distinction between disturbances and measurement errors is that the properties of the disturbance process affect the behavioral responses of the agents in the model, whereas the properties of the measurement error only affect the job of the econometrician. For instance, if productivity disturbances are more persistent, agents in the model will engage in less intertemporal substitution in consumption and hours worked, altering the response of all endogenous variables. Instead, more persistent measurement errors only mechanically drive a difference between the endogenous variables and the observations. Second, from an econometric perspective, while both Ireland's and our approaches exploit the state-space representation of the model, Ireland's focus is on dealing with the measurement equation, while ours is on the state equation. Third, we take a Bayesian approach, we allow for VARs of higher order than one, and we focus on implications for business cycles.

Del Negro and Schorfheide (2009) also emphasize the need for robustifying inferences from DSGEs. Their preferred approach is to merge the versatility of a VAR with the tight restrictions of a DSGE in an innovative method that uses the DSGE to provide priors for the VAR. They also contrast their approach with the alternative of allowing for flexible processes for the disturbances as we do. As they note, our approach fits into their general framework for dealing with misspecification in policy analysis. Their empirical analysis is constrained to independent AR(2) processes though, and part of their criticisms focus on researchers judiciously picking which correlations to model. We instead allow for a more flexible and more general correlation structure for the disturbances. Finally, they emphasize the difficult issue of policy invariance, while we are more worried with the positive properties of the models.

A third related paper is Chib and Ramamurthy (2010). Like us, they use the insight of Gibbs sampling to propose an alternative to Metropolis estimation. However, while our blocks are suggested by the structure of the model, in their work it is the statistical properties of the data that guides the blocking of parameters. More concretely, at each step they randomly cluster the parameters into arbitrary blocks to reduce the number of draws that are necessary to characterize the posterior distribution. Instead, our algorithm clusters the parameters into two groups, the economic parameters and those related to the disturbance processes. While our clustering is not efficient in the sense of minimizing the number of draws, and there may be considerable correlation between the two groups, its virtue is that one of the groups has exact or approximate conjugate distributions. Therefore, while we may need more draws than Chib and Ramamurthy (2010), our algorithm is easier to implement and much faster to execute.

A few papers have moved beyond the assumption of independent AR(1) disturbances, but typically in only special ways. Within closed-economy models, Chari, Kehoe, and McGrattan (2007) allow for a restricted VAR(1) where the productivity disturbance is special in that it Granger-causes all others, and Smets and Wouters (2007) allow two of their seven disturbances to follow an ARMA(1,1) and two others to be contemporaneously correlated. Schmitt-Grohe and Uribe (2010) find that a common shock to total factor productivity and investment-specific productivity explain an important share of the business cycle.

In the open-economy literature, it is more common to assume that disturbances are correlated across countries, starting with the work of Backus, Kehoe and Kydland (1992). More recently, Justiniano and Preston (2010) estimate an open-economy DSGE model and find that correlated cross-country disturbances can partially account for the exchange rate disconnect puzzle. Rabanal, Rubio-Ramirez and Tuesta (2010) allow for cointegration among technological disturbances and find they can explain the volatility of real exchange rates.

As these papers on closed-economy and open-economy business cycles show, as models grow larger, with more disturbances and more emphasis on accounting for the data beyond just a few moments, there is a natural tendency to allow for correlated disturbances. We take a step further than this literature and allow for a richer and more general correlation between disturbances.

1.2 Three issues: simplicity, identification and orthogonalization

A natural objection to allowing for correlated disturbances is that it is harder to give them a structural interpretation than, say, AR(1) disturbances. While we are sympathetic with this objection, we are uncomfortable with its implications. Even though the estimates from independent AR(1)s for a vector of variables are simpler to interpret than those from a VAR, few (if any) researchers would argue in favor of the former instead of the latter. This apparent simplicity comes with great estimation biases and incorrect inferences. Moreover, as the two applications in this paper show, it is possible to interpret estimates with correlated disturbances. Once this is done, what becomes hard to understand is what was captured by estimates that assumed, for instance, that government spending was exogenous. Looking forward, we would expect that once researchers become used to models with correlated disturbances, this objection will become mute as it did just a few years after VARs became popular. In any case, the contribution of this paper is to argue that even when researchers prefer to assume independent disturbances, they should apply our specification check on whether their inferences are robust to allowing for correlated disturbances.

A second, more difficult, issue is identification. As noted by Sargent (1978) in estimating dynamic labor demands, it will often be difficult to empirically distinguish between endogenous sluggishness mechanisms, and exogenous persistent disturbances.[2] More generally, the issue is similar to the old argument (Griliches, 1967) that it is difficult to separately identify a linear regression with both a lagged dependent variable and an autocorrelated disturbance. Komunjer and Ng (2010) have provided a set of conditions for identification of DSGE models involving the rank of the information matrix, and which includes the case of correlated disturbances. In all of the applications of this paper, we exhaustively checked that their condition was satisfied, and never found a problem. Looking forward, we find compelling the argument that when there is an identification problem, the disturbance parameter responsible for it is set to zero so that the endogenous mechanisms have primacy in explaining the data.

Third, whenever disturbances are contemporaneously correlated, one must orthogonalize them to produce impulse responses and variance decompositions. In our empirical study of business cycles using the Smets and Wouters (2007) model, we consider the special case where disturbances are dynamically but not contemporaneously correlated, so that this issue does not arise. In the simple RBC model, different orthogonalizations give similar results so the issue is empirically negligible. Therefore, our applications are robust to different orthogonalizations.

More generally, we think it is a virtue rather than a vice to bring attention to the need for thinking hard about identification and orthogonalization in estimating DSGE models.[3] These are central issues in all empirical work, and should not be assumed away as the assumption of independent disturbances implicitly does. In any case, the particular methods and results in this paper do not depend on which stand one takes on identification and orthogonalization more generally.

---

[2] It is important to note that even if the exogenous disturbances could follow an arbitrary process, e.g. an infinite order VARMA, in many DSGE models, the economic and statistical parameters would still be identified. As noted by Sargent (1978) and many others, rational expectations models impose cross-equation restrictions that both identify the models as well as give them testable predictions.

[3] Reis (2008) discusses other identification issues in DSGE modelling.

2 Correlated disturbances in a canonical DSGE model

The best-known and simplest DSGE model is due to Prescott (1986), and we extend it to include government spending following Baxter and King (1993) and Christiano and Eichenbaum (1992). This model has three merits for our purposes. First, it is sufficiently simple that the effect of correlated disturbances can be grasped intuitively. Second, it has generated some puzzles that we can re-examine. And third, it only has a few parameters, which makes the estimation method transparent and allows us to conduct Monte Carlo experiments to gauge the efficiency of the method.

2.1 The model of fluctuations

A social planner chooses sequences of consumption and hours, $\{C_t, N_t\}_{t=0}^{\infty}$, to maximize

$$E_0 \left[\sum_{t=0}^{\infty} \beta^t \left\{ \frac{\left[C_t \left(1 - N_t\right)^{\theta} \right]^{1-1/\gamma} - 1}{1 - 1/\gamma} + V(G_t) \right\} \right], \tag{1}$$

subject to

$$Y_t = C_t + K_t - (1-\delta)K_{t-1} + G_t, \tag{2}$$

$$Y_t = \left(A_t N_t\right)^{1-\alpha} K_{t-1}^{\alpha}. \tag{3}$$

The notation is standard.[4] Utility increases with consumption and leisure and the benefits of government spending enter additively through the function $V(.)$, so they have no effect on the positive predictions of the model. Equation (2) states that output equals consumption plus investment plus government spending, and equation (3) is a neoclassical production function. We use this DSGE model to explain the business cycle in output and hours worked (Y_t, N_t) in response to disturbances to productivity and government spending (A_t, G_t).

[4] In particular: C_t is private consumption, G_t is government consumption, N_t is the fraction of hours in a quarter spent at work, K_t is capital, Y_t is output, A_t is total factor productivity, β is the discount factor, γ is the intertemporal elasticity of substitution, θ determines the relative utility from leisure and consumption, δ is the geometric depreciation rate, and α is the labor share.

Some of the parameters are easily pinned down by steady-state relations.[5] Two of the parameters are not, and they are crucial to the model's business-cycle predictions. First, the elasticity of intertemporal substitution, γ, determines the willingness of households to shift resources over time. It is a key determinant of how strongly savings and labor supply respond to persistent productivity changes, and thus of the model's ability to generate sizeable output fluctuations. Second, the parameter θ pins down the steady-state elasticity of labor supply with respect to wages. It is the key determinant of the size of the fluctuations in hours worked. We collect these *economic parameters* in the vector $\varepsilon = (\gamma, \theta)$.

Collecting the disturbances in the vector $s_t = (\ln(A_t), \ln(G_t/\bar{G})) \equiv (\hat{A}_t, \hat{G}_t)$, they follow a vector autoregression of order k:

$$s_t = \Phi(L)s_{t-1} + e_t \quad \text{with } e_t \sim N(0, \Omega), \tag{4}$$

where $\Phi(L) = \Phi_1 + .. + \Phi_k L^{k-1}$, the Φ_i are 2x2 matrices, and Ω is a positive-definite symmetric 2x2 matrix. This is a quite general representation; beyond assuming linearity and covariance stationarity, it merely assumes that the order k is large enough to approximate well an arbitrary Wold process. It nests three cases:

1) Independent AR(1) disturbances. This is the typical assumption in the literature, which in our notation maps into k being one and Φ_1 and Ω both being diagonal. These assumptions are hard to accept in this context. Government spending is certainly not an independent process in the data, and via the payment of unemployment benefits or countercyclical fiscal policy, G_t typically responds to A_t at least with a lag. In the other direction, perhaps private productivity responds with a lag to some forms of government spending like infrastructures or the enforcement of contracts.

2) Dynamically correlated disturbances. In this case, $k \geq 1$ and the Φ_i are unrestricted, but Ω is still diagonal. Because, in the model, G_t and A_t are exogenous, their correlations cannot be explained but must be assumed. It is then desirable to assume as little as possible

[5] In particular, the discount factor, β, is set at 0.995, to generate a steady-state risk-free annual real interest rate of 2%, the production parameter, α, is 0.33, to match the capital income share, the depreciation rate, δ, is 0.015 to roughly match econometric estimates and the average U.S. capital-ouput ratio, the average level of productivity, $\bar{A}$, is normalized to 1, and the average government spending $\bar{G}$ equals its historical average of 20% of GDP.

on these measures of our ignorance and focus instead on the tight restrictions imposed by the model on the endogenous variables. Imposing the assumption that Ω is diagonal has the virtue that we can still give a structural interpretation to the elements of e_t as innovations to productivity and government spending.

3) Contemporaneously correlated disturbances. Now Ω is not diagonal but it is left unrestricted. The elements of e_t no longer have a structural interpretation, unless we add orthogonalization assumptions as in the VAR literature. However, the inferences on the economic parameters ε are invariant to these restrictions.

We model disturbances in the RBC model as being both dynamically and contemporaneously correlated, so that we impose as little structure on the general specification in equation (4) as possible. With only two variables and two disturbances, only one orthogonalization condition is needed and it is easy to check alternatives and their implications for impulse responses and variance decompositions. Inspired by the results of Evans (1992) discussed in the introduction, in our baseline we use a Choleski decomposition with the innovations to government spending ordered first. We discuss robustness to this orthogonalization later.

One argument for assuming independent AR(1) disturbances is that it reduces the number of parameters. Letting σ denote the vector of *statistical parameters* in $\Phi(L)$ and Ω that describe the dynamics of the disturbances, with independent AR(1)s, σ has four elements. With unrestricted correlated disturbances, there are $3 + 4k$ statistical parameters. There is a curse of dimensionality as k increases, since the computational complexity of most estimation algorithms explodes even for modest values of k. However, as we show next, this is not a limitation of the theory, but rather of the particular algorithms being used.

2.2 Estimating the model

Log-linearizing the solution of the model around a non-stochastic steady state:

$$x_t = \lambda_1 \hat{K}_{t-1} + \lambda_2 s_t, \tag{5}$$

$$\hat{K}_t = \lambda_3 \hat{K}_{t-1} + \lambda_4 s_t, \tag{6}$$

where $x_t = (\hat{Y}_t, \hat{L}_t)$ are the observables, and a hat over a variable denotes its log-deviation. The state vector of the problem includes the exogenous s_t and the endogenous capital stock $\hat{K}_t$, and the λ_i are conformable matrices of coefficients that are functions of both ε and σ. These functions can be complicated, but are nowadays easily computed by many algorithms. Substituting out the unobserved capital stock, the reduced-form of the DSGE is:

$$x_t = \lambda_3 x_{t-1} + \lambda_2 s_t + (\lambda_1\lambda_4 - \lambda_3\lambda_2)s_{t-1}, \tag{7}$$

$$s_t = \Phi(L)s_{t-1} + e_t \quad \text{with } e_t \sim N(0, \Omega), \tag{8}$$

together with initial conditions s_0, x_0, and a transversality condition.

While the model has a state-space form, the non-linear function mapping the structural parameters to the reduced-form parameters, the cross-equation restrictions embedded in this map, and the absence of errors in the measurement equation, make this class of problems quite different from the conventional problem in state-space estimation (Durbin and Koopman, 2001). It is nowadays popular to take a Bayesian perspective to estimate models like this one.[6] Starting with a prior distribution for the parameters, $q(\varepsilon, \sigma)$, we can use the reduced-form in equation (7)-(8) to compute the likelihood function $\mathcal{L}(x^T | \varepsilon, \sigma)$ for a sample of data $x^T \equiv \{x_t\}_{t=1}^T$, and obtain the posterior distribution for the parameters via Bayes rule:

$$p(\varepsilon, \sigma | x^T) = \mathcal{L}(x^T | \varepsilon, \sigma) q(\varepsilon, \sigma) / p(x^T). \tag{9}$$

The marginal posterior density of the data $p(x^T)$ is unknown, and there is no convenient analytical form for the posterior distribution, so it must be characterized numerically. This is usually done with Markov Chain Monte Carlo (MCMC) algorithms, that draw a new (ε, σ) pair from an approximate distribution conditional on the last draw, in a way that ensures convergence of the draws to the posterior distribution.

The typical algorithm used is a random-walk Metropolis. At step j, it draws a proposal $(\varepsilon, \sigma)^{(j)}$ from a normal density with mean $(\varepsilon, \sigma)^{(j-1)}$ and some pre-defined covariance matrix, accepting this draw with a probability that depends on the ratio $p(\varepsilon, \sigma)^{(j)} / p(\varepsilon, \sigma)^{(j-1)}$, keeping $(\varepsilon, \sigma)^{(j-1)}$ in case of rejection. This algorithm is robust in the sense that it usually

[6] See Fernandez-Villaverde (2009) for a survey and a defense of the virtues of the Bayesian approach.

explores well the posterior distribution with minimal input from the researcher. The other side to this robustness is that, because it uses almost no knowledge of the shape of the posterior, the algorithm can take many draws to converge. Experience with DSGE models has found that it takes millions of draws to converge if there are more than ten parameters to estimate. With correlated disturbances, this algorithm quickly hits the curse of dimensionality and becomes infeasible.

We propose an alternative algorithm that avoids the curse of dimensionality by exploiting the economic structure of the model. Because its central observation is to use knowledge that some conditional posterior distributions are exactly or approximately conjugate, we label it the *conjugate-conditionals* algorithm. It is based on three observations.

First, by the principle of Gibbs sampling, we can break the sampling from the joint posterior at step j into drawing $\sigma^{(j)}$ from the conditional $p(\sigma^{(j)} \mid x^T, \varepsilon^{(j-1)})$ followed by drawing $\varepsilon^{(j)}$ from the conditional $p(\varepsilon^{(j)} \mid x^T, \sigma^{(j)})$. This well-known alternative to the random-walk Metropolis has here a natural application in separating statistical and economic parameters.

Second, note that while we are interested in the parameters, there is also uncertainty on the realization of the innovations e^T and thus the disturbances s^T. Focusing on the first Gibbs step, note that by the definition of a marginal distribution $p(\sigma \mid x^T, \varepsilon) = \int p(\sigma, s^T \mid x^T, \varepsilon) ds^T$, so drawing from the conditional for the statistical parameters is equivalent to drawing from the joint distribution for σ and s^T, retaining only the σ draws. This is often referred to in the statistics literature as data augmentation.

Third, note that drawing from $p(\sigma, s^T \mid x^T, \varepsilon)$ can be split by Gibbs sampling again into drawing from $p(s^T \mid x^T, \varepsilon, \sigma)$ and $p(\sigma \mid x^T, \varepsilon, s^T)$ in succession. But, conditional on the parameters, the reduced-form in equations (7)-(8) is a state-space system and the uncertainty on the disturbances s^T fits into a standard signal extraction problem. Therefore, the conditional distribution $p(s^T \mid x^T, \varepsilon, \sigma)$ is normal with mean and variance given by variants of the Kalman smoother. Moreover, conditional on the disturbances s^T, equation (8) is a standard vector autoregression. If the prior distribution Ω is an inverse-Wishart, then the posterior distribution is also an inverse-Wishart. Moreover, if the variability in the innovations e_t is much larger than the variability in the initial disturbances, then approximately all of the information about Φ in the system (7)-(8) is contained only in the second equation, and a normal prior for Φ leads to a normal posterior distribution. Both for the exact inverse-Wishart distribution and for the approximate normal distribution, we have easily computable analytical expressions for their moments.

Combining these three observations provides our algorithm. It draws from the expanded parameter vector $(\varepsilon, \sigma, s^T)$ in turn, exploiting the knowledge that the conditional distribution for s^T is known, while we have a good guess for the conditional distribution for σ. Only the conditional for ε is unknown, but this involves just two parameters, regardless of the assumptions on the disturbances. Allowing for correlated disturbances may dramatically increase the number of parameters in σ, but because the conditional posterior distribution for the covariance matrix is known analytically, and because we have a good approximating distribution for the conditional posterior distribution for the correlation coefficients, then the curse of dimensionality is broken. Estimating a DSGE with correlated disturbances is not significantly harder than one with independent AR(1) disturbances, because it is not harder to draw from normals and inverse-Wishart distributions of higher dimension. Because it uses our knowledge of particular slices of the posterior distribution that we are trying to characterize, this algorithm should be more efficient than the standard Metropolis algorithm.[7]

2.3 Data, priors, and Monte Carlo evidence on the efficiency of the algorithm

We now turn to the data to demonstrate the use of our method and its potential. Because the model is so simple, and such a large literature in the last twenty years has identified and partly remedied its weaknesses, we do not want to take the estimates too seriously. Our goal here is instead to show how some apparent puzzles when comparing likelihood-based estimates of this DSGE with other estimates can be resolved by allowing for correlated disturbances. Section 4 is more concerned with fitting the data. Here, we use U.S. data for non-farm business sector hours and output per capita that is quarterly, HP-filtered, and goes from 1948:1 to 2008:2, although we use the data before 1960:1 only to calibrate the

[7] The statement has to be qualified, because it is possible that the co-dependence between ε and σ is so strong that the Metropolis algorithm ends up dominating the Gibbs-sampler. In our experience, this is not the typical case.

priors.

The priors are summarized in table 1. Following the convention in the literature, we set the prior modes for the economic parameters at $\gamma = 2/3$ and $\theta = 4.85$ (to generate a steady-state value of 0.2 for N) and they have a gamma distribution. For the statistical parameters, the modes of the four AR(1) parameters (the diagonal terms of Φ_0 and Ω_0) are set to match four moments in the the data before 1960: the variances and serial correlations of output and hours. For the remainder statistical parameters, we consider two cases. In the first case, we follow the literature and assume independent AR(1)s. The priors for all the correlated-disturbance terms are zero with zero variance. We include this case both because it provides the comparison point for the correlated-disturbances case, and because it provides an illustration of the relative efficiency of our new algorithm. Our focus is on the correlated case, and we present results for an unrestricted VAR(1). The three non-diagonal elements in Φ_0 and Ω still have a prior mode of zero, but now have a non-zero variance set according to the extension of the Minnesota prior discussed in Kadiyala and Karlsson (1997), tighter around zero the further away we move from the diagonal.[8] We estimated VARs of orders 1 to 6 with very similar results. While the marginal likelihood is higher for order 6, we focus on the VAR(1) case because the results are easier to interpret and the difference in marginal likelihood is less than 3 log points.[9]

We start by investigating the efficiency of the conjugate-conditionals algorithm versus the Metropolis random-walk through a Monte Carlo experiment We simulated data of the same length as the sample using the priors for the independent AR(1), estimated the model on the simulated data using the two algorithms with four parallel chains, and then compared their relative efficiency at converging to the posterior distribution. The propositions in section 4 will show that the estimates from both algorithms converge to the truth, and our simulations confirmed this.

We use four metrics to assess convergence. First, the R statistic of Gelman and Rubin (1992), which compares the variance of each parameter estimate between and within chains,

[8]Section 3 discusses these priors in more detail as well as alternatives within the conjugate-conditionals family.

[9]Here and everywhere, we calculated the marginal likelihood using a harmonic mean with the truncated multivariate normal distribution as the weighting function, as described in Geweke (2005).

to estimate the factor by which these could be reduced by continuing to take draws. This statistic is always larger or equal than one, and a cut-off of 1.001 is often used. We report the maximum of these statistics across all the parameters. Second, the number of effective draws, *neff*, in each chain for each parameter, which corrects for the serial correlation across draws following Geweke (1992). The larger this is, the more efficient the algorithm, and we again report the minimum of these statistics. Third, the number of effective draws in total, *mneff*, which combines the previous two corrections applied to the mixed simulations from the four chains (Gelman et al, 1998: 298), where again we report the minimum across parameters. Finally, the number of rejections at the 5% level of the z-test that the mean of the parameter draws in two separated parts of the chain is the same. This is the separated partial means test, *SPM*, of Geweke (1992) and fewer rejections implies being closer to convergence.

Figure 1 shows the results.[10] In the horizontal axis are the number of draws, and in the vertical axes are the convergence metrics. The conjugate-conditionals algorithm clearly dominates the Metropolis random-walk. The number of effective draws is almost always higher, and commonly used thresholds like 1.01 for R, or 300 for *neff*, are reached earlier while the SPM tests have always fewer rejections for the same number of draws.[11]

Recall that, in these simulations, disturbances are uncorrelated. There are only four statistical parameters, so this Monte Carlo experiment provides a conservative estimate on the improvement to be had in switching to the conjugate-conditionals algorithm. When the disturbances are a VAR of high order, the benefits from the conjugate-conditional approach over a random-walk Metropolis are larger.

[10] The proposal density for ε in the conjugate-conditionals algorithm is a random-walk Metropolis. The covariance matrix for the Metropolis algorithm is the Hessian at the mode of the posterior (found by numerical maximization), multiplied by a scale factor to obtain approximately a 25% acceptance rate. This is updated after 20,000 draws to the covariance matrix of these draws, and the algorithm is then re-started. We report the draws in this second run, after discarding the initial 12,500 for burn-in, and average over 20 Monte Carlo simulations.

[11] Closely related to *neff* are efficiency factors (Geweke, 2005), which just equal *neff* divided by the number of draws. In the Monte Carlo, the efficiency factor of the conjugate conditionals algorithm is 0.0167, whereas that of a Metropolis random walk is only 0.0076.

2.4 Estimates and inferences with correlated disturbances

Starting with the independent AR(1)s case, the first panel of table 2 reports moments of the posterior distributions, and the top panel of figure 2 plots the distribution of impulse responses to one standard-deviation innovations to the two disturbances with the legend showing the median unconditional variance decomposition between parentheses. Four features of the estimates show well-known problems with this model.

1) The IES disconnect. The mean intertemporal elasticity of substitution is 1.4, not just above the prior, but especially substantially higher than the usual value of 0.2 that comes from Euler-equation estimates (Hall, 1988, Yogo, 2004).

2) The output persistence puzzle. In response to an improvement in productivity, output increases both because of the higher productivity, and also because the representative household chooses to work longer today when the returns to working are higher. However, as Cogley and Nason (1995) noted, the persistence of the output response closely mirrors the persistence of the productivity disturbance, whereas most reduced-form estimates of these responses are more gradual.

3) The hours-productivity puzzle. Gali (1999), Francis and Ramey (2005), and Basu, Fernald and Kimball (2006) estimated that hours fall after improvements in productivity, while Uhlig (2004) and Dedola and Neri (2007) find a response of hours close to zero. In figure 2 though, hours increase strongly after a productivity improvement.

4) The sources-of-business-cycles puzzle. According to the variance decompositions, government spending disturbances account for half of the variance of output and most of the variance of hours, against the findings in typical VAR studies (e.g., Shapiro and Watson, 1986, Fisher 2006), which attribute a larger role to productivity.[12]

One other feature of the estimates is worth discussing. An increase in public spending lowers resources inducing households to work harder, but because the shock is temporary they borrow from the future, de-accumulating capital. The first effect is stronger on impact so output rises, but as capital falls, within a few periods, the second effect becomes stronger and output turns negative. The empirical size of the fiscal multiplier is still under debate,

[12] The 90% credible sets for the variance decompositions output are (17, 79) and (21, 83) and for hours (3, 12) and (89, 97), for A_t and G_t respectively.

and the model predicts very different responses to transitory and permanent shocks (Baxter and King, 1993) so it is hard to compare these estimates to other evidence.

We now turn to the unrestricted VAR(1). The second panels of table 2 and figure 2 summarize the posterior distributions, impulse responses and variance decompositions. The three non-diagonal terms of the Φ and Ω matrices do not include zero in their 90% credible sets, unlike the assumption in the independent case. This is reflected in the log marginal predictive density of the model, which is 26 points higher with correlated disturbances than with independent AR(1)s, so the posterior odds ratio is an overwhelming e^{26} in favor of the former.[13] The largest correlated-disturbance term is the lagged productivity term in the law of motion for government spending. According to these estimates, when productivity falls, there is a lagged increase in government spending, matching what we would expect from the automatic and discretionary stabilizers in U.S. fiscal policy.

We can then re-examine the puzzles, now that we have allowed for this lagged response of government spending to productivity that the data strongly favors. First, the elasticity of intertemporal substitution is much lower, with a mean of 0.43 and a 5% bound of 0.29, bringing the DSGE estimates in line with the single-equation Euler equation estimates. Second, the response of output to a productivity disturbance is now significantly more delayed. An increase in productivity now leads to a subsequent fall in government spending. While this initially makes the impact on output smaller, after a few periods, it boosts output up partially solving the output persistence puzzle. Third, an improvement in productivity lowers hours. While the improvement in productivity increases hours, the subsequent fall in government spending lowers them and the net impact is close to zero, matching the results from the literature that followed Gali (1999). Fourth, productivity now accounts for a much larger fraction of the business cycle, and as much of the earlier predominance of government spending was due to its response to productivity.[14] In line with the VAR evidence, productivity now accounts for three quarters of the variance of output and 64%

[13] Note that because our prior was still centered around the independent-disturbances model, and had shrinking variances as we moved towards the cross-correlations, the marginal likelihood would be, if anything, biased against correlated distrubances.

[14] The 90% credible sets for the variance decompositions output are (58, 83) and (17, 42) and for hours (44, 75) and (25, 56), for A_t and G_t respectively.

of the variance of hours.[15]

Introducing correlated disturbances therefore solves four apparent puzzles with the real business cycle model. By imposing the strict and unjustified assumption that disturbances are independent AR(1)s, researchers would face a discrepancy between the DSGE full-information estimate and those that come from independent VARs and limited-information estimates. Allowing for correlated disturbances showed that the robust inference is instead that the dynamics of the model are broadly consistent with these other facts. Moreover, the estimates showed that the direction for improving the model is to account for countercyclical fiscal policy.

3 The general theory of the conjugate-conditionals method

Consider an economic model that relates the following vectors:

- x_t : observables, of dimension n_x;
- y_t : endogenous economic variables, of dimension n_y;
- s_t : exogenous disturbances, of dimension n_s;
- e_t : exogenous mean-zero innovations to the disturbances, of dimension n_s;
- ε : economic parameters, of dimension n_ε,
- σ : statistical parameters, of dimension n_σ,

in a sample $t = 1, ..., T$ with the convention that a variable dated t is determined at that date. The sample realization of a variable, say x_t, from $t = 1$ to date j is denoted by $x^j \equiv \{x_t\}_{t=1}^j$. We use $p(.)$ to denote a general posterior distribution, $f(.)$ to denote a sampling distribution, and $q(.)$ to denote a prior distribution. The inference problem is the following: given the observations x^T, and a prior distribution $q(\varepsilon, \sigma)$, to characterize the posterior distribution $p(\varepsilon, \sigma | x^T) \propto f\left(x^T \middle| \varepsilon, \sigma\right) q(\varepsilon, \sigma)$ numerically by simulating J draws.

[15] These results identify the impulse responses and variance decompositions with a Choleski decomposition ordering government spending first. We also tried ordering productivity first, as well as estimating a model with dynamic but not contemporaneous correlation between the disturbances. The solution of the four puzzles was robust to these alternatives. The results are also robust to the order of the VAR.

3.1 Two assumptions characterizing the problem

Two assumptions define our problem, and we will return to them at the end of this section to discuss whether they can be relaxed. The first assumption characterizes the economic models to which our methods apply, which includes most DSGE models in the literature:[16]

Assumption 1. *The economic model is:*

$$x_t = H_1(\varepsilon) + H_2(L)y_t + H_3(L)s_t, \tag{10}$$

$$y_t = \Lambda_1(\varepsilon,\sigma)y_{t-1} + \Lambda_2(\varepsilon,\sigma)s_t + \Lambda_3(\varepsilon,\sigma)(L)s_{t-1} \tag{11}$$

$$s_t = \Phi(\sigma)(L)\, s_{t-1} + e_t, \quad \text{with } e_t \text{ i.i.d. and } Var(e_t) = \Omega(\sigma). \tag{12}$$

$\Phi(\sigma)(L) = \sum_{j=1}^{k} \Phi_j(\sigma) L^{j-1}$, is a matrix lag polynomial of order k, and similarly for $H_2(L)$, $H_3(L)$ and $\Lambda_3(\varepsilon,\sigma)(L)$. All matrices are conformable and their elements are functions of the sub-set of parameters of (ε,σ) that are indicated in brackets. Moreover:

a) The distributions $p(\varepsilon|\, x^T,\sigma)$, $p(\sigma|\, x^T,\varepsilon)$ and $p(s^T\big|\, x^T,\varepsilon,\sigma)$ are not point masses, that is they are not degenerate in the sense of the random variables being almost surely constant.

b) The $\Lambda_j(\varepsilon,\sigma)$ matrices depend on the parameters in $\Phi(\sigma)$ but not on the parameters in $\Omega(\sigma)$.

c) The statistical parameters $\sigma = vec(\Phi,\Omega)$, and the matrices Φ and Ω are unrestricted.

Equation (10) links observables to endogenous variables and disturbances in a general linear way, including allowing for constant deviations between the two through $H_1(\varepsilon)$.[17] In applications with de-meaned data, often $x_t = y_t$ so the endogenous variables are observed, as was the case in section 2. More usually, the data is not de-meaned and $H_1(\varepsilon)$ includes the steady-state of the model, which depends on the economic parameters, while $H_2(L)$ and $H_3(L)$ are typically simple data transformations that adjust units.

We abstract from measurement error in these observations to avoid confusion with the economic disturbances specified in the model. Including measurement error does not change

[16] Beyond the two models in this paper, a supplementary appendix available form the authors shows that other popular DSGE models in the literature satisfy this assumption.

[17] We will treat y_t as deviations from a steady-state, so we omit constants from (11)-(12), but it is straightforward to include these.

our conclusions significantly, although it requires a clear distinction between them and the disturbances.

Assumption 1a) simply states that there is a statistical estimation problem. At the most elementary level, this requires that, given the observed data, there is more than one set of parameters that could have generated it with non-zero probability. The assumption strengthens this basic requirement in two ways. First, it requires that conditioning on each of the two sub-sets of parameters, ε or σ, again we still have a non-trivial statistical estimation problem. If this was not the case, some of the steps would be trivial or redundant. Second, it requires that the observables are not enough to recover the disturbances. Otherwise the statistical problem would boil down to estimating the VAR in (12).

The second equation in the assumption, equation (11), nests most linear (or linearized) dynamic economic models that are described by a system of equations:

$$\Psi_0(\varepsilon)y_t = \Psi_1(\varepsilon)y_{t-1} + \Psi_2(\varepsilon)(L)s_t + \Psi_3(\varepsilon)w_t, \tag{13}$$

where the vector of endogenous disturbances w_t has the property that $E_{t-1}w_t = 0$ and can capture terms involving $E_t(y_{t+1})$. The Ψ_i matrices typically have many zero elements and have more elements than n_ε, embodying the cross-equation restrictions that come from optimal behavior, technologies and other constraints and which are affected by the economic parameters ε. As Blanchard and Khan (1982) and Sims (2002) among many others have shown, equation (11) is the solution, or reduced-form, of these models.

The matrices $\Lambda_j(\varepsilon, \sigma)$ in this solution are typically complicated non-linear functions of all parameters. Therefore, while the model has a state-space representation, estimating it requires moving well beyond the standard techniques in state-space estimation (Durbin and Koopman, 2001).[18] In general, little can be said about the functions $\Lambda_j(\varepsilon, \sigma)$, as their form will depend on the model, but there is one exception stated as assumption 1b): the principle of certainty equivalence, that the parameters in the reduced-form solution of the model do not depend on the covariances in Ω.

[18] There is another difference relative to state-space models that one should not get confused about. In our model and notation, y_t are not the state variables. Rather, y_t includes *all* of the variables in the economic model, including states, controls, or any other variable.

The third equation requires that disturbances are linear processes that are well approximated by a vector autoregression of finite order k. It implies that $n_\sigma = kn_s + n_s(n_s+1)/2$, so the number of statistical parameters may be quite large.

Assumption 1c) clarifies that we are not imposing any restriction on the VAR. In many cases, we may want to impose some restriction, such as stationarity. Section 3.4 will discuss how this assumption can be relaxed.

The second assumption defines the distributions for the different propositions and variables:

Assumption 2. *The prior and likelihood distributions are:*
a) $f(e_t|\, \varepsilon, \sigma)$ *is a normal distribution*
b) $f(s^k|\, \varepsilon, \sigma)$ *is a normal distribution.*
c) $q(\Omega)$ *is an inverse-Wishart distribution,*
d) $q(\Phi|\, \Omega)$ *is a normal distribution,*
e) $q(\varepsilon)$ *is a non-degenerate distribution, that is* ε *is not almost surely constant.*

The first part of the assumption is standard in the literature: innovations are independent and identically normally distributed. The second part assumes that the initial unobserved states in the k lags of the VAR are also normal, so that the observations x^T are normally distributed. The third and fourth part set the priors for the statistical parameters. These are standard in the VAR literature, although not as common in the DSGE literature. Finally, the fifth part puts only the weakest restriction on the prior for the economic parameters for our method to work.

3.2 Two results on which the method rests

The first result breaks the problem into several sub-problems using the powerful result on Gibbs-sampling:

Proposition 1 *Starting at step* j *with* $(\varepsilon^{(j-1)}, \sigma^{(j-1)})$*, then:*
a) drawing $\sigma^{(j)}$ *from the conditional* $p(\sigma^{(j)}|\, x^T, \varepsilon^{(j-1)})$ *and then drawing* $\varepsilon^{(j)}$ *from the conditional* $p(\varepsilon^{(j)}|\, x^T, \sigma^{(j)})$ *converges in distribution to a set of draws from* $p(\varepsilon, \sigma|\, x^T)$*.*
b) drawing $\sigma^{(j)}$ *and* $s^{T(j)}$ *from the joint distribution* $p(\sigma^{(j)}, s^{T(j)}|\, x^T, \varepsilon^{(j-1)})$*, and storing*

only the $\sigma^{(j)}$ draws gives a set of draws from $p(\sigma^{(j)} | x^T, \varepsilon^{(j-1)})$
c) drawing $s^{T(j)}$ from the conditional $p(s^{T(j)} | x^T, \varepsilon^{(j-1)}, \sigma^{(j-1)})$ and then drawing $\sigma^{(j)}$ from the conditional $p(\sigma^{(j)} | x^T, \varepsilon^{(j-1)}, s^{T(j)})$ converges in distribution to a set of draws from $p(\sigma^{(j)}, s^{T(j)} | x^T, \varepsilon^{(j-1)})$.

Proof: Result b) just states the definition of a marginal distribution in relation to the joint distribution. Results a) and c) are applications of the convergence of the Gibbs sampler. The proof follows the same steps as Tierney (1994), where the crucial assumptions are 1a and 2e ensuring that the Markov chain defined by the Gibbs sampler is irreducible.⧫

Focusing on result c) of the previous proposition, we can further show that there are two conjugate families of priors-posteriors within our problem:

Proposition 2 *The following two distributions belong to known families, with analytical means and variances:*
a) the posterior distribution for the disturbances, conditional on the data and the parameters, $p(s^T | x^T, \varepsilon, \sigma)$ is normal.
b) the posterior distribution for the variance of the innovations, conditional on the data, the other parameters, and the disturbances $p(\Omega | x^T, \varepsilon, \Phi, s^T) = p(\Omega | s^T)$, and it is an inverse-Wishart.

Proof: Equations (10)-(12) in assumption 1 define a linear state-space system. Assumptions 2a and 2b state that innovations and initial conditions are normal. Therefore, the disturbances are normal, proving result a). To prove the second result, note that assumption 1b implies that only equation (12) involves the covariance matrix Ω. Moreover, no x^T or ε appear in that equation. It therefore follows that $p(\Omega | x^T, \varepsilon, s^T, \Phi) = p(\Omega | s^T)$. But then, using assumption 1c) and 1d), it is a standard result from linear regression that, since the prior is an inverse-Wishart, so is the posterior.⧫

As a final remark on this second proposition, note that the parameters of these distributions are known analytically. The mean and covariance of the normal distribution of result a) are the output of the Kalman smoother, and the parameter of the Wishart in result b) can be found in most Bayesian statistics textbooks (e.g., Geweke, 2005).

3.3 The conjugate-conditionals method

Based on the two propositions, the output of the following hybrid, Metropolis-within-Gibbs (or block-Metropolis) algorithm will converge to a set of draws from the posterior distribution of the parameters:

Algorithm *At draw j:*

Step 1) draw $s^{T(j)}$ from $p(s^{T(j)}|\, x^T, \varepsilon^{(j-1)}, \sigma^{(j-1)})$, the known distribution in proposition 2;

Step 2) draw $\Omega^{(j)}$ from $p(\Omega|\, s^{T(j)})$, the known distribution in proposition 2;

Step 3) draw $\Phi^{(j)}$ from a proposal distribution that approximates $p(\Phi|\, x^T, \varepsilon^{(j-1)}, s^{T(j)}, \Omega^{(j)})$ and accept or reject this draw with some probability;

Step 4) draw $\varepsilon^{(j)}$ from a proposal distribution that approximates $p(\varepsilon|\, x^T, \sigma^{(j)})$ and accept or reject this draw with some probability.

The first two steps are easy even for a very large number of disturbances n_s, number of lags, k, and number of observations T. Most software programs can take draws from the multivariate normal quickly and, while the Kalman filter recursions can take some time, they were required anyway in order to calculate the likelihood function of the problem. The Kalman smoother provides the posterior means and variances recursively, although as shown by Carter and Kohn (1994), sampling from the joint distribution is considerably more efficient. We use their approach as described in Chib's (2001) algorithm 14.

As for the third step, while we do not have the exact distribution, we have a good guess. The autocorrelation parameters Φ enter both the reduced-form solution of the model in (11), as well as the VAR in (12). But, if the variance of the innovations e^T is much smaller than the variance of the prior for the initial states and endogenous variables, then this filtering problem has an approximate solution where only the information in the VAR is relevant.[19] That is, in this limit case, $p(\Phi|\, x^T, \varepsilon, s^T, \Omega) \approx p(\Phi|\, s^T, \Omega)$. But then, we have another conjugate conditional, since assumption 2c) implies that the posterior for Φ is also normal and the formulae for the mean and variance are the standard linear regression formulae.

We have found that a particular implementation of this approximate proposal works

[19] It is common practice to set the prior variance of the initial conditions equal to the unconditional variance predicted by the system. If the economic system has significant propagation and magnification, then this variance should be considerably larger than the variance of the innovations.

remarkably well, converging quickly. Following Geweke (1989), we use an independence-Metropolis step sampling from a t-distribution instead of the normal in the previous paragraph, to allow for fatter tails. To be clear, recall that this is the proposal density, so there are many other alternatives that would lead to consistent estimates. What the argument in the previous paragraph strongly suggests, and our experience confirms, is that a t-distribution for $p(\Phi|\, s^T, \Omega)$, using the mean and variance from normal conjugate formulae, provides a good approximation to the target distribution $p(\Phi|\, x^T, \varepsilon, s^T, \Omega)$, as judged by how quickly the draws converge and the very high acceptance rate that we obtain for the draws.[20]

Finally, for step 4, our algorithm does not make any significant improvement over the literature. We neither know $p(\varepsilon|\, x^T, \sigma)$, nor is there any hope of having even an approximate result beyond very specific models, since the parameters ε usually enter the system in a highly non-linear way. In practice, we used a random-walk Metropolis for this step, drawing $\varepsilon^{(j)}$ from a normal with mean $\varepsilon^{(j-1)}$ and covariance matrix equal to the inverse-Hessian at the mode of the posterior, scaled to reach an acceptance rate around one quarter. We have tried several alternatives: independent Metropolis, rejection sampling, and modifying the random-walk Metropolis to have the new draws depend on $\sigma^{(j)}$, but none clearly dominated the more conventional random-walk Metropolis.

3.4 Relaxing the two assumptions

Some features of assumption 1 are central to our method. First, the state-space linear form of equations (11)-(12) are important, especially for proposition 2. Second, the fact that the matrices $\Lambda_i(\varepsilon, \sigma)$ with the model solution are non-linear functions of the parameters, with many cross-equation restrictions imposed by the theory, distinguishes our problem from an unrestricted state-space system. Both of these features define what a linear DSGE is, so they restrict us only insofar as we are dealing with this class of models.

Third, assumption 1a) ensures that the inference problem is not trivial. Fourth, assumption 1b) is important for the conjugate distribution of Ω in proposition 2, but this

[20] Some readers, accostumed to the Metropolis-Hastings random walk sampling often used in economics, may find it puzzling that we refer to high acceptance rates as a measure of efficiency. This is the case because we are using *independence* Metropolis.

principle of certainty equivalence applies to all linearized DSGE models. Fifth and finally, the assumption that the economic parameters ε do not affect the law of motion for the disturbances in (12) is crucial for the ability to deal separately with the two types of parameters, but it is as much an assumption as it the modelling definition of what σ and ε are.

The other assumptions can all be relaxed in many ways. Starting with equation (10), we could allow for $H_2(\varepsilon)(L)$ and $H_3(\varepsilon)(L)$, so the measurement equation linking the endogenous variables to the observable can depend on the economic parameters. This requires no change in the algorithm, although we have trouble finding economic models to which this extension would be useful.

Assumption 1c) can be significantly relaxed, as it is easy to accommodate many types of restrictions on the VAR matrices. One case is when disturbances follow independent AR(k)s, so Φ_j and Ω matrices are all diagonal. Adapting the priors in assumption 2 to $i = 1, ..., n_s$ independent normals for $[\Phi_j(i)]_{j=1}^k$, and $i = 1, ...n_s$ independent inverse-gammas squared for each of $\Omega(i)$, our results follow. A second case is to have dynamic but not contemporaneous correlation, so the Φ_j are unrestricted but the Ω must be diagonal. In this case, using the normal priors for Φ from assumption 2, and the independent inverse gamma priors for $\Omega(i)$ just described, again our results follow. Finally, more generally, we may wish to impose that some of the elements of Φ_j and Ω are either zero, or appear more than once in the matrices. In this case, equation (12) is a system of seemingly unrelated regressions (SUR). Collecting the disturbances into the vector $\bar{s}$ of size $n_s(T-k)$, it is written as $\bar{s} = Z\beta + \varepsilon$, with $\varepsilon \sim N(0, \Omega \otimes I_{t-k})$, where Z contains the lagged states as well as blocks of zeros allowing for a rich set of restrictions on the VAR. The coefficients β include the elements of Φ. As long as the prior for $\beta | \Omega$ is normal and the prior for Ω^{-1} is the Wishart distribution as described in assumption 2, then our results on conjugate distributions still hold (Zellner, 1962).

Another important class of prior restrictions comes from the common desire to impose the constraint that the VAR in equation (12) is stationary. This affects the distribution for Φ in step 3, which is now truncated to the stationarity region. However, our experience is that still using as proposal the t-distribution based on the approximate-normal result, but truncating it to only accept stationary draws, has almost no effect on the performance of the algorithm. This is perhaps not entirely surprising; the truncation does not affect the relative density of different draws in the stationary region, so it has little effect on the importance sampling algorithm.

Turning to assumption 2, the normality of the errors and initial conditions is important to our method in order to obtain the conjugate conditional distribution for the disturbances in proposition 2. Assumption 2c) and 2d) can be somewhat relaxed. There are alternative conjugate priors to the normal-inverse-Wishart family. Kadiyala and Karlsson (1997) discuss combinations of diffuse, normal, Wishart and Minnesota prior distributions that deliver conjugate families for VARs. Sims and Zha (1998) propose an alternative, with a normal conjugate family for the distribution of Φ conditional on Ω, which puts fewer restrictions on the prior variance than the one in assumption 2 and has some computational advantages, although the posterior for the covariance matrix Ω stops being conjugate.

4 New Keynesian cycles with correlated disturbances

The medium-scale DSGE models estimated by Smets and Wouters (2003, 2007) has been very influential in the study of business cycles and monetary policy. In this section, we ask to what extent their inferences are robust to the possibility of correlated disturbances.

4.1 The model

Smets and Wouters (2007) found that a new Keynesian model augmented with a variety of frictions, including sticky prices and wages, habits in consumption, and investment adjustments costs, can fit the U.S. data well. They focused on seven series: output, consumption, investment, hours worked, real wages, inflation and nominal interest rate.[21]

The appendix lays out the equations of the model. Our focus is on its seven exogenous disturbances: total factor productivity (A), investment-specific productivity (EI), risk premium (B), government spending (G), monetary policy (ER), price markups (EP), and wage markups (EW). Following the DSGE tradition, Smets and Wouters assume that they

[21] Del Negro et al (2007) document more exhaustively the empirical strengths and weaknesses of the model.

all follow independent AR(1)s, with only two exceptions. First, the model includes two first-order moving average terms for the price and wage markup disturbances to fit high-frequency movements in the data. Second, Smets and Wouters allow for contemporaneously correlated disturbances between government spending and total factor productivity.

We re-estimate the Smets and Wouters model, using the same data and priors, but with two different assumptions for the dynamics of the disturbances.[22] In the first case, we assume that disturbances are independent AR(1)s, just like Smets and Wouters, with only one modification: we set to zero the contemporaneous correlation between productivity and government spending. This way, all of the disturbances are independent. In the second case, we allow for correlated disturbances that follow a VAR(1). We impose the restriction on this VAR that disturbances are dynamically correlated, but not contemporaneously so. That is, using the notation from section 3, the matrix of dynamic-correlation coefficients Φ is left unrestricted, but the variance matrix Ω is required to be diagonal. We restrict ourselves to this case because, with seven disturbances, any orthogonalization would be controversial.[23]

4.2 The economic parameters and impulse responses

The full set of estimates and impulse responses is reported in the appendix. While there are a few differences between the estimates of the economic parameters across the two specifications of the disturbances, the posterior credible sets typically overlap. Since, in turn, these sets are wide, it is hard to draw any firm conclusions. One exception is in the estimates of the parameter that measures the elasticity of adjustment costs to investment, which are well above the prior with independent disturbances, but slightly below it with correlated disturbances.

Focusing instead on the impulse responses, figure 3 shows the impulse responses of output, hours and inflation to disturbances to productivity, fiscal spending, and monetary policy. They are all qualitatively similar. An improvement in productivity still lowers

[22] All the estimates are based on 3 million draws, preceeded by another 6 million draws used to update the covariance matrix in the Metropolis proposals.

[23] It is important to emphasize that, with our algorithm, allowing the disturbances to follow vecor autoregressions of much higher order would be feasible and fast.

hours and inflation, more government spending still boosts output only in the first quarter, and higher interest rates cause a recession and a hump-shaped decline in inflation.

Quantitatively though, there are some noticeable differences. With correlated disturbances, an increase in government spending lowers output by more after the first quarter, partly because if has a more modest and temporary impact on hours. Also, the response of inflation to monetary policy is more than twice as large with dynamically-correlated rather than independent disturbances.

The models also both do roughly well at explaining the data: the log marginal predictive densities of the models with independent and dynamically-correlated disturbances are within 5 log points of each other. Part of this similarity is explained by the many sources of endogenous dynamics in the model that reduce the impact of correlated disturbances. Another part is due to the imprecision with which many of these moments are estimated.

4.3 Variance decompositions and statistical parameters

Table 3 shows the median variance decompositions for output, hours, real wages and inflation in the short run (1 quarter ahead), the long run (unconditionally), and at business cycle frequencies (2 years and 8 years ahead).[24] With independent disturbances, the fluctuations in output and hours are accounted mostly by government spending, risk premium and investment-specific productivity at the shorter horizon. At longer horizons, as Smets and Wouters (2007) emphasized, it is the wage-markup disturbance that dominates.

With correlated disturbances, the conclusions are similar at the short-run frequencies but very different at the business-cycle and long-run frequencies. Focusing on the variance of output, wage markups go from accounting for 47% and 49% at the 8-year and infinite horizon with independent disturbances, to only 25% and 9% with correlated disturbances. The two productivity disturbances and government spending now explain 80% of the variance of output in the long run, and as much as 35% at the 2-year horizon. Looking instead at the variance of inflation, again the role of wage markup declines significantly when we allow for correlated disturbances, and the difference from the independent-disturbances case increases with the horizon. Across all series, there is an increase in the role of productivity

[24] The 90% credible sets are in the appendix.

and government spending in accounting for the business cycle.

Next, looking at the nuisance parameters in the (dynamic-)correlation structure, there are a few that are significantly different from zero. (The whole posterior distribution is in the appendix.) First, the correlation between total factor productivity and government spending is large, and goes in both directions (Φ_{AG} and Φ_{GA}). In part, this justifies the Smets and Wouters (2007) modelling assumption of allowing for contemporaneous correlation between these two variables. It also shows that while the new Keynesian model is more involved than the simple RBC model from section 2, it is still missing an important role for fiscal policy rules.[25]

Second, all of the other significant correlations (Φ_{BA}, Φ_{BEI}, Φ_{BER}, Φ_{ERB}, Φ_{EIA}, Φ_{EIB}, Φ_{EIG}, Φ_{REI}) involve either the risk-premium disturbance or investment-specific productivity. This suggests that an important direction for future research building on this model should focus on endogenizing these disturbances. Combining this observation with the differing estimates for the elasticity of investment adjustment costs, suggests that models with financial imperfections (e.g., Cúrdia and Woodford, 2009) seem particularly promising.

4.4 Lessons from correlated disturbances in the new Keynesian model

To conclude, allowing for correlated disturbances confirmed some of the lessons from previous estimates of the Smets-Wouters model. There are three changes though, that again highlight the need to allow for correlated disturbances to robustify inference and to point the direction of future research. First, the size of the response of the economy to government spending and monetary policy disturbances depends on how disturbances are modelled, recommending caution in using this model to too finely tune the economy. Second, we found that the much debated finding that markup disturbances are important is not robust. The role of wage markups is much reduced for all variables and becomes insignificant for output and wages beyond a few quarters.[26] Third, as in the simple model of section 2, the results

[25] Alternatively, both models assume a closed economy, so G may be capturing net exports. Since both theory and casual empirics suggest that the trade balance is sensitive to the business cycle, this reinforces our point that it is inappropriate to assume an exogenous G.

[26] See Chari, Kehoe, McGrattan (2009) for some of the debate, and Justiniano and Primiceri (2008) for an alternative estimation approach that converges with our results that wage markups are not as important as previously thought.

showed that it is important to account for endogenous fiscal policy responses to the business cycle. Moreover, the main missing element in the endogenous dynamics of the model is in modelling risk and investment.

5 Conclusion

DSGE modelling has made great strides in the last decade, in particular in the area of estimation and statistical inference. Because this work is in its infancy, there are still some clear holes in our knowledge that must be filled. This paper identified one of these holes: the strong and incredible restrictions that models typically place on the exogenous disturbances. Using well-known points in simultaneous-equation econometrics, we argued that these restrictions could severely hamper the model's ability to fit the data and severely bias inferences on key parameters and model predictions. We proposed the alternative of allowing for correlated disturbances, in the tradition of Zellner (1962).

The main obstacle to allowing for correlated disturbances is that it introduces a large number of nuisance parameters. We proposed a new method for estimating DSGE models, based on using conjugate families for some conditional posterior distributions. The algorithm is also valid and useful with uncorrelated disturbances, and with correlated disturbances it makes previously infeasible estimation now possible.

We applied the method to a simple real business cycle mode, and found that many apparent empirical puzzles in this model were easily accounted for by its omission of the strong correlation between government spending and productivity disturbances. This suggests that endogenous countercyclical fiscal policy is the main missing element to make this model roughly consistent with the data.

We then studied the impact of correlated disturbances in a more involved monetary business cycle model. We found that disturbances to markups are much less important once one accounts for correlated disturbances. Rather, it is productivity and fiscal policy that drives the significant part of the business cycle previously ascribed to markups, and again endogenizing it is the priority for future research. Our method again not only showed which of the inferences were not robust to correlated disturbances, but also provided hints on where modelling should turn to next. Endogenously modelling risk premia and investment-specific productivity disturbances seems to be the most promising avenue to bringing this model closer to the data.

Appendix

The New Keynesian business cycle model. We follow Smets and Wouters (2007) closely, including keeping their notation in this appendix as much as we can. The only change is for the statistical parameters to fit our general setup in section 3. The notation refers to: y_t is output, c_t is consumption, i_t is investment, q_t is the value of capital, l_t is hours worked, z_t is capital utilization, r_t is the nominal interest rate, π_t is inflation, w_t is the real wage, k_t is capital installed, μ_t^p is the price mark-up, and μ_t^w is the wage mark-up. The disturbance are all denoted by s_t with the superscript denoting the type of shock. The estimates of the model with independent AR(1) disturbances and dynamic correlated VAR(1) disturbances are in tables A.1 and A.2, respectively. Impulse responses at the median of the posterior are in figure A.1 and the credible sets for the variance decompositions are in table A.3.

The model has the following equations:

$$
\begin{aligned}
y_t &= (0.82 - i_y)\, c_t + i_y i_t + R_*^k k_y z_t + s_t^g \\
c_t &= c_1 c_{t-1} + (1 - c_1) E_t c_{t+1} + c_2 (l_t - E_t l_{t+1}) - c_3 (r_t - E_t \pi_{t+1} + s_t^b) \\
i_t &= i_1 i_{t-1} + (1 - i_1) E_t i_{t+1} + i_2 q_t + s_t^i \\
q_t &= q_1 E_t q_{t+1} + (1 - q_1) E_t \left(l_{t+1} - k_{t+1} + w_{t+1}\right) - (r_t - E_t \pi_{t+1} + s_t^b) \\
y_t &= \phi \left[\alpha k_{t-1} + \alpha z_t + (1 - \alpha) l_t + s_t^a\right] \\
z_t &= \left[(1 - \psi)/\psi\right] (l_t - k_t + w_t) \\
k_t &= k_1 k_{t-1} + (1 - k_1) i_t + k_2 s_t^i \\
\pi_t &= \pi_1 \pi_{t-1} + \pi_2 E_t \pi_{t+1} - \pi_3 \mu_t^p + s_t^p \\
w_t &= w_1 w_{t-1} + (1 - w_1) \left(E_t w_{t+1} + E_t \pi_{t+1}\right) - w_2 \pi_t + w_3 \pi_{t-1} - w_4 \mu_t^w + s_t^w \\
\mu_t^p &= \alpha (k_{t-1} + z_t - l_t) - w_t + s_t^a \\
\mu_t^w &= w_t - \left[\sigma_l l_t + (c_t - c_{t-1} \lambda/\gamma) / (1 - \lambda/\gamma)\right] \\
r_t &= \rho r_{t-1} + (1 - \rho)[r_\pi \pi_t + r_y (y_t - y_t^p)] + r_{\Delta y} \left(\Delta y_t - \Delta y_t^p\right) + s_t^r
\end{aligned}
$$

The reduced-form parameters are linked to structural parameters according to: $i_y = (\gamma - 0.975) k_y$, $c_1 = (\lambda/\gamma)(1 + \lambda/\gamma)$, $c_2 = [(\sigma_c - 1)(W_*^h L_*/C_*)/[\sigma_c(1 + \lambda/\gamma)]$, and $c_3 =$

$(1-\lambda/\gamma)/[(1+\lambda/\gamma)\sigma_c]$, $i_1 = 1/(1+\beta\gamma^{(1-\sigma_c)})$, $i_2 = i_1/\gamma^2\varphi$, $q_1 = 0.975\beta\gamma^{-\sigma_c}$, $k_1 = 0.975/\gamma$, $k_2 = (1-k_1)(1+\beta\gamma^{1-\sigma_c})\gamma^2\varphi$, $\pi_1 = \iota_p/(1+\beta\gamma^{1-\sigma_c}\iota_p)$, $\pi_2 = \pi_1\beta\gamma^{1-\sigma_c}/\iota_p$, $\pi_3 = (\pi_1/\iota_p)\left\{(1-\beta\gamma^{1-\sigma_c}\xi_p)(1-\xi_p)/\left\{\xi_p\left[10(\phi-1)+1\right]\right\}\right\}$, $w_1 = i_1$, $w_2 = w_1(1+\beta\gamma^{1-\sigma_c}\iota_w)$, $w_3 = \iota_w w_1$, $w_4 = w_1\left\{(1-\beta\gamma^{1-\sigma_c}\xi_w)(1-\xi_w)/\left\{\xi_w\left[10(\Phi_{SW}-1)+1\right]\right\}\right\}$, $\gamma^* = 100(\gamma-1)$, and k_y is the steady-state capital-output ratio and R_*^k is the steady-state rental rate of capital,

The structural parameters are: $\gamma^* = 100(\gamma-1)$ is the steady-state growth rate, l^* is the steady-state hours worked, π^* is the steady-state inflation rate, β is the discount factor, ϕ is one plus the share of fixed costs in production, σ_c is the elasticity of intertemporal substitution keeping labor fixed, λ is the degree of habit formation, ξ_w is the degree of wage stickiness, σ_l is the wage elasticity of labor supply, ξ_p is the degree of price stickiness, ι_w is the degree of wage indexation, ι_p is the degree of price indexation, ψ is a positive function of the steady-state elasticity of the capital utilization adjustment cost function that is φ, Φ_{SW} is the gross steady-state labor markup, ρ_{SW}, r_π, r_y and $r_{\Delta y}$ are the monetary policy-rule parameters, and α is the capital share.

References

Backus David K., Patrick J. Kehoe and Finn E. Kydland (1992). "International Real Business Cycles." *Journal of Political Economy*, 100 (4), 745-775.

Baxter, Marianne and Robert G. King (1993). "Fiscal Policy in General Equilibrium." *American Economic Review*, 83 (3), 315-334.

Basu Susanto, John G. Fernald and Miles S. Kimball (2006). "Are Technology Improvements Contractionary?" *American Economic Review*, 96 (5), 1418-1448.

Blanchard, Olivier J. and Charles M. Kahn (1980). "The Solution of Linear Difference Models under Rational Expectations." *Econometrica*, 48 (5), 1305-1311.

Carter, C. K. And R. Kohn (1994). "On Gibbs Sampling for State Space Models." *Biometrika*, 81, 541-553.

Chari, V. V., Patrick J. Kehoe and Ellen R. McGrattan (2007). "Business Cycle Accounting." *Econometrica*, 75 (3), 781-836.

Chari, V. V., Patrick J. Kehoe and Ellen R. McGrattan (2009). "New Keynesian Models:

Not Yet Useful for Monetary Policy." *American Economic Journal: Macroeconomics*, 1 (1), 242-266.

Chib, Siddartha (2001). "Markov Chain Monte Carlo Methods: Computation and Inference." In *Handbook of Econometrics*, edited by J. J. Heckman and E. E. Leamer, volume 5, chapter 57, pp. 3569-3649.

Chib, Siddartha and Srikanth Ramamurthy (2010). "Tailored Randomized-block MCMC Methods with Application to DSGE Models." *Journal of Econometrics*, 155, 19-38.

Christiano, Lawrence J. and Martin Eichenbaum (1992). "Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations." *American Economic Review*, 82 (3), 430-450.

Cochrane, D. and G. Orcutt (1949). "Application of Least Squares Regression to Relationships Containing Autocorrelated Error Terms." *Journal of the American Statistical Association*, 44, 32-61.

Cogley, Timothy and James M. Nason (1995). "Output Dynamics in Real-Business-Cycle Models." *American Economic Review*, 85 (3), 492-511.

Cúrdia, Vasco and Michael Woodford (2009). "Credit Frictions and Optimal Monetary Policy." Columbia University and FRB New York, manuscript.

Dedola, Luca and Stefano Neri (2007). "What does a technology shock do? A VAR analysis with model-based sign restrictions." *Journal of Monetary Economics*, 54 (2), 512-549.

Del Negro, Marco and Frank Schorfheide (2009). "Monetary Policy Analysis with Potentially Misspecified Models. " *American Economic Review*, 99 (4), 1415-1450.

Del Negro, Marco, Frank Schorfheide, Frank Smets and Rafael Wouters (2007). "On the Fit and Forecasting Performance of New Keynesian Models." *Journal of Business and Economic Statistics*, 25 (2), 123-162.

Durbin, James and Siem J. Koopman (2001). *Time Series Analysis by State Space Methods*. Oxford University Press: Oxford, UK.

Evans, Charles L. (1992). "Productivity Shocks and Real Business Cycles." *Journal of Monetary Economics*, 29 (2), 191-208.

Fair, Ray S. (2004). *Estimating How the Macroeconomy Works*. Harvard University Press.

Fernandez-Villaverde, Jesus (2009). "The Econometrics of DSGE Models." Working paper.

Fisher, Jonas M. (2006). "The Dynamic Effects of Neutral and Investment-Specific Technology Shocks" *Journal of Political Economy*, 114 (3), 413-451.

Francis, Neville and Valerie A. Ramey (2005). "Is the technology-driven real business cycle hypothesis dead? Shocks and aggregate fluctuations revisited." *Journal of Monetary Economics*, 52 (8), 1379-1399.

Gali, Jordi (1999). "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" *American Economic Review*, 89 (1), 249-271.

Gelman, Andrew and Donald B. Rubin (1992). "Inference from Iterative Simulation Using Multiple Sequences." *Statistical Science*, 7, 457-511.

Gelman, Andrew, John B. Carlin, Hal S. Stern and Donald B. Rubin (1998). *Bayesian Data Analysis*. Chapman & Hall / CRC Press.

Geweke, John (1989). "Bayesian Inference in Econometric Models Using Monte Carlo Integration." *Econometrica*, 57 (6), 1317-1339

Geweke, John (1992). "Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments." In J. M. Bernardo et al, eds., *Bayesian Statistics vol. 4*, Clarendon Press: Oxford UK.

Geweke, John (2005). *Contemporary Bayesian Econometrics and Statistics*. John Wiley and Sons.

Greenwood, Jeremy, Zvi Hercowitz, and Per Krusell (1997). "Long-Run Implications of Investment-Specific Technological Change." *American Economic Review*, 87 (3), 342-362.

Griliches, Zvi (1967). "Distributed Lags: A Survey." *Econometrica*, 35, 16-49.

Hall, Robert E. (1988). "Intertemporal Substitution in Consumption." *Journal of Political Economy*, 96 (2), 339-57.

Hall, Robert E. (1997). "Macroeconomic Fluctuations and the Allocation of Time." *Journal of Labor Economics*, 15 (1), S223-250.

Ireland, Peter N. (2004). "A method for taking models to the data." *Journal of Economic Dynamics and Control*, 28 (6), 1205-26.

Jorgenson, Dale W. (1966). "The Embodiement Hypothesis." *Journal of Political Economy*, 74 (1), 1-17.

Justiniano, Alejandro and Bruce Preston (2010). "Can Structural Small Open Economy Models Account for the Influence of Foreign Shocks?" *Journal of International Economics*, 81 (1), 61-74.

Justiniano, Alejandro and Giorgio E. Primiceri (2008). "Potential and Natural Output." Working paper.

Kadiyala, K. Rao and Sune Karlsson (1997). "Numerical Methods for Estimation and Inference in Bayesian VAR-Models." *Journal of Applied Econometrics*, 12 (2), 99-132.

Komunjer, Ivana and Serena Ng (2010). "Dynamic Identification of DSGE Models." Manuscript.

Newey, Whitney K. and Kenneth D. West (1987). "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55 (3), 703-708.

Parkin, Michael (1988). "A Method for Determining Whether Parameters in Aggregative Models Are Structural." *Carnegie-Rochester Conference Series on Public Policy*, 29, 212-252.

Prescott, Edward C. (1986). "Theory Ahead of Business Cycle Measurement." *FRB Minneapolis Quarterly Reiview,* 9-22.

Rabanal, Pau, Juan F. Rubio-Ramirez and Vicente Tuesta (2010). "Cointegrated TFP Shocks and International Business Cycles." *Journal of Monetary Economics*, forthcoming.

Reis, Ricardo (2008). "Using VARs to Identify Models of Fiscal Policy: A Comment on Perotti." *NBER Macoeconomics Annual 2007*, 22, 227-236.

Rotemberg, Julio J. and Michael Woodford (1992). "Oligopolistic Pricing and the Effects of Aggregate Demand on Economic Activity." *Journal of Political Economy*, 100 (6), 1153-1207.

Sargent, Thomas J. (1978). "Estimation of Dynamic Labor Demand Schedules under Rational Expectations." *Journal of Political Economy*, 86(6), 1009-1044.

Schmitt-Grohe, Stephanie and Martin Uribe (2010). "Business Cycles with a Common

Trend in Neutral and Investment-Specific Productivity." *Review of Economic Dynamics*, forthcoming.

Shapiro, Matthew D. and Mark W. Watson (1986). "Sources of Business Cycle Fluctuations." *NBER Macroeconomics Annual*, 3, 111-148.

Sims, Christopher A. (2002). "Solving Linear Rational Expectations Models." *Computational Economics*, 20 (1-2), 1-20.

Sims, Christopher A. and Tao Zha. (1998). "Bayesian Methods for Dynamic Multivariate Models." *International Economic Review*, 39 (4), 949-968.

Smets, Frank and Rafael Wouters (2003). "An Estimated Stochastic Dynamic General Equilibrium Model of the Euro Area." *Journal of the European Economic Association*, 1 (5), 1123-1175.

Smets, Frank and Rafael Wouters (2007). "Shocks and Frictions in U.S. Business Cycles: A Bayesian DSGE Approach." *American Economic Review*, 97 (3), 586-606.

Solow, Robert M. (1957). "Technical Change and the Aggregate Production Function." *Review of Economics and Statistics*, 39, 312-320.

Stock, James H. (2010). "The Other Transformation in Econometric Practice: Robust Tools for Inference." *Journal of Economic Perspectives*, 24 (2), 83-94.

Tierney, L. (1994). "Markov Chains for Exploring Posterior Distributions." *Annals of Statistics*, 22, 1701-62.

Uhlig, Harald (2004). "Do Technology Shocks Lead to a Fall in Total Hours Worked?" *Journal of the European Economic Association*, 2 (2-3), 361-371.

Yogo, Motohiro (2004). "Estimating the Elasticity of Intertemporal Substitution When Instruments Are Weak." *Review of Economics and Statistics*, 86 (3), 797-810.

Zellner, Arnold (1962). "An Efficient Method of Estimating Seemingly Unrelated Regressions and Test for Aggregation Bias" *Journal of the American Statistical Association*, 57, 348-368.

Table 1. Prior Distributions for RBC model

Parameter	Density[a]	Mode	Percentile 5	Percentile 50	Percentile 95
Economic					
γ	*G*	0.6667	0.2347	1.1559	3.3311
θ	*G*	4.8480	2.7326	5.3290	9.2117
Statistical					
Panel A. Independent AR(1)s					
Φ_A	*N*	0.7525	0.5141	0.7459	0.9356
Φ_G	*N*	0.4255	0.1994	0.4297	0.6554
Ω_A	IG^2	.00010	.00006	.00015	.00050
Ω_G	IG^2	0.2297	0.1452	0.4002	1.7917
Panel B. Unrestricted VAR(1)					
Φ_{AA}	*N*	0.7525	0.5070	0.7413	0.9308
Φ_{AG}	*N*	0.0000	-0.0041	0.0001	0.0041
Φ_{GA}	*N*	0.0000	-12.960	-0.0558	12.720
Φ_{GG}	*N*	0.4255	0.1788	0.4237	0.6540
Ω_{AA}	*IW*	.00008	.00006	.00015	.00048
Ω_{AG}	*IW*	0.0000	-0.0080	0.0000	0.0081
Ω_{GG}	*IW*	0.2583	0.2012	0.4820	1.5017

a. The densities are the gamma (*G*), normal (*N*) and the inverse-Wishart (*IW*).

Table 2. Posterior Distributions for RBC model

Parameter	Mean	Mode	Percentile		
			5	50	95
Panel A. Independent AR(1)s					
Economic					
γ	1.4029	1.4234	0.4970	1.2435	2.8629
θ	0.6184	0.4896	0.2632	0.5471	1.2036
Statistical					
Φ_A	0.8173	0.8106	0.7422	0.8174	0.8923
Φ_G	0.7505	0.7518	0.6713	0.7520	0.8234
Ω_A	.00014	.00014	.00012	.00014	.00017
Ω_G	0.2706	0.2475	0.1928	0.2645	0.3684
Panel B. Unrestricted VAR(1)					
Economic					
γ	0.4301	0.4304	0.2892	0.4170	0.6060
θ	4.8550	4.3184	1.9072	4.6302	8.5641
Statistical					
Φ_{AA}	0.9385	0.9355	0.9058	0.9402	0.9656
Φ_{AG}	0.0048	0.0048	0.0041	0.0049	0.0054
Φ_{GA}	-8.62	-8.26	-11.21	-8.50	-6.25
Φ_{GG}	0.8805	0.8828	0.8362	0.8811	0.9232
Ω_{AA}	.00013	.00013	.00011	.00013	.00016
Ω_{AG}	0.0084	0.0071	0.0045	0.0080	0.0138
Ω_{GG}	2.0718	1.3527	0.7942	1.6752	4.6432

Table 3. Variance Decompositions in the Smets-Wouters model

	Shock						
Variable	Total productivity	Risk premium	Government spending	Investment productivity	Monetary Policy	Price markup	Wage markup
Panel A. Independent AR(1) disturbances							
1-quarter ahead							
Output	0.016	0.289	0.475	0.1160	0.065	0.025	0.003
Hours	0.421	0.160	0.274	0.082	0.034	0.006	0.014
Real wage	0.010	0.016	0.000	0.006	0.012	0.268	0.682
Inflation	0.026	0.003	0.001	0.007	0.014	0.807	0.137
2-years ahead							
Output	0.177	0. 075	0.184	0.191	0.095	0.083	0.163
Hours	0.158	0.075	0.203	0.149	0.087	0.059	0.242
Real wage	0.098	0.018	0.000	0.061	0.055	0.269	0.474
Inflation	0.050	0.008	0.003	0.022	0.050	0.408	0.443
8-years ahead							
Output	0.200	0.023	0.134	0.072	0.033	0.034	0.474
Hours	0.064	0.026	0.170	0.066	0.033	0.026	0.596
Real wage	0.330	0.011	0.001	0.080	0.044	0.189	0.304
Inflation	0.046	0.007	0.004	0.021	0.044	0.321	0.542
Unconditional							
Output	0.133	0.014	0.208	0.043	0.019	0.020	0.489
Hours	0.047	0.015	0.257	0.041	0.019	0.015	0.558
Real wage	0.379	0.010	0.001	0.073	0.040	0.172	0.280
Inflation	0.044	0.006	0.006	0.019	0.037	0.279	0.593
Panel B. Dynamic VAR(1) disturbances							
1-quarter ahead							
Output	0.008	0.454	0.392	0.032	0.016	0.053	0.023
Hours	0.468	0.229	0.212	0.023	0.007	0.014	0.032
Real wage	0.045	0.036	0.004	0.002	0.007	0.325	0.564
Inflation	0.054	0.009	0.020	0.022	0.023	0.638	0.202
2-years ahead							
Output	0.108	0.161	0.173	0.066	0.019	0.103	0.308
Hours	0.210	0.180	0.069	0.136	0.012	0.046	0.289
Real wage	0.300	0.023	0.057	0.021	0.029	0.227	0.275
Inflation	0.058	0.020	0.067	0.056	0.063	0.288	0.386
8-years ahead							
Output	0.237	0.051	0.275	0.096	0.011	0.027	0.252
Hours	0.128	0.104	0.059	0.097	0.019	0.035	0.484
Real wage	0.452	0.046	0.199	0.120	0.011	0.059	0.069
Inflation	0.068	0.027	0.081	0.066	0.060	0.258	0.383
Unconditional							
Output	0.381	0.044	0.287	0.129	0.005	0.011	0.094
Hours	0.177	0.089	0.094	0.107	0.016	0.028	0.407
Real wage	0.456	0.047	0.258	0.142	0.005	0.019	0.036
Inflation	0.224	0.036	0.175	0.098	0.032	0.145	0.231

Figure 1. Convergence in simulated RBC model with AR(1) disturbances: random-walk Metropolis versus conjugate-conditionals methods

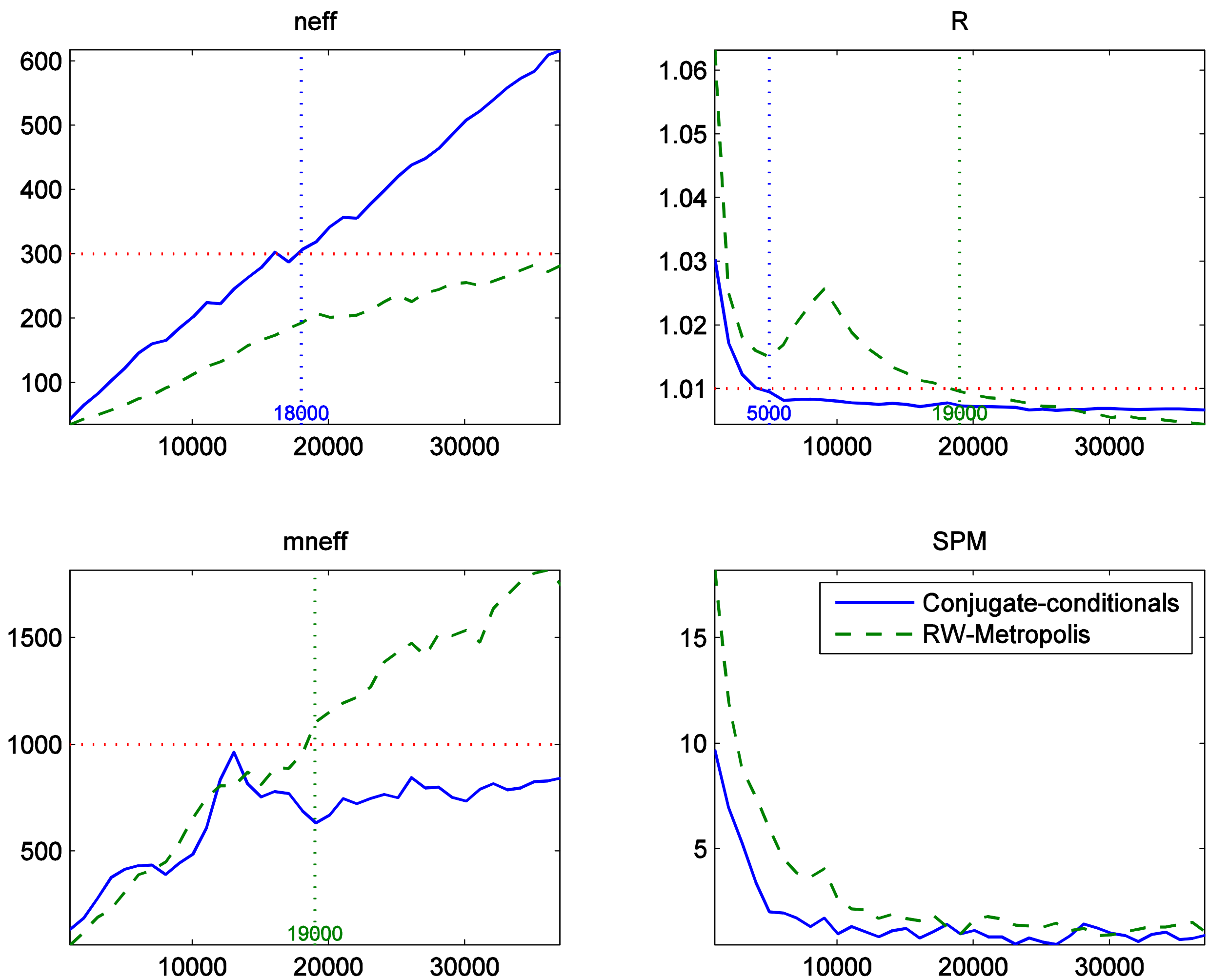
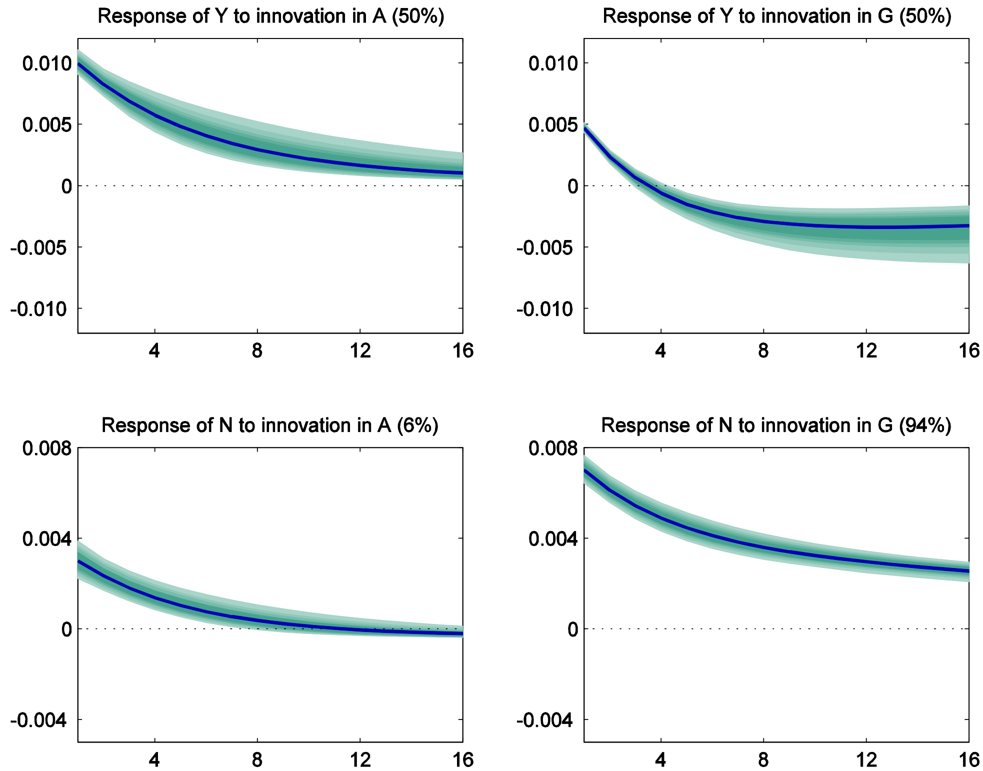


Figure 2. Impulse response functions in RBC model, median and distributions

Panel A. Independent AR(1)s case



Panel B. Unrestricted VAR(1)

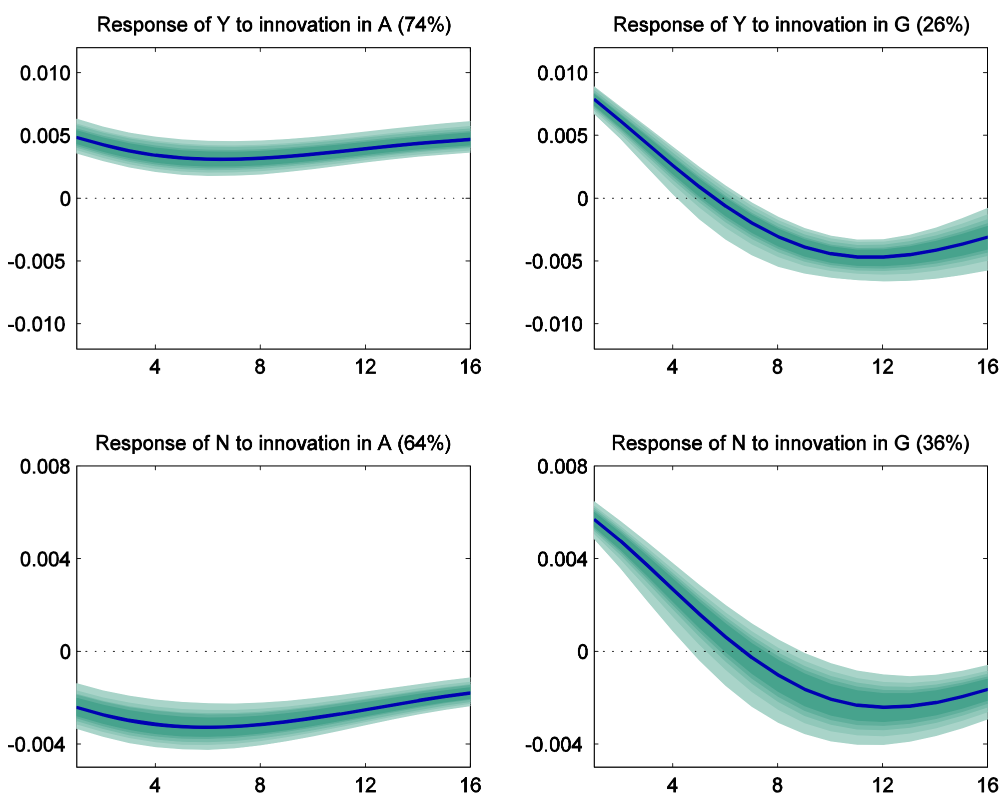


Figure 3. Median impulse response functions in the Smets-Wouters model, with independent and correlated disturbances

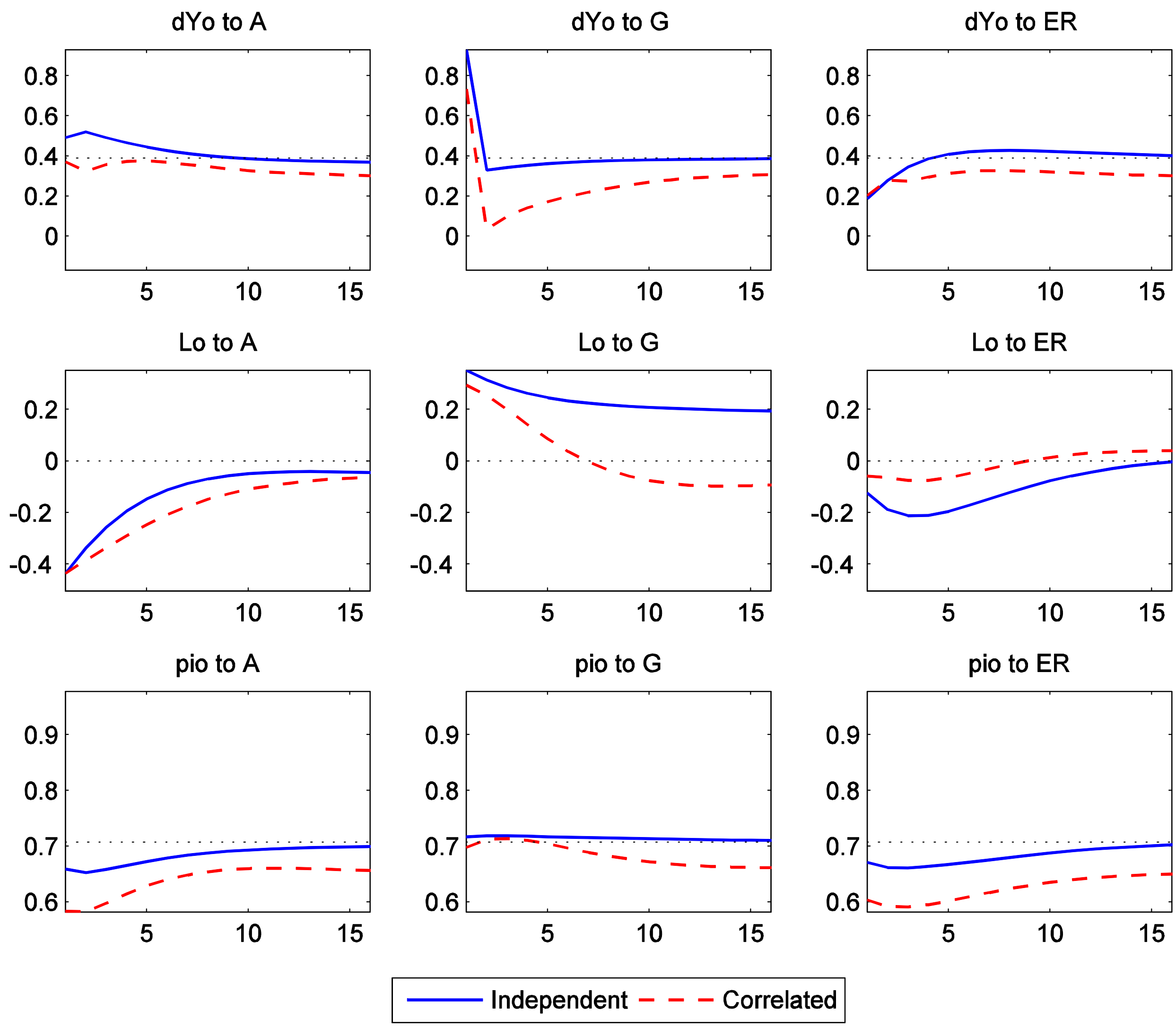


Table A.1. Prior and posterior distribution for MBC model, independent AR(1) disturbances

	Prior				Posterior					
	Dist	5%	Median	95%	Mode	Mean	SE	5%	Median	95%
γ^*	N	0.2355	0.4000	0.5645	0.3898	0.3864	0.0191	0.3523	0.3881	0.4145
l^*	N	-0.4935	0.0000	0.4935	0.0000	0.0003	0.3006	-0.4989	0.0020	0.4954
π^*	G	0.4652	0.6146	0.7931	0.6873	0.7100	0.1024	0.5454	0.7071	0.8843
$100(\beta^{-1} - 1))$	G	0.1111	0.2368	0.4339	0.1470	0.1698	0.0592	0.0830	0.1643	0.2765
ϕ	N	1.5327	4.0000	6.4673	6.1285	6.1955	1.1311	4.4047	6.1516	8.1194
σ_c	N	0.8832	1.5000	2.1168	1.4058	1.3673	0.1409	1.1508	1.3589	1.6113
λ	B	0.5242	0.7068	0.8525	0.7024	0.7083	0.0486	0.6218	0.7122	0.7807
ξ_w	B	0.3351	0.5000	0.6649	0.7056	0.6756	0.0701	0.5562	0.6788	0.7861
σ_l	N	0.7664	2.0000	3.2336	1.7248	1.7625	0.5421	0.9467	1.7220	2.7179
ξ_p	B	0.3351	0.5000	0.6649	0.7011	0.6845	0.0572	0.5850	0.6872	0.7735
ι_w	B	0.2526	0.5000	0.7474	0.5110	0.5137	0.1259	0.3061	0.5142	0.7203
ι_p	B	0.2526	0.5000	0.7474	0.2645	0.3024	0.1109	0.1438	0.2899	0.5046
ψ	B	0.2526	0.5000	0.7474	0.6195	0.6366	0.0693	0.5299	0.6326	0.7585
Φ_{SW}	N	1.0526	1.2500	1.4474	1.6617	1.6628	0.0764	1.5398	1.6608	1.7914
r_π	N	1.0888	1.5000	1.9112	1.9834	2.0435	0.1724	1.7654	2.0392	2.3341
ρ_{SW}	B	0.5701	0.7595	0.8971	0.8015	0.8008	0.0258	0.7562	0.8020	0.8405
r_y	N	0.0378	0.1200	0.2022	0.0846	0.0884	0.0207	0.0566	0.0872	0.1243
$r_{\Delta y}$	N	0.0378	0.1200	0.2022	0.2257	0.2257	0.0289	0.1788	0.2254	0.2739
α	N	0.2178	0.3000	0.3822	0.1676	0.1698	0.0179	0.1408	0.1695	0.1998
$\Phi_{A,1}$	N	0.1986	0.4964	0.7732	0.9601	0.9609	0.0139	0.9369	0.9618	0.9822
$\Phi_{B,1}$	N	0.2010	0.4959	0.7805	0.2021	0.2382	0.1478	0.0274	0.2206	0.5267
$\Phi_{G,1}$	N	0.1869	0.4994	0.7780	0.9945	0.9910	0.0062	0.9795	0.9922	0.9986
$\Phi_{EI,1}$	N	0.1957	0.4975	0.7853	0.7119	0.7147	0.0570	0.6204	0.7149	0.8089
$\Phi_{ER,1}$	N	0.1925	0.4958	0.7764	0.1698	0.1779	0.0713	0.0604	0.1787	0.2934
$\Phi_{EP,1}$	N	0.1967	0.4983	0.7772	0.7203	0.7053	0.0982	0.5365	0.7098	0.8575
$\Phi_{EW,1}$	N	0.1834	0.4979	0.7882	0.9802	0.9794	0.0098	0.9616	0.9807	0.9931
$-\Psi_{EP}$	B	0.1718	0.5000	0.8282	0.5470	0.5228	0.1363	0.2866	0.5291	0.7358
$-\Psi_{EW}$	B	0.1718	0.5000	0.8282	0.8926	0.8540	0.0641	0.7331	0.8653	0.9367
Ω_A	IG2	0.0291	0.0823	0.3889	0.2076	0.2143	0.0257	0.1758	0.2122	0.2596
Ω_B	IG2	0.0291	0.0823	0.3889	3.4472	3.9211	2.0819	1.0706	3.6287	7.8164
Ω_G	IG2	0.0291	0.0823	0.3889	0.3182	0.3285	0.0376	0.2723	0.3255	0.3946
Ω_{EI}	IG2	0.0291	0.0823	0.3889	0.2170	0.2272	0.0453	0.1621	0.2223	0.3092
Ω_{ER}	IG2	0.0291	0.0823	0.3889	0.0615	0.0648	0.0078	0.0528	0.0642	0.0786
Ω_{EP}	IG2	0.0291	0.0823	0.3889	0.0264	0.0279	0.0051	0.0203	0.0275	0.0371
Ω_{EW}	IG2	0.0291	0.0823	0.3889	0.0673	0.0701	0.0117	0.0526	0.0693	0.0906

Table A.2. Prior and posterior distributions for MBC model with correlated VAR(1) disturbances

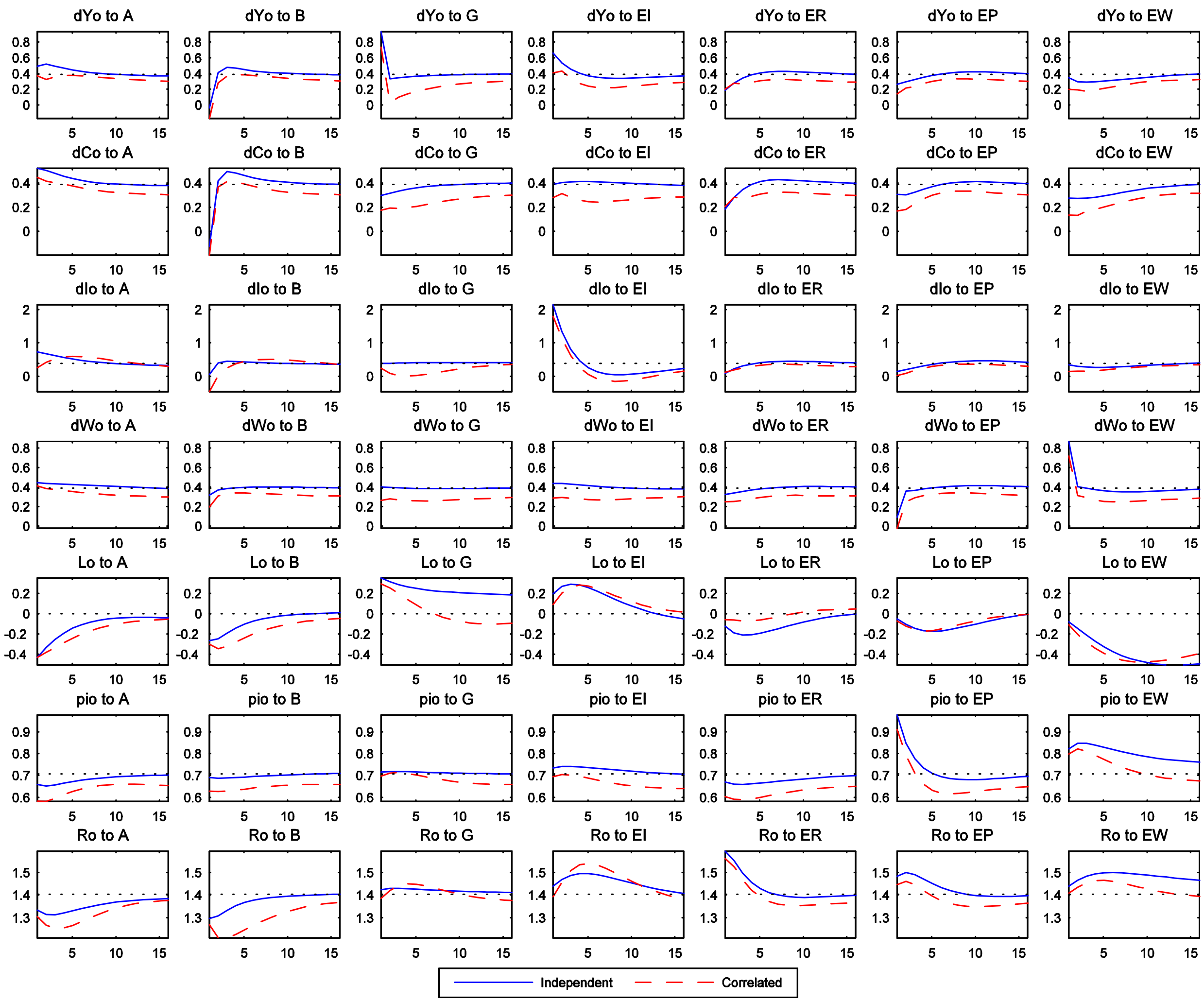
	Prior				Posterior					
	Dist	5%	Median	95%	Mode	Mean	SE	5%	Median	95%
γ^*	N	0.2355	0.4000	0.5645	0.2753	0.2964	0.0201	0.2623	0.2976	0.3271
l^*	N	-0.4935	0.0000	0.4935	-0.0000	-0.0001	0.2976	-0.4911	-0.0003	0.4878
π^*	G	0.4652	0.6146	0.7931	0.6090	0.6559	0.1024	0.4937	0.6525	0.8295
$100(\beta^{-1}-1))$	G	0.1111	0.2368	0.4339	0.2335	0.2645	0.0907	0.1304	0.2563	0.4259
ϕ	N	1.5327	4.0000	6.4673	5.1223	5.3067	1.1751	3.3973	5.2952	7.2588
σ_c	N	0.8832	1.5000	2.1168	1.4438	1.5421	0.2238	1.2028	1.5300	1.9303
λ	B	0.5242	0.7068	0.8525	0.5250	0.6873	0.0629	0.5668	0.6965	0.7745
ξ_w	B	0.3351	0.5000	0.6649	0.6400	0.5441	0.0545	0.4560	0.5430	0.6356
σ_l	N	0.7664	2.0000	3.2336	0.9592	1.2453	0.5283	0.4435	1.2086	2.1817
ξ_p	B	0.3351	0.5000	0.6649	0.5150	0.5832	0.0628	0.4772	0.5845	0.6841
ι_w	B	0.2526	0.5000	0.7474	0.4756	0.5619	0.1284	0.3453	0.5644	0.7687
ι_p	B	0.2526	0.5000	0.7474	0.2083	0.2912	0.1105	0.1324	0.2789	0.4932
ψ	B	0.2526	0.5000	0.7474	0.3518	0.4891	0.0585	0.3944	0.4885	0.5857
Φ_{SW}	N	1.0526	1.2500	1.4474	1.4191	1.4946	0.0734	1.3765	1.4930	1.6180
r_π	N	1.0888	1.5000	1.9112	1.5055	1.7383	0.1887	1.4361	1.7327	2.0606
ρ_{SW}	B	0.5701	0.7595	0.8971	0.7611	0.7535	0.0325	0.6979	0.7552	0.8035
r_y	N	0.0378	0.1200	0.2022	0.0564	0.0801	0.0302	0.0330	0.0787	0.1314
$r_{\Delta y}$	N	0.0378	0.1200	0.2022	0.2250	0.1913	0.0305	0.1418	0.1909	0.2420
α	N	0.2178	0.3000	0.3822	0.0395	0.0994	0.0183	0.0713	0.0984	0.1311
$\Phi_{A,A,1}$	N	0.1787	0.4932	0.7524	0.9302	0.9141	0.0348	0.8578	0.9137	0.9719
$\Phi_{A,B,1}$	N	-0.2779	-0.0033	0.2907	-0.0088	0.0096	0.0239	-0.0366	0.0135	0.0399
$\Phi_{A,G,1}$	N	-0.2820	-0.0025	0.2848	-0.2914	-0.2712	0.0521	-0.3588	-0.2702	-0.1876
$\Phi_{A,EI,1}$	N	-0.2886	-0.0049	0.2828	-0.0961	-0.0922	0.1051	-0.2675	-0.0899	0.0753
$\Phi_{A,ER,1}$	N	-0.2817	0.0021	0.2944	0.0222	-0.0142	0.1290	-0.2228	-0.0162	0.2012
$\Phi_{A,EP,1}$	N	-0.2768	-0.0006	0.2862	-0.0588	0.0074	0.1573	-0.2537	0.0096	0.2625
$\Phi_{A,EW,1}$	N	-0.2859	0.0016	0.3046	0.1247	0.0872	0.0995	-0.0767	0.0873	0.2513
$\Phi_{B,A,1}$	N	-0.2828	0.0012	0.2923	0.0693	0.2099	0.1217	0.0385	0.1974	0.4282
$\Phi_{B,B,1}$	N	0.1942	0.4910	0.7528	0.6947	0.2397	0.2056	-0.0321	0.1961	0.6489
$\Phi_{B,G,1}$	N	-0.2874	0.0048	0.3024	-0.1308	-0.1692	0.1635	-0.4602	-0.1550	0.0707
$\Phi_{B,EI,1}$	N	-0.2864	-0.0036	0.2867	-0.4177	-0.8591	0.4121	-1.5980	-0.8148	-0.2752
$\Phi_{B,ER,1}$	N	-0.2799	0.0076	0.3013	-0.5817	-1.2483	0.6686	-2.4081	-1.2064	-0.2460
$\Phi_{B,EP,1}$	N	-0.2786	-0.0013	0.2933	0.2720	0.7074	0.6548	-0.2060	0.6163	1.9356
$\Phi_{B,EW,1}$	N	-0.2873	-0.0018	0.3019	0.1370	0.4821	0.4196	-0.1093	0.4332	1.2362
$\Phi_{G,A,1}$	N	-0.2914	-0.0008	0.2920	-0.1899	-0.1787	0.0394	-0.2445	-0.1779	-0.1155
$\Phi_{G,B,1}$	N	-0.2791	0.0017	0.2847	-0.0630	-0.0196	0.0252	-0.0596	-0.0191	0.0187
$\Phi_{G,G,1}$	N	0.1996	0.4931	0.7622	0.7551	0.6769	0.0660	0.5686	0.6770	0.7846
$\Phi_{G,EI,1}$	N	-0.2826	0.0010	0.2724	0.4629	0.1924	0.1345	-0.0144	0.1843	0.4281
$\Phi_{G,ER,1}$	N	-0.2816	0.0036	0.2758	-0.0990	0.0240	0.1482	-0.2175	0.0228	0.2704
$\Phi_{G,EP,1}$	N	-0.2839	0.0016	0.2958	0.1392	0.1926	0.1760	-0.0969	0.1927	0.4808
$\Phi_{G,EW,1}$	N	-0.2794	-0.0026	0.2871	0.3713	0.1419	0.1190	-0.0489	0.1394	0.3405
$\Phi_{EI,A,1}$	N	-0.2904	-0.0039	0.2980	0.0595	0.0643	0.0280	0.0203	0.0632	0.1117
$\Phi_{EI,B,1}$	N	-0.2901	-0.0019	0.2762	-0.0385	-0.0324	0.0175	-0.0638	-0.0305	-0.0083
$\Phi_{EI,G,1}$	N	-0.2882	-0.0026	0.2836	-0.1085	-0.0779	0.0387	-0.1463	-0.0749	-0.0199
$\Phi_{EI,EI,1}$	N	0.1922	0.4911	0.7543	0.6312	0.6918	0.0575	0.5961	0.6928	0.7854
$\Phi_{EI,ER,1}$	N	-0.2855	0.0030	0.2898	-0.1230	-0.0878	0.0964	-0.2495	-0.0852	0.0646
$\Phi_{EI,EP,1}$	N	-0.2782	0.0022	0.2846	0.0318	0.0180	0.1008	-0.1424	0.0148	0.1882
$\Phi_{EI,EW,1}$	N	-0.2901	0.0012	0.2830	0.0213	0.0756	0.0912	-0.0741	0.0781	0.2184

	Prior				Posterior					
	Dist	5%	Median	95%	Mode	Mean	SE	5%	Median	95%
$\Phi_{ER,A,1}$	N	-0.2710	0.0030	0.2851	-0.0283	-0.0330	0.0174	-0.0624	-0.0327	-0.0052
$\Phi_{ER,B,1}$	N	-0.2781	0.0038	0.2789	-0.0810	-0.0335	0.0148	-0.0620	-0.0308	-0.0149
$\Phi_{ER,G,1}$	N	-0.2786	-0.0002	0.2843	0.0075	0.0192	0.0276	-0.0260	0.0190	0.0652
$\Phi_{ER,EI,1}$	N	-0.2735	0.0015	0.2936	0.1007	0.1143	0.0438	0.0459	0.1126	0.1896
$\Phi_{ER,ER,1}$	N	0.2035	0.4908	0.7542	0.1292	0.1838	0.0771	0.0577	0.1832	0.3102
$\Phi_{ER,EP,1}$	N	-0.2715	0.0017	0.2855	-0.0099	0.0011	0.0842	-0.1391	0.0017	0.1384
$\Phi_{ER,EW,1}$	N	-0.2837	0.0003	0.2837	0.0423	-0.0052	0.0640	-0.1122	-0.0040	0.0976
$\Phi_{EP,A,1}$	N	-0.2843	-0.0005	0.2877	-0.0107	-0.0058	0.0058	-0.0157	-0.0056	0.0034
$\Phi_{EP,B,1}$	N	-0.2901	0.0049	0.2851	-0.0007	0.0014	0.0055	-0.0071	0.0012	0.0107
$\Phi_{EP,G,1}$	N	-0.2960	-0.0016	0.2891	-0.0273	-0.0036	0.0102	-0.0204	-0.0035	0.0126
$\Phi_{EP,EI,1}$	N	-0.2846	-0.0007	0.2847	-0.0007	0.0072	0.0171	-0.0207	0.0078	0.0336
$\Phi_{EP,ER,1}$	N	-0.2898	0.0014	0.2903	0.0253	0.0062	0.0399	-0.0570	0.0051	0.0735
$\Phi_{EP,EP,1}$	N	0.1988	0.4934	0.7435	0.8069	0.6629	0.0842	0.5205	0.6660	0.7962
$\Phi_{EP,EW,1}$	N	-0.2785	0.0025	0.2842	-0.0243	-0.0082	0.0222	-0.0484	-0.0057	0.0235
$\Phi_{EW,A,1}$	N	-0.2686	-0.0004	0.2937	0.0099	0.0055	0.0083	-0.0083	0.0055	0.0191
$\Phi_{EW,B,1}$	N	-0.2822	-0.0018	0.2723	-0.0008	0.0029	0.0081	-0.0091	0.0027	0.0149
$\Phi_{EW,G,1}$	N	-0.2895	0.0008	0.2891	0.0124	0.0194	0.0132	-0.0003	0.0183	0.0429
$\Phi_{EW,EI,1}$	N	-0.2707	0.0043	0.2966	-0.0247	-0.0024	0.0261	-0.0434	-0.0034	0.0427
$\Phi_{EW,ER,1}$	N	-0.2839	-0.0002	0.2809	-0.0082	-0.0194	0.0544	-0.1077	-0.0193	0.0676
$\Phi_{EW,EP,1}$	N	-0.2856	-0.0028	0.2851	0.0002	-0.0035	0.0532	-0.0854	-0.0061	0.0867
$\Phi_{EW,EW,1}$	N	0.1819	0.4911	0.7657	0.9735	0.9422	0.0331	0.8826	0.9481	0.9830
$-\Psi_{EP}$	B	0.1718	0.5000	0.8282	0.5134	0.3397	0.1212	0.1422	0.3382	0.5421
$-\Psi_{EW}$	B	0.1718	0.5000	0.8282	0.9665	0.6739	0.1094	0.4751	0.6877	0.8264
Ω_{A}	IG2	0.0291	0.0823	0.3889	0.2431	0.2257	0.0284	0.1828	0.2235	0.2756
Ω_{B}	IG2	0.0291	0.0823	0.3889	0.4768	5.1996	3.3725	0.6226	4.8122	11.5165
Ω_{G}	IG2	0.0291	0.0823	0.3889	0.2563	0.2584	0.0310	0.2120	0.2562	0.3126
Ω_{EI}	IG2	0.0291	0.0823	0.3889	0.0582	0.0912	0.0334	0.0473	0.0854	0.1552
Ω_{ER}	IG2	0.0291	0.0823	0.3889	0.0497	0.0542	0.0066	0.0444	0.0537	0.0659
Ω_{EP}	IG2	0.0291	0.0823	0.3889	0.0296	0.0257	0.0052	0.0181	0.0252	0.0351
Ω_{EW}	IG2	0.0291	0.0823	0.3889	0.0855	0.0724	0.0132	0.0528	0.0712	0.0961

Table A.3. Variance Decompositions, Smets-Wouters model, 5% and 95% in the posterior

	Shock						
Variable	Total productivity	Risk premium	Government spending	Investment productivity	Monetary Policy	Price markup	Wage markup
Panel A. Independent AR(1) disturbances							
1-quarter ahead							
Output	.004, .044	.234, .355	.397, .545	.072, .168	.043, .099	.016, .038	.000, .017
Hours	.349, .491	.123, .210	.229, .326	.053, .120	.021, .058	.002, .011	.006, .028
Real wage	.002, .029	.004, .049	.000, .001	.002, .013	.003, .030	.208, .341	.583, .756
Inflation	.011, .051	.001, .012	.000, .003	.001, .023	.005, .033	.680, .907	.069, .222
2-years ahead							
Output	.109, .264	.047, .137	.121, .267	.107, .299	.051, .162	.054, .124	.069, .297
Hours	.101, .226	.047, .144	.147, .269	.092, .229	.048, .145	.036, .093	.122, .403
Real wage	.028, .231	.003, .046	.001, .002	.004, .107	.024, .099	.293, .389	.332, .624
Inflation	.029, .083	.006, .026	.000, .007	.028, .065	.026, .091	.181, .544	.315, .557
8-years ahead							
Output	.111, .323	.013, .047	.066, .246	.034, .143	.014, .072	.017, .065	.315, .633
Hours	.039, .103	.014, .055	.093, .279	.034, .121	.015, .069	.014, .049	.420, .752
Real wage	.121, .590	.004, .027	.000, .002	.031, .164	.018, .087	.112, .323	.147, .520
Inflation	.023, .079	.002, .022	.001, .010	.005, .064	.021, .083	.223, .435	.408, .674
Unconditional							
Output	.048, .278	.005, .031	.051, .642	.012, .106	.005, .051	.006, .048	.201, .762
Hours	.018, .088	.005, .035	.076, .674	.012, .096	.006, .048	.005, .035	.227, .828
Real wage	.130, .708	.004, .024	.000, .005	.025, .157	.013, .084	.084, .306	.108, .512
Inflation	.018, .081	.002, .018	.002, .014	.004, .062	.015, .076	.146, .405	.429, .786
Panel B. Dynamic VAR(1) disturbances							
1-quarter ahead							
Output	.000, .067	.355, .560	.259, .488	.003, .076	.001, .060	.021, .114	.005, .073
Hours	.383, .547	.173, .303	.152, .275	.003, .050	.000, .031	.003, .041	.014, .075
Real wage	.014, .101	.006, .093	.000, .026	.000, .016	.000, .032	.219, .440	.419, .695
Inflation	.021, .104	.000, .067	.001, .062	.001, .074	.001, .086	.482, .790	.111, .308
2-years ahead							
Output	.035, .220	.082, .300	.093, .288	.017, .190	.002, .098	.028, .221	.173, .460
Hours	.110, .323	.069, .342	.035, .136	.036, .289	.001, .074	.006, .144	.145, .487
Real wage	.152, .458	.008, .090	.007, .182	.003, .116	.001, .125	.097, .404	.110, .477
Inflation	.021, .120	.001, .130	.009, .170	.005, .180	.007, .182	.187, .426	.247, .528
8-years ahead							
Output	.098, .395	.023, .121	.143, .428	.027, .225	.001, .053	.008, .084	.110, .460
Hours	.059, .250	.037, .254	.022, .154	.033, .263	.003, .099	.008, .120	.286, .671
Real wage	.311, .590	.009, .126	.094, .323	.032, .243	.001, .050	.021, .156	.028, .162
Inflation	.028, .142	.004, .124	.017, .194	.013, .184	.009, .173	.167, .377	.241, .523
Unconditional							
Output	.189, .552	.010, .126	.162, .432	.030, .254	.001, .030	.002, .048	.022, .308
Hours	.076, .391	.028, .225	.030, .250	.037, .252	.003, .085	.007, .103	.154, .636
Real wage	.309, .607	.005, .134	.135, .398	.038, .264	.000, .028	.004, .071	.008, .125
Inflation	.074, .431	.007, .116	.059, .321	.033, .203	.006, .116	.048, .267	.080, .436

Figure A.1. Median impulse response functions in the Smets-Wouters model, with independent and correlated disturbances



Variables: dY is output growth, dCo is consumption growth, dIo is investment, dWo is wage growth, Lo is hours, pio is inflation, and Ro is the nominal interest rate.

Disturbances: total factor productivity (A), risk premium (B), government spending (G), investment-specific productivity (EI), nominal interest rates (ER), price markups (EP), wage markups (EW).

Correlated Disturbances and U.S. Business Cycles*

Suplementary Appendix

(Not for publication)

Vasco Cúrdia[†]
Federal Reserve Bank of New York

Ricardo Reis[‡]
Columbia University

This technical appendix describes four dynamic stochastic general equilibrium (DSGE) models log-linear representation and how they map to the assumptions in the main paper. The models chosen are meant to represent different degrees of complexity of the DSGE. In section 1 we present the first model discussed in the model — a real business cycle (RBC) DSGE model. In sections 2 and 3 we present two small scale New Keynesian DSGE models. The first is discussed in Lubik and Schorfheide (2004) and is shown here as the most basic New Keynesian (NK) framework which is taken to the data – but more complex than the RBC one. The second, discussed in Cúrdia, Ferrero, and Tambalotti (2010), is an extension of the NK framework to include some more complexity, while keeping it fairly simple, estimated on the same three observable variables as the previous one. Finally, in section 4 we present the model of Smets and Wouters (2007), which is also discussed in the main paper as a medium scale NK model. For more complex large scale DSGE models see for example Christiano, Motto, and Rostagno (2009). The first model is described in detail, including the exact matrices in the state space representation. For the other models we only show the key pieces in terms of our notation, without much discussion.

1 Simple Real Business Cycle DSGE Model

The first model is a basic RBC with productivity and government expenditures shocks, discussed in section 2 of the paper. It is based on Christiano and Eichenbaum (1992).

*The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Federal Reserve Bank of New York or the Federal Reserve System.

[†]*E-mail*: vasco.curdia@ny.frb.org

[‡]*E-mail*: rreis@columbia.edu

Log-linear representation:

$$0 = E_t\left[\frac{\alpha\beta Y}{K}\left(\hat{Y}_{t+1}-\hat{K}_t\right)+\gamma^{-1}\left(\hat{Y}_t-\hat{Y}_{t+1}\right)+\left(\theta(1-\gamma^{-1})\frac{N}{1-N}-\frac{\gamma^{-1}}{1-N}\right)\left(\hat{N}_t-\hat{N}_{t+1}\right)\right] \quad (1.1)$$

$$0 = (C-Y)\,\hat{Y}_t+G\hat{G}_t-\frac{C}{1-N}\hat{N}_t+K\hat{K}_t-(1-\delta)K\hat{K}_{t-1}, \quad (1.2)$$

$$0 = \hat{Y}_t-(1-\alpha)(\hat{A}_t+\hat{N}_t)-\alpha\hat{K}_{t-1}, \quad (1.3)$$

Endogenous variables:

$$y_t \equiv \left(\hat{Y}_t,\hat{K}_t,\hat{N}_t\right),$$

corresponding to output, capital and labor, in log-deviations from steady state; exogenous disturbances

$$s_t \equiv \left(\hat{A}_t,\hat{G}_t\right),$$

corresponding to productivity and government expenditure shocks; (Y,K,N) correspond to the steady state values; parameters (α,β,δ) are calibrated; and parameters (γ,θ) are estimated.

Observation equations:

$$Y_t^{obs} = \hat{Y}_t, \quad (1.4)$$

$$N_t^{obs} = \hat{N}_t, \quad (1.5)$$

where $x_t \equiv \left(Y_t^{obs},N_t^{obs}\right)$ is the vector of observables.

The shock structure takes the form of

$$s_t = \Phi\left(L\right)s_{t-1}+e_t, \quad (1.6)$$

$$Var\left(e_t\right) = \Omega, \quad (1.7)$$

just as mentioned in the paper, with $e_t \equiv \left(e_t^A,e_t^G\right)$ the vector of innovations. The vector of statistical parameters to be estimated depends on the specific case considered. For the AR(1) case, we have

$$\Phi_{AR1} \equiv \begin{bmatrix} \Phi_A & 0 \\ 0 & \Phi_G \end{bmatrix},$$

$$\Omega_{AR1} \equiv \begin{bmatrix} \Omega_A & 0 \\ 0 & \Omega_G \end{bmatrix},$$

and the statistical parameters vector is

$$\sigma_{AR1} \equiv \left(\Phi_A,\Phi_G,\Omega_A,\Omega_G\right).$$

For the VAR(1) case then we get

$$\Phi_{VAR1} \equiv \begin{bmatrix} \Phi_{AA} & \Phi_{AG} \\ \Phi_{GA} & \Phi_{GG} \end{bmatrix},$$

$$\Omega_{VAR1} \equiv \begin{bmatrix} \Omega_{AA} & \Omega_{AG} \\ \Omega_{GA} & \Omega_{G} \end{bmatrix},$$

and the statistical parameters vector is

$$\sigma_{VAR1} \equiv \left(\Phi_{AA}, \Phi_{AG}, \Phi_{GG}, \Phi_{GA}, \Omega_{AA}, \Omega_{AG}, \Omega_{GG}\right),$$

and notice that because Ω is a covariance matrix then it must be the case that $\Omega_{GA} = \Omega_{AG}$.

The observation equations above imply that $H_1 = H_3\left(L\right) = 0$ and H_2 is a matrix with ones and zeros mapping the endogenous variables y_t to the observable variables x_t,

$$H_2 \equiv \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}.$$

The log-linear equations satisfy the canonical form proposed in Sims (2002):

$$\Psi_0\left(\varepsilon\right) y_t = \Psi_1\left(\varepsilon\right) y_{t-1} + \Psi_2\left(\varepsilon\right) s_t + \Psi_3\left(\varepsilon\right) w_t, \tag{1.8}$$

where the matrices are:

$$\Psi_0\left(\varepsilon\right) \equiv \begin{bmatrix} \gamma^{-1} - \frac{\alpha\beta Y}{K} & \theta(1-\gamma^{-1})\frac{N}{1-N} - \frac{\gamma^{-1}}{1-N} & 0 \\ C - Y & -\frac{C}{1-N} & K \\ 1 & -(1-\alpha) & 0 \end{bmatrix},$$

$$\Psi_1\left(\varepsilon\right) \equiv \begin{bmatrix} \gamma^{-1} & \theta(1-\gamma^{-1})\frac{N}{1-N} - \frac{\gamma^{-1}}{1-N} & -\frac{\alpha\beta Y}{K} \\ 0 & 0 & (1-\delta)K \\ 0 & 0 & \alpha \end{bmatrix},$$

$$\Psi_2\left(\varepsilon\right) \equiv \begin{bmatrix} 0 & 0 \\ 0 & -G \\ 1-\alpha & 0 \end{bmatrix},$$

$$\Psi_3\left(\varepsilon\right) \equiv \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix},$$

and w_t is a single endogenous expectations shock, defined as

$$w_t \equiv \left(\frac{\alpha\beta Y}{K} - \gamma^{-1}\right)\left(\hat{Y}_t - E_{t-1}\hat{Y}_t\right) - \left(\theta(1-\gamma^{-1})\frac{N}{1-N} - \frac{\gamma^{-1}}{1-N}\right)\left(\hat{N}_t - E_{t-1}\hat{N}_t\right). \tag{1.9}$$

As discussed in that paper, the solution to this forward-backward looking problem, combining (1.6) with (1.8), is given by

$$y_t = \Lambda_1(\varepsilon,\sigma)\, y_{t-1} + \Lambda_2(\varepsilon,\sigma)\, s_t \tag{1.10}$$

where $\Lambda_1(\varepsilon,\sigma)$ are nonlinear functions of (ε,σ).[1] This fits the framework assumed in the our paper. Furthermore, under this state space representation, and because the number of observables, n_x, is equal to the number of disturbances, n_s, it is possible to apply the Kalman filter and generate a well behaved likelihood function, as described in Hamilton (1994). With proper priors defined, then assumption 1a) in the paper holds. Because certainty equivalence applies to all models that satisfy the canonical form (1.8), then assumption 1b) also holds. Assumption 1c) is also satisfied by construction as written in (1.6) — all that it implies is that we do not constrain any individual elements of the matrices Φ and Ω, other than setting them to be diagonal in the AR(1) case.

2 Basic New Keynesian DSGE Model

Next we consider the model proposed in Lubik and Schorfheide (2004). This is a prototypical NK-DSGE model in its most basic form. A similar model is also considered in An and Schorfheide (2006). All variables and parameters are explained in Lubik and Schorfheide (2004). Here we just show the representation in terms of our assumptions, without any further interpretation.

Log-linear equations:

$$\tilde{x}_t = E_t\tilde{x}_{t+1} - \tau\left(\tilde{R}_t - E_t\tilde{\pi}_{t+1}\right) + g_t, \tag{2.1}$$

$$\tilde{\pi}_t = \beta E_t\tilde{\pi}_{t+1} + \kappa\left(\tilde{x}_t - z_t\right), \tag{2.2}$$

$$\tilde{R}_t = \rho_R\tilde{R}_{t-1} + (1-\rho_R)\left(\psi_1\tilde{\pi}_t + \psi_2\left[\tilde{x}_t - z_t\right]\right) + r_t. \tag{2.3}$$

Observation equations:

$$x_t^{obs} = \tilde{x}_t, \tag{2.4}$$

$$\pi_t^{obs} = \pi^* + 4\tilde{\pi}_t, \tag{2.5}$$

$$R_t^{obs} = (r^* + \pi^*) + 4\tilde{R}_t \tag{2.6}$$

[1] Notice that this representation is true for solutions to the rational expectations equilibrium (REE) that exist and are unique. In the implementation it is usual for authors to discard parameter values that do not satisfy existence and uniqueness, which is equivalent to considering a joint prior that is truncated so as to give zero probability to such outcomes.

where x_t^{obs} is the log of real per capita GDP detrended using HP filter and multiplied by 100; π_t^{obs} is CPI inflation, in annual percentage points; and R_t^{obs} is the average Federal Funds Rate, in annual percentage points.

Observable variables:

$$x_t \equiv \left(x_t^{obs}, \pi_t^{obs}, R_t^{obs}\right).$$

Endogenous variables:

$$y_t \equiv \left(\tilde{x}_t, \tilde{\pi}_t, \tilde{R}_t\right).$$

Exogenous disturbances:

$$s_t \equiv (g_t, z_t, r_t).$$

Innovations:

$$e_t \equiv (e_t^g, e_t^z, e_t^r).$$

Economic parameters:

$$\varepsilon \equiv \left(\psi_1, \psi_2, \rho_R, \pi^*, r^*, \kappa, \tau^{-1}\right).$$

(Notice that $\beta = 1/(1 + r^*)$.)

Statistical parameters: depends on the assumption,

$$\begin{aligned}
\sigma_{AR1} &\equiv (\Phi_g, \Phi_z, \Phi_r, \Omega_g, \Omega_z, \Omega_r), \text{ or} \\
\sigma_{VAR1} &\equiv (\Phi_{gg}, \Phi_{gz}, \Phi_{gr}, \Phi_{zg}, \Phi_{zz}, \Phi_{zr}, \Phi_{rg}, \Phi_{rz}, \Phi_{rr}, \Omega_{gg}, \Omega_{gz}, \Omega_{gr}, \Omega_{zz}, \Omega_{zr}, \Omega_{rr}).
\end{aligned}$$

Equations (2.1)-(2.3) satisfy the canonical form (1.8),[2] hence the solution to this model can be represented as in (1.10), consistent with our paper's framework. As in the previous model, the number of observables, n_x, is equal to the number of disturbances, n_s, hence again it is possible to apply the Kalman filter and generate a well behaved likelihood function. With proper priors defined, then assumption 1a) in the paper holds. Because certainty equivalence applies to all models that satisfy the canonical form (1.8), then assumption 1b) also holds. Assumption 1c) is also satisfied by construction, just as previously discussed.

3 Small New Keynesian DSGE Model

In this section we offer an alternative NK-DSGE model that is more complex than the previous one, considering habit in consumption, unit root in the productivity shock, and price indexation. As an example of such a setup we use the specific model described in Cúrdia, Ferrero, and Tambalotti (2010) and Sbordone, Tambalotti, Rao, and Walsh (2010). All variables and parameters

[2] Lubik and Schorfheide (2004) actually describe the exact matrices and how the model fits in the canonical form, much like our description of the previous model.

are explained in Cúrdia, Ferrero, and Tambalotti (2010). Here we just show the representation in terms of our assumptions, without any further interpretation.

Log-linear equations:

$$\tilde{x}_t = E_t\tilde{x}_{t+1} - \varphi_\gamma^{-1}\left(i_t - E_t\pi_{t+1} - r_t^e\right), \tag{3.1}$$

$$\tilde{x}_t = \left[x_t^e - \eta_\gamma\left(\tilde{y}_t^L - \tilde{y}_t^{e,L}\right)\right] - \beta\eta_\gamma E_t\left(x_{t+1}^* - \eta_\gamma x_t^*\right), \tag{3.2}$$

$$x_t^e = \tilde{y}_t - \tilde{y}_t^e \tag{3.3}$$

$$\tilde{y}_t^e = -\frac{\varphi_\gamma}{\omega}\left\{\tilde{y}_t^e - \eta_\gamma\left(\tilde{y}_t^{e,L} - \gamma_t\right) - \beta\eta_\gamma E_t\left[\tilde{y}_{t+1}^e + \gamma_{t+1}^y - \eta_\gamma\tilde{y}_t^e\right]\right\} + \frac{\beta\eta_\gamma\omega^{-1}}{1-\beta\eta_\gamma}E_t\delta_{t+1}^y, \tag{3.4}$$

$$r_t^e = E_t\gamma_{t+1}^y + E_t\delta_{t+1}^y - \omega\left(E_t\tilde{y}_{t+1}^e - \tilde{y}_t^e\right), \tag{3.5}$$

$$\tilde{\pi}_t = \xi\left(\omega x_t^e + \varphi\tilde{x}_t\right) + \beta E_t\tilde{\pi}_{t+1} + u_t, \tag{3.6}$$

$$\tilde{\pi}_t = \pi_t - \zeta\pi_{t-1}, \tag{3.7}$$

$$i_t = \rho i_{t-1} + (1-\rho)\left(\phi_\pi\pi_t + \phi_x x_t^e\right) + r_t, \tag{3.8}$$

$$\tilde{y}_t^{e,L} = \tilde{y}_{t-1}^e, \tag{3.9}$$

$$\tilde{y}_t^L = \tilde{y}_{t-1}, \tag{3.10}$$

$$\Delta y_t = \tilde{y}_t - \tilde{y}_{t-1} + \gamma_t, \tag{3.11}$$

$$\gamma_t^y = \gamma_t, \tag{3.12}$$

$$\delta_t^y = \delta_t. \tag{3.13}$$

Observation equations:

$$\Delta Y_t^a = \gamma^a + 400\Delta y_t \tag{3.14}$$

$$\pi_t^a = \pi^* + 400\pi_t \tag{3.15}$$

$$i_t^a = (r^a + \pi^*) + 400 i_t, \tag{3.16}$$

where ΔY_t^a is the growth rate of real GDP, π_t^a is the inflation rate, i_t^a is the Federal Funds target rate.

Observable variables:

$$x_t \equiv \left(\Delta Y_t^a, \pi_t^a, i_t^a\right).$$

Endogenous variables:

$$y_t \equiv \left(\tilde{y}_t, \tilde{\pi}_t, i_t, \tilde{x}_t, x_t^e, x_t^*, \tilde{y}_t^e, r_t^e, \Delta y_t, \tilde{y}_t^L, \tilde{y}_t^{e,L}, \gamma_t^y, \delta_t^y\right).$$

Exogenous disturbances:

$$s_t \equiv \left(\gamma_t, \delta_t, u_t, r_t\right).$$

Innovations:

$$e_t \equiv \left(e_t^\gamma, e_t^\delta, e_t^u, e_t^r\right).$$

Economic parameters:

$$\varepsilon \equiv (\omega, \xi, \eta, \zeta, \rho, \phi_\pi, \phi_x, \pi^*, r^a, \gamma^a).$$

(Other parameters showing up in equations are either calibrated or some combination of estimated parameters.)

Statistical parameters: depends on the assumption,

$$\begin{aligned}
\sigma_{AR1} &\equiv (\Phi_\gamma, \Phi_\delta, \Phi_u, \Phi_r, \Omega_\gamma, \Omega_\delta, \Omega_u, \Omega_r), \text{ or} \\
\sigma_{VAR1} &\equiv (\Phi_{\gamma\gamma}, \Phi_{\gamma\delta}, \Phi_{\gamma u}, \Phi_{\gamma r}, \Phi_{\delta\gamma}, \Phi_{\delta\delta}, \Phi_{\delta u}, \Phi_{\delta r}, \Phi_{u\gamma}, \Phi_{u\delta}, \Phi_{uu}, \Phi_{ur}, \Phi_{r\gamma}, \Phi_{r\delta}, \Phi_{ru}, \Phi_{rr}, \\
&\quad \Omega_{\gamma\gamma}, \Omega_{\gamma\delta}, \Omega_{\gamma u}, \Omega_{\gamma r}, \Omega_{\delta\delta}, \Omega_{\delta u}, \Omega_{\delta r}, \Omega_{uu}, \Omega_{ur}, \Omega_{rr}).
\end{aligned}$$

Equations (3.1)-(3.13) satisfy the canonical form (1.8), hence the solution to this model can be represented as in (1.10), consistent with our paper's framework. (Notice that, as before, in this model we defined several auxiliary variables that allow us to map the model exactly to the framework needed in our paper.) The number of observables, n_x, is smaller than the number of disturbances, n_s, hence again it is possible to apply the Kalman filter and generate a well behaved likelihood function. With proper priors defined, then assumption 1a) in the paper holds. Because certainty equivalence applies to all models that satisfy the canonical form (1.8), then assumption 1b) also holds. Assumption 1c) is also satisfied by construction, just as previously discussed.

4 Medium Scale New Keynesian DSGE Model

The fourth, and last, model considered is the one discussed in the section 4 of the paper, based on Smets and Wouters (2007). This is an example of a medium scale NK-DSGE model, incorporating everything as in the previous one and, in addition, capital, investment adjustment costs, and wage nominal rigidities.

Log-linear equations:

$$\tilde{y}_t = (0.82 - i_y)\, c_t + i_y i_t + R_*^k k_y z_t + s_t^g, \tag{4.1}$$

$$c_t = c_1 c_t^L + (1 - c_1) E_t c_{t+1} + c_2 (l_t - E_t l_{t+1}) - c_3 (r_t - E_t \pi_{t+1} + s_t^b), \tag{4.2}$$

$$i_t = i_1 i_t^L + (1 - i_1) E_t i_{t+1} + i_2 q_t + s_t^i, \tag{4.3}$$

$$q_t = q_1 E_t q_{t+1} + (1 - q_1) E_t \left(l_{t+1} - k_{t+1} + w_{t+1}\right) - (r_t - E_t \pi_{t+1} + s_t^b), \tag{4.4}$$

$$\tilde{y}_t = \phi \left[\alpha k_{t-1} + \alpha z_t + (1 - \alpha) l_t + s_t^a\right], \tag{4.5}$$

$$z_t = [(1-\psi)/\psi](l_t - k_t + w_t), \tag{4.6}$$
$$k_t = k_1 k_{t-1} + (1-k_1) i_t + k_2 s_t^i, \tag{4.7}$$
$$\pi_t = \pi_1 \pi_t^L + \pi_2 E_t \pi_{t+1} - \pi_3 \mu_t^p + s_t^p, \tag{4.8}$$
$$w_t = w_1 w_t^L + (1-w_1)(E_t w_{t+1} + E_t \pi_{t+1}) - w_2 \pi_t + w_3 \pi_t^L - w_4 \mu_t^w + s_t^w, \tag{4.9}$$
$$\mu_t^p = \alpha(k_{t-1} + z_t - l_t) - w_t + s_t^a, \tag{4.10}$$
$$\mu_t^w = w_t - [\sigma_l l_t + (c_t - c_{t-1}\lambda/\gamma)/(1-\lambda/\gamma)], \tag{4.11}$$
$$r_t = \rho r_{t-1} + (1-\rho)[r_\pi \pi_t + r_y(\tilde{y}_t - \tilde{y}_t^p)] + r_{\Delta y}(\Delta y_t - \Delta y_t^p) + s_t^r, \tag{4.12}$$
$$\tilde{y}_t^p = (0.82 - i_y) c_t^p + i_y i_t^p + R_*^k k_y z_t^p + s_t^g, \tag{4.13}$$
$$c_t^p = c_1 c_t^{pL} + (1-c_1) E_t c_{t+1}^p + c_2(l_t^p - E_t l_{t+1}^p) - c_3(r_t^{rp} + s_t^b), \tag{4.14}$$
$$i_t^p = i_1 i_t^{pL} + (1-i_1) E_t i_{t+1}^p + i_2 q_t^p + s_t^i, \tag{4.15}$$
$$q_t^p = q_1 E_t q_{t+1}^p + (1-q_1) E_t \left(l_{t+1}^p - k_{t+1}^p + w_{t+1}^p\right) - (r_t^{rp} + s_t^b), \tag{4.16}$$
$$\tilde{y}_t^p = \phi\left[\alpha k_{t-1}^p + \alpha z_t^p + (1-\alpha) l_t^p + s_t^a\right], \tag{4.17}$$
$$z_t^p = [(1-\psi)/\psi](l_t^p - k_t^p + w_t^p), \tag{4.18}$$
$$k_t^p = k_1 k_{t-1}^p + (1-k_1) i_t^p + k_2 s_t^i, \tag{4.19}$$
$$w_t^p = \alpha(k_{t-1}^p + z_t^p - l_t^p) + s_t^a, \tag{4.20}$$
$$w_t^p = \left[\sigma_l l_t^p + \left(c_t^p - c_{t-1}^p \lambda/\gamma\right)/(1-\lambda/\gamma)\right], \tag{4.21}$$
$$c_t^L = c_{t-1}, \tag{4.22}$$
$$i_t^L = i_{t-1}, \tag{4.23}$$
$$\pi_t^L = \pi_{t-1}, \tag{4.24}$$
$$w_t^L = w_{t-1}, \tag{4.25}$$
$$c_t^{pL} = c_{t-1}^p, \tag{4.26}$$
$$i_t^{pL} = i_{t-1}^p, \tag{4.27}$$
$$\Delta y_t = \tilde{y}_t - \tilde{y}_{t-1}, \tag{4.28}$$
$$\Delta y_t^p = \tilde{y}_t^p - \tilde{y}_{t-1}^p, \tag{4.29}$$
$$\Delta c_t = c_t - c_{t-1}, \tag{4.30}$$
$$\Delta i_t = i_t - i_{t-1}, \tag{4.31}$$
$$\Delta w_t = w_t - w_{t-1}. \tag{4.32}$$

Observation equations:

$$\begin{aligned}
\Delta y_t^{obs} &= \bar{\gamma} + \Delta y_t, & (4.33)\\
\Delta c_t^{obs} &= \bar{\gamma} + \Delta c_t, & (4.34)\\
\Delta i_t^{obs} &= \bar{\gamma} + \Delta i_t, & (4.35)\\
\Delta w_t^{obs} &= \bar{\gamma} + \Delta w_t, & (4.36)\\
l_t^{obs} &= \bar{l} + l_t & (4.37)\\
\pi_t^{obs} &= \bar{\pi} + \pi_t & (4.38)\\
r_t^{obs} &= \bar{r} + r_t & (4.39)
\end{aligned}$$

where $\left(\Delta Y_t^a, \Delta c_t^{obs}, \Delta i_t^{obs}, \Delta w_t^{obs}\right)$ represent the real growth rates of GDP, consumption, investment and wages, l_t^{obs} hours worked, π_t^{obs} inflation rate, and r_t^{obs} the federal funds rate.

Observable variables:

$$x_t \equiv \left(\Delta Y_t^a, \Delta c_t^{obs}, \Delta i_t^{obs}, \Delta w_t^{obs}, l_t^{obs}, \pi_t^{obs}, r_t^{obs}\right).$$

Endogenous variables:

$$\begin{aligned}
y_t \equiv & \left(\tilde{y}_t, c_t, i_t, z_t, l_t, r_t, q_t, k_t, w_t, \pi_t, \mu_t^p, \mu_t^w, \tilde{y}_t^p, c_t^p, i_t^p, z_t^p, l_t^p, r_t^{rp}, i_t^p, q_t^p, z_t^p, k_t^p, w_t^p,\right.\\
& \left. c_t^L, i_t^L, \pi_t^L, w_t^L, c_t^{pL}, i_t^{pL}, \Delta y_t, \Delta y_t^p, \Delta i_t, \Delta c_t, \Delta w_t\right).
\end{aligned}$$

Exogenous disturbances:

$$s_t \equiv \left(s_t^g, s_t^b, s_t^i, s_t^a, s_t^p, s_t^w, s_t^r\right).$$

Innovations:

$$e_t \equiv \left(e_t^g, e_t^b, e_t^i, e_t^a, e_t^p, e_t^w, e_t^r\right).$$

Economic parameters:

$$\varepsilon \equiv \left(\bar{\gamma}, \bar{l}, \bar{\pi}, 100\left(\beta^{-1} - 1\right), \varphi, \sigma_c, \lambda, \xi_w, \sigma_l, \xi_p, \iota_w, \iota_p, \psi, \Phi_{sw}, r_\pi, \rho_{sw}, r_y, r_{\Delta y}, \alpha\right).$$

(Other parameters showing up in equations are either calibrated or some combination of estimated parameters.)

Statistical parameters: depends on the assumption,

$$\begin{aligned}
\sigma_{AR1} \equiv & \left(\Phi_g, \Phi_b, \Phi_i, \Phi_a, \Phi_p, \Phi_w, \Phi_r, \Omega_g, \Omega_b, \Omega_i, \Omega_a, \Omega_p, \Omega_w, \Omega_r\right), \text{ or}\\
\sigma_{VAR1^*} \equiv & \left(\Phi_{gg}, \Phi_{gb}, \Phi_{gi}, \Phi_{ga}, \Phi_{gp}, \Phi_{gw}, \Phi_{gr}, ...\Phi_{rg}, \Phi_{rb}, \Phi_{ri}, \Phi_{ra}, \Phi_{rp}, \Phi_{rw}, \Phi_{rr},\right.\\
& \left.\Omega_g, \Omega_b, \Omega_i, \Omega_a, \Omega_p, \Omega_w, \Omega_r\right),
\end{aligned}$$

where σ_{VAR1^*} stands for the dynamic VAR(1) — dynamic correlated shocks, but contemporaneously independent — which is the case discussed in our paper (the full list of parameters is shown in Table A.2. of the paper).

Equations (4.1)-(4.32) satisfy the canonical form (1.8), hence the solution to this model can be represented as in (1.10), consistent with our paper's framework. (Notice that, as before, in this model we defined several auxiliary variables that allow us to map the model exactly to the framework needed in our paper.) The number of observables, n_x, is equal to the number of disturbances, n_s, hence again it is possible to apply the Kalman filter and generate a well behaved likelihood function. With proper priors defined, then assumption 1a) in the paper holds. Because certainty equivalence applies to all models that satisfy the canonical form (1.8), then assumption 1b) also holds. Assumption 1c) is also satisfied by construction, just as previously discussed.

References

An, S., and F. Schorfheide (2006): "Bayesian Analysis of DSGE Models," *Econometric Reviews, forthcoming.*

Christiano, L., R. Motto, and M. Rostagno (2009): "Financial Factors in Economic Fluctuations," *Unpublished.*

Christiano, L. J., and M. Eichenbaum (1992): "Liquidity Effects and the Monetary Transmission Mechanism," *American Economic Review*, 82(2), 346–353.

Cúrdia, V., A. Ferrero, and A. Tambalotti (2010): "Evaluating Interest Rate Rules in an Estimated DSGE Model," .

Hamilton, J. D. (1994): *Time Series Analysis.* Princeton University Press, Princeton, New Jersey.

Lubik, T. A., and F. Schorfheide (2004): "Testing for Indeterminacy: An Application to U.S. Monetary Policy," *American Economic R*, 94(1), 190–217.

Sbordone, A. M., A. Tambalotti, K. Rao, and K. Walsh (2010): "Policy Analysis Using DSGE Models: An Introduction," *Federal Reserve Bank of New York Economic Policy Review*, 16(2), 23–44.

Sims, C. A. (2002): "Solving Linear Rational Expectations Models," *Computational Economics*, 20(1-2), 1–20.

Smets, F., and R. Wouters (2007): "Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach," *American Economic Review*, 97(3), 586–606.