Correlated Disturbances and U.S. Business Cycles*

Vasco Curdia
Federal Reserve Bank of New York

Ricardo Reis
Columbia University

June 2009

Abstract

The dynamic stochastic general equilibrium (DSGE) models 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 to make the estimation of DSGE models with correlated disturbances feasible and quick. Our second contribution is a re-examination 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 a key missing ingredient in the models is countercyclical fiscal policy. According to our estimates, government spending and technology disturbances play a larger role in the business cycle than previously ascribed, while changes in markups become much less important. (*JEL* E30, E10)

*We are grateful to Craig Burnside, Lawrence Christiano, Martin Eichenbaum, Alejandro Justiniano, Jonathan Parker, Giorgio Primiceri, Juan Rubio-Ramirez, 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.

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 argues that these assumptions are unwarranted, 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 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, since it has been 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 incorrectly-identified disturbances.[1] However, DSGE macroeconometric models routinely assume that disturbances are independent first-order autoregressions, AR(1)s. While this makes interpretation and estimation easier, and even feasible given the limits of existing algorithms in dealing with many nuisance parameters, it is still arbitrary and potentially dangerous for inference.

In this paper, we develop new Bayesian econometric techniques to incorporate correlated disturbances in dynamics macroeconomic models. We show that the economic structure of the models implies that key conditional posterior distributions belong 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 estimates. Our algorithm significantly speeds up estimation with independent AR(1)s. More importantly, because the parameters associated with the disturbances are part of the conjugate conditional distributions, it allows for estimation of DSGE models with correlated disturbances

[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.

that were previously prohibitively numerically costly.

We envision two 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 heteroskedasticity and autocorrelation in the disturbances (Stock and Watson, 2007). 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 more freely on the dimensions along which the model is inadequate. Finding a strong correlation between different elements of the disturbance vector highlights the ways in which the endogenous part of the model is failing to match the data, and suggests the path to building future models that endogenize these correlations.

The second contribution of this paper is to the study of 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 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 1 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 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, four long-standing puzzles for full-information estimates of this model.

Section 2 present the estimation method more generally. We show that the conjugate conditionals arise in a broad class of equilibrium macroeconomic models, and discuss a few ways to exploit the knowledge of this known slice of the posterior distribution in making inferences.

Section 3 focuses on a richer business-cycle model, due to Smets and Wouters (2007). Allowing for correlated disturbances does not significantly improve the fit of the model, nor does it affect many of its predictions on the impact of policy changes in the economy. Still, it does change inferences along two dimensions. First, disturbances to wage markups are now much less important sources of business cycles, being replaced by productivity and government spending as key drivers. Second, 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 4 concludes with a brief review of the main results.

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, whereas 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.

A few papers have moved beyond the assumption of independent AR(1) disturbances, but typically in only special ways. 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. In the open-economy literature, it is common to assume that productivity and monetary disturbances are correlated across countries, and recently Rabanal, Ramirez and Tuesta (2008) allow for cointegration among disturbances.[2] We take a step further allowing for a much richer correlation between disturbances.

Del Negro and Schorfheide (forthcoming) also emphasize the need for robustifying inferences from DSGEs. They 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, and contrast this with allowing for flexible processes for the disturbances as we do. As they note, our approach is a special case that 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 are on judiciously modelling some correlations instead of the flexible and more general approach that we propose. They emphasize the difficult issue of policy invariance, while we are more worried with the positive properties of the models.

Finally, our paper fits into a burgeoning literature extending the ability to estimate more general DSGE models. Justiniano and Primiceri (2008a) allow for time-varying volatility,

[2] This literature is too large to do full justice to here, but see Lubik and Schorfheide (2007) for an early estimated open-economy DSGE model and Justiniano and Preston (2008) for a recent one.

Ramirez and Villaverde (2007) consider non-normal innovations, and Binsbergen, Koijen, Ramirez and Villaverde (2008) deal with recursive non-expected utility preferences. Our methods are complementary to these. Chib and Ramamurthy (2008) propose a multiple-block Metropolis-Hasting approach to DSGE estimation, with some resemblances with ours. A key difference is that 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.

Simplicity, identification and orthogonalization

An objection to allowing for correlated disturbances is that it makes their structural interpretation harder. 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 easier to interpret than those from a VAR, few (if any) researcher would argue in favor of the former instead of the latter, given the estimation bias that this apparent simplicity entails. Moreover, as the application in this paper shows, it is possible to interpret estimates with correlated disturbances and once this is done, what becomes hard to understand is what was captured by estimates that assumed, for instance, that government spending was exogenous. Finally, 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.

A 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. 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. In all of the applications of this paper, we exhaustively insured against this danger, both by checking that the information matrix of each model we estimated had full column rank, as well as by looking at the rank of the Hessian of the posterior distribution at many randomly drawn points. In the future, 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.

Finally, 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. 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. Anyway, the particular methods and results in this paper do not depend on which stand one takes on identification and orthogonalization more generally.

1 Correlated disturbances in a canonical DSGE model

The best-known and simple DSGE model is due to Prescott (1986), and we extend it to include government spending following Baxter and King (1992) 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.

1.1 The model of fluctuations

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

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

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

subject to

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

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

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).

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. 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. 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:

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

[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.

[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.

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 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 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 will model disturbances 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.

One argument for assuming uncorrelated AR(1)s is that it reduces the number of parameters. Letting σ denote 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.

1.2 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:

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

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

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

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:

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

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

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 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.

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 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 $p(\sigma^{(j)} | x^T, \varepsilon^{(j-1)})$ followed by drawing $\varepsilon^{(j)}$ from the conditional $p(\varepsilon^{(j)} | x^T, \sigma^{(j)})$. 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 $p(\sigma | x^T, \varepsilon) = \int p(\sigma, s^T | x^T, \varepsilon) ds^T$, 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 $p(\sigma, s^T | x^T, \varepsilon)$ can be split by Gibbs sampling again into drawing from $p(s^T | x^T, \varepsilon, \sigma)$ and $p(\sigma | x^T, \varepsilon, s^T)$ 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 $p(s^T | x^T, \varepsilon, \sigma)$ is normal with mean and variance given by a disturbance smoother filter. Moreover, conditional on the disturbances s^T, equation (8) is a standard vector autoregression, for which there are conjugate priors. In particular, if the prior distribution for $(\Phi(L), \Omega)$ is normal-inverse-Wishart, then the posterior distribution for the statistical parameters is also normal-inverse-Wishart with analytical expressions for its moments.

Combining these three observations provides our algorithm. It draws from the expanded parameter vector $\left(\varepsilon, \sigma, s^T\right)$ in turn, exploiting the knowledge that two of the three needed conditional distributions are known analytically. 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 σ is known analytically, 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]

1.3 Data, priors, and 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 has identified and partly remedied its weaknesses, we do not want to take the estimates too seriously. 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

[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.

2008:2, although we use the data before 1960:1 only to calibrate the 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 $\theta = 4.9$ (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 of the correlated-disturbance terms is 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, and the marginal likelihood was higher for order 6, but 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.

Our first set of results address the efficiency of the conjugate-conditionals algorithm versus the Metropolis random-walk. 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.

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, 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

[8] Section 2 discusses these priors in more detail as well as alternatives within the conjugate-conditionals family.

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 means test, *SPM*, of Geweke (1992) and fewer rejections implies being closer to convergence.

Figure 1 shows the results.[9] In the horizontal axes 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, 300 for $neff$, or 2000 for $mneff$, are reached earlier. Since in this case, disturbances are uncorrelated and the number of statistical parameters is small, these figures provide 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.

1.4 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

[9] 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 20% 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 10 Monte Carlo simulations.

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 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 hour, against most VAR studies (e.g., Shapiro and Watson, 1986), which attribute a larger role to productivity.[10]

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, 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

[10]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.

former. 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.

Let us 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 0.42 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.[11] In line with the VAR evidence, productivity now accounts for three quarters of the variance of output and 64% of the variance of hours.[12]

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

[11] 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.

[12] These results identify the impulse reponses 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. Likewise, the results are robust to the order of the VAR.

fiscal policy.

2 The general theory of the conjugate-conditionals method

This section starts by laying out the set of models to which the method applies. It then presents the two conjugate conditional distributions, followed by the block Metropolis algorithm that exploits them. Finally, we discuss alternative priors and complications that may arise from restrictions on the statistical parameters.

2.1 The model

We consider an economic model that relates the following vectors of variables:

- y_t : endogenous economic variables, of dimension n_y;
- s_t : exogenous disturbances, of dimension n_s;
- e_t : exogenous innovations to the disturbances, of dimension n_s;
- x_t : observables, of dimension n_x;
- 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, and $q(.)$ to denote a prior distribution.

As first noted by Blanchard and Kahn (1980), the dynamics of many linear (or linearized) economic models are described by a system of equations:

$$\Psi_0(\varepsilon)y_t = \Psi_1(\varepsilon)y_{t-1} + \Psi_2(\varepsilon)s_t + \Psi_3(\varepsilon)w_t, \tag{10}$$

together with a set of boundary conditions, coming from initial states and transversality conditions. The vector of endogenous disturbances w_t has the property that $E_{t-1}w_t = 0$ and can capture terms involving $E_t(y_{t+1})$ (Sims, 2002). The Ψ_i matrices typically have many zero elements and their size is higher than n_ε, embodying the cross-equation restrictions that come from optimal behavior, technologies and other constraints.[13]

The disturbances, s_t, are assumed to be linear covariance-stationary processes that are well approximated by a vector autoregression of finite order k:

$$s_t = \Phi(\sigma)(L)\, s_{t-1} + e_t, \text{ with } e_t \sim N(0, \Omega(\sigma)) \tag{11}$$

where $\Phi(\sigma)(L) = \sum_{j=1}^{k} \Phi_j(\sigma) L^j$, a matrix lag polynomial. If these matrices are left unrestricted, then the number of statistical parameter $n_\sigma = kn_s + n_s(n_s+1)/2$, which could be quite large. If there are restrictions on the disturbance processes, n_σ may be smaller and its components may have to lie in restricted sets.

The reduced-form solution to the system in equations (10)-(11) is:

$$\begin{pmatrix} y_t \\ \tilde{s}_t \end{pmatrix} = \begin{pmatrix} \Lambda_1(\varepsilon,\sigma) & \Lambda_2(\varepsilon,\sigma) \\ 0 & \tilde{\Phi}(\sigma) \end{pmatrix} \begin{pmatrix} y_{t-1} \\ \tilde{s}_{t-1} \end{pmatrix} + \begin{pmatrix} \Lambda_3(\varepsilon,\sigma) \\ \Upsilon \end{pmatrix} e_t. \tag{12}$$

The first row has the rational expectations solution of the model, where the coefficient matrices Λ_i are typically complicated non-linear functions of the economic and shock parameters.[14] The second row re-writes the VAR in state-space form, where $\tilde{s}_t = \left(s'_t, ..., s'_{t-k+1}\right)'$, the large matrix $\tilde{\Phi}(\sigma)$ includes all the elements of $\Phi(\sigma)(L)$, and Υ is a selection matrix of zeros and ones. Letting $z_t \equiv (y'_t, \tilde{s}'_t)'$ be the state vector, we write this more compactly as:

$$z_t = \Gamma_1(\varepsilon,\sigma) z_{t-1} + \Gamma_2(\varepsilon,\sigma) e_t \tag{13}$$

The data is for a vector of variables x_t, which is a linear combination of the elements of z_t that may or may not involve the economic parameters:

$$x_t = H(\varepsilon)\, z_t. \tag{14}$$

We consider the usual case where the matrix H does not have full rank, so we cannot directly observe z_t (otherwise typically there will not be an estimation problem). We abstract from

[13] We will treat y_t as deviations from a steady-state so we ommit constants from the system, but it is straightforward to include these.

[14] By the principle of certainty equivalence, the Λ_i might depend on Ψ_i and Φ_j, but will not depend on Ω.

measurement error in these observations to avoid confusion with the economic disturbances specified in the model, but including measurement error does not change our conclusions significantly.

Given this setup, the estimation problem is the following: we observe x^T and we know its dynamics are determined by the state-space system in equations (13)-(14). Starting from a prior distribution $q(\varepsilon, \sigma)$, the goal is to characterize the posterior $p(\varepsilon, \sigma | x^T)$, and this is done numerically by simulating J draws. Note that this is a different problem than the one in unrestricted state-space estimation because the economic theory puts tight constraints on how the parameters in (ε, σ) affect $\Gamma_1(\varepsilon, \sigma)$ and $\Gamma_2(\varepsilon, \sigma)$, and because the $\tilde{s}_t$ component of the z_t vector involves only σ, not ε. It is this last feature that leads to the conjugate-conditionals method.

2.2 The conjugate conditional distributions

Our first result characterizes the conditional distribution for the statistical parameters.

Proposition 1 *A sufficient statistic for the conditional* $p\left(\sigma | x^T, \varepsilon, z^T\right)$ *is* s^T *and, since the elements of* σ *are* $\Phi(L)$ *and* Ω*, from the definition of a conditional probability:*

$$p\left(\sigma | x^T, \varepsilon, z^T\right) = p(\tilde{\Phi} \Big| s^T, \Omega) p(\Omega | s^T). \tag{15}$$

Then, if $q(\Omega)$ *is an inverse-Wishart with* ν_0 *degrees of freedom and scale matrix* Ξ_0*, while* $\breve{\Phi} = [\Phi_1, ..., \Phi_k]'$ *and* $\bar{\Phi}$ *vectorizes it,* $\bar{\Phi} = [vec(\Phi_1')', ..., vec(\Phi_k')']'$*, so that* $\bar{\Phi}$ *is a* $n_s^2 k$ *vector with mean* Φ_0 *and variance* $\Omega \otimes \Theta_0$*, then the posterior distributions belong to the same families, with posterior parameters:*

$$\begin{aligned}
\nu_1 &= \nu_0 + T - k && (16)\\
\Xi_1 &= \Xi_0 + S + \breve{\Phi}_0' \Theta_0^{-1} \breve{\Phi}_0 + \breve{\Phi}_{OLS}' X' X \breve{\Phi}_{OLS} - \breve{\Phi}_1' \Theta_1^{-1} \breve{\Phi}_1 && (17)\\
\breve{\Phi}_1 &= \Theta_1 \Theta_0^{-1} \breve{\Phi}_0 + \Theta_1 X' X \breve{\Phi}_{OLS} && (18)\\
\Theta_1 &= \left(\Theta_0^{-1} + X'X\right)^{-1} && (19)
\end{aligned}$$

where the VAR is written as $Y = X\breve{\Phi} + E$*, with* Y *a* $(T-k)xn_s$ *matrix,* X *a* $(T-k)xn_s k$

matrix and $\breve{\Phi}_{OLS} = (X'X)^{-1}X'Y$ *and* $S = \left(Y - X\breve{\Phi}_{OLS}\right)'\left(Y - X\breve{\Phi}'_{OLS}\right)$.

The first part in the proposition notes that once we know z^T and extract s^T, then the second block of the reduced-form state-space system is simply a VAR. The second part shows that the family of normal-inverse-Wishart priors are conjugate for this problem (Kadiyala and Karlsson, 1997).

If the disturbances were observed in the data, this first proposition would suffice to characterize the statistical parameters.[15] Even without observing s^T, Proposition 1 is useful once combined with a second result. The distribution for the disturbances, conditional on all the parameters and the data $p(z^T | x^T, \varepsilon, \sigma)$, is a standard filtering problem on the state-space system in equations (13)-(14). This has a well-known (e.g., Kim and Nelson, 1999, Durbin and Koopman, 2001) analytical solution:

Proposition 2 *The state vector realizations* z^T *are normally distributed, conditionally on the observations* x^T *and the parameters* (ε, σ). *Their density can be factored into a product of normal densities:*

$$p\left(z^T \middle| x^T, \varepsilon, \sigma\right) = f\left(z_T \middle| x^T, \varepsilon, \sigma\right) \prod_{t=1}^{T-1} f\left(z_t \middle| z_{t+1}, x^t, \varepsilon, \sigma\right). \tag{20}$$

with means and covariances denoted by:

$$\begin{aligned} z_{t|t} &\equiv E(z_t | x^t, \varepsilon, \sigma), & P_{t|t} &\equiv E\left(\left(z_t - z_{t|t}\right)\left(z_t - z_{t|t}\right)' \middle| x^t, \varepsilon, \sigma\right), \\ z_{t|t+1} &\equiv E(z_t | z_{t+1}, x^t, \varepsilon, \sigma), & P_{t|t+1} &\equiv E\left(\left(z_t - z_{t|t+1}\right)\left(z_t - z_{t|t+1}\right)' \middle| z_{t+1}, x^t, \varepsilon, \sigma\right), \end{aligned} \tag{21}$$

[15] To characterize the full posterior distribution, in a first step, we could take J draws for σ quickly and easily even for large n_σ using the result in Proposition 1, and then use these to numerically characterize the unknown conditional $p(\varepsilon | x^T, \sigma)$ for the smaller number n_ε of economic parameters . To give one example where this would work, consider estimating a Calvo New Keynesian Phillips curve:

$$\pi_t = \beta E(\pi_{t+1}) + \chi^{-1}(1-\chi)(1-\beta\chi)mc_t + \nu_t,$$

where π_t is endogenous inflation, mc_t is the observed exogenous marginal cost, ν_t is a measurement error, and the two economic parameters are the discount factor, β, and the fraction of fixed prices, χ. The disturbance in this model is mc_t, so allow it to follow an AR(k), with $k+1$ parameters. The conventional Metropolis random-walk algorithm would sample from the $k+3$ parameters in one block. Our method would instead first obtain J draws for the $k+1$ statistical parameters quickly from the known conditional for an AR(k) fit to the observed data on mc_t. Then it would use each of these draws as regressors in the Phillips curve to characterize the distribution of β and χ.

These means and covariances are given by the Kalman filter recursions:

$$\begin{aligned} z_{t|t-1} &= \Gamma_1 z_{t-1|t-1}, & x_{t|t-1} &= H z_{t|t-1}, \\ P_{t|t-1} &= \Gamma_1 P_{t-1|t-1} \Gamma_1' + \Gamma_2 \Omega \Gamma_2', & K_t &= P_{t|t-1} H' \left(H P_{t|t-1} H' \right)^{-1}, \\ z_{t|t} &= z_{t|t-1} + K_t \left(x_t - x_{t|t-1} \right), & P_{t|t} &= \left(I - K_t H \right) P_{t|t-1}. \end{aligned} \tag{22}$$

followed by a second set of recursions for the disturbance smoother:

$$M_t = P_{t|t} \Gamma_1 P_{t+1|t}^{-1}, \quad z_{t|t+1} = z_{t|t} + M_t \left(z_{t+1} - \Gamma_1 z_{t|t} \right), \quad P_{t|t+1} = P_{t|t} - M_t P_{t+1|t} M_t' \tag{23}$$

Sampling from $p\left(z^T \middle| x^T, \varepsilon, \sigma \right)$ is an easy matter, as even for very large n_z, most software programs can take draws from the multivariate normal quickly. Moreover, while the Kalman filter recursions can take some time, they were required anyway in order to calculate the likelihood function. Given a draw for z^T, we can read off directly the draw for the state vector s^T or its extended counterpart $\tilde{s}^T$, and apply the result in Proposition 1.

2.3 Hybrid algorithms to exploit the conjugate conditionals

We now describe algorithms, sometimes referred to as hybrid, Metropolis-within-Gibbs, or block-Metropolis, that exploit both propositions. These algorithms rely on the following application of a fundamental result (Tierney, 1994).

Proposition 3 *A Markov Chain Monte Carlo that at step j starts with $(\varepsilon^{(j-1)}, \sigma^{(j-1)}, z^{T(j-1)})$ and then (i) draws $z^{T(j)}$ from the conditional, $p(z^{T(j)} | x^T, \varepsilon^{(j-1)}, \sigma^{(j-1)})$ in proposition 2; (ii) draws $\sigma^{(j)}$ from the conditional $p(\sigma^{(j)} | x^T, \varepsilon^{(j-1)}, z^{T(j)})$ in proposition 1, and (iii) draws $\varepsilon^{(j)}$ from the conditional $p(\varepsilon^{(j)} | x^T, \sigma^{(j)}, z^{T(j)})$, will produce a set of J draws that converge to draws from the posterior distribution $p(\varepsilon, \sigma, z^T | x^T)$.*

If we knew the conditional distribution for the economic parameters, then this would be a straightforward Gibbs algorithm. However, there is little hope that any distribution involving the economic parameters has an analytical form given that they are complicated non-linear functions of the elements of the state-space representation. We must instead draw from a proposal distribution that approximates $p(\varepsilon | x^T, \sigma, z^T)$. There are a few approaches to doing this:

1) The random-walk Metropolis. Starting with a draw $\varepsilon^{(j-1)}$, at step j this (i) draws a candidate ε^{*} from a proposal density, $\pi(\varepsilon|\,\varepsilon^{(j-1)})$, (ii) computes the ratio $r = p\left(\varepsilon^{*}|\,x^{T},\sigma,z^{T}\right)$ $/p\left(\varepsilon^{(j-1)}\big|\,x^{T},\sigma,z^{T}\right)$, (iii) with probability $\min(r,1)$ sets $\varepsilon^{(j)} = \varepsilon^{*}$ and otherwise sets $\varepsilon^{(j-1)} = \varepsilon^{*}$. In practice, we followed convention and used for proposal a normal distribution with mean $\varepsilon^{(j-1)}$ and covariance $c\mathcal{H}^{-1}$ where $\mathcal{H}$ is the Hessian at the mode of the posterior for the conditional distribution for ε , and c is a scaling factor adjusted to control the rate of convergence. As for the ratio r, using Bayes theorem,

$$r = \frac{\mathcal{L}\left(x^{T}|\varepsilon^{*},\sigma^{(j)},z^{T(j)}\right)q(\varepsilon^{*})}{\mathcal{L}\left(x^{T}|\varepsilon^{(j-1)},\sigma^{(j)},z^{T(j)}\right)q(\varepsilon^{(j-1)})}, \tag{24}$$

where $\mathcal{L}\,(.)$ is the likelihood function and $q(.)$ the prior.

2) The independent Metropolis. The algorithm is identical to the previous one with the exception of the proposal density $\pi(.)$, which is now independent of $\varepsilon^{(j-1)}$. One strategy that often leads to good results looks for local maxima of the posterior distribution. If there is only one, pick $\pi(.)$ to be a normal centered at the mode with covariance $c\mathcal{H}^{-1}$ where c is a factor well above 1. If there are many, pick $\pi(.)$ to be a mixture of normal centered at these modes, with mixing weights proportional to the mode's posterior densities.

3) Rejection sampling. In this algorithm, we pick an auxiliary density $\pi(\varepsilon)$, from which we sample a candidate ε^{*}, which is accepted with probability $\min(r',1)$ where

$$r' = \mathcal{L}\left(x^{T}|\varepsilon^{*},\sigma^{(j)},z^{T(j)}\right)q(\varepsilon^{*})/k\pi(\varepsilon^{*}). \tag{25}$$

The scalar k is set to ensure that r' is close to 1, and a typical choice for the proposal density is a normal centered at the mode, or a mixture of normals around local maxima.

We have not found that any of these algorithms clearly dominates the others. It seems to depend on the model and the data. Finally, in our experience, it improved inference considerably to importance re-sample the J draws for (ε,σ), that is to draw from the original set of J draws, with probabilities proportional to their posterior density.

2.4 Alternative priors and parameter restrictions

Proposition 1 used a particular choice of prior distributions. There are alternatives, especially if we want to impose restrictions on the VAR matrices for the coefficients Φ_j in equation (11), as in the case of independent AR(1)s. We now summarize some of the alternative conjugate families.

If we assume that disturbances follow independent AR(k)s, then the Φ_j and Ω matrices are all diagonal. In that case, a family of conjugate priors and posteriors is: $i = 1, ..., n_s$ independent normals for $[\Phi_j(i)]_{j=1}^k$, and $i = 1, ...n_s$ independent inverse-gammas squared for each of $\Omega(i)$. The means and variances of the prior distributions can be chosen freely.

A second possibility 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 proposition 1, and the independent inverse gamma priors for $\Omega(i)$ from the previous paragraph, we have a conjugate family.

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, the system in equation (11) 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; for instance, with an unrestricted VAR, the $\bar{s}$ are stacked along the diagonal blocks of the Z matrix. In this general SUR case, as long as $\beta|\,\Omega$ is normal and Ω^{-1} has a Wishart distribution, then so will the posteriors (Zellner, 1962).

There are alternative conjugate priors to the normal-inverse-Wishart priors in proposition 1. 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 proposition 1 and has some computational advantages. The priors and posteriors for the covariance matrix Ω stop being conjugate.

Finally, in some models, the transversality conditions impose the constraint that the

VAR in equation (11) is stationary. There is no conjugate family for both Φ and Ω in this case, but there is a conjugate family for Φ conditional on Ω, the truncated normal in the region of stationarity. In this case, the conditional distribution of Ω is not known and must be characterized together with the economic parameters using proposal distributions as in the previous sub-section. Still, we at least know the conditional distribution for Φ and since this is where the curse of dimensionality bites as k increases, even in this worst-scenario case, our conjugate-conditionals should offer some efficiency improvements.

3 Monetary business cycles with correlated disturbances

Smets and Wouters (2007) proposed a new Keynesian model of monetary policy and business cycles with a variety of frictions, including sticky prices and wages, habits in consumption, and investment adjustments costs. They found that this monetary business-cycle (MBC) model fit the U.S. data on seven series—output, consumption, investment, hours worked, real wages, inflation and nominal interest rate—quite well and, in a slightly different version, the Euro-area data as well (Smets and Wouters, 2003). Versions of this model are being adopted by central banks around the world.[16]

The model has 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)—and, following the DSGE tradition, these are assumed to follow independent AR(1)s. Smets and Wouters (2007) made two exceptions to the AR(1) restriction. First, they included two first-order moving average terms for price and wage markup disturbances, in order to fit high-frequency movements in the data. Second, they allowed for contemporaneously correlated disturbances between government spending and total factor productivity.

We compare two treatments of the disturbances in this MBC model, using the same data and priors as Smets and Wouters (2007). In the first case, we impose the restriction that disturbances are independent. To make the contrast clearer, we eliminate the contemporaneous correlation between productivity and government spending; this is the only

[16] Del Negro et al (2008) document more exhaustively the empirical strengths and weaknesses of the model.

difference between this model and the one estimated by Smets and Wouters (2007). In the second case, we allow for dynamically correlated disturbances according to a VAR(1). We still restrict the disturbances to be contemporaneously independent, because any orthogonalization would be controversial with seven innovations. We also maintain the moving average terms, again to maintain the comparison limited to independent versus correlated disturbances.

The full set of estimates and impulse responses is reported in the appendix, and we focus here on three main lessons. The first one is that the responses of output, hours and inflation to productivity, fiscal spending, and monetary policy are qualitatively similar in the two cases. These are plotted in figure 3. The inferences on the response of the economy to policy changes, one of the main uses of this model, are therefore robust to correlated disturbances. This is reflected also in the log marginal predictive densities, that are within 5 log points with independent or correlated disturbances. Part of this is due to the many sources of endogenous dynamics in the model, and another part is due to the imprecisions with which many of these moments are estimated, so that they are not distinguishable across the two cases.

The second lesson is that, in spite of the close fit of the two models, one inference changes significantly once we allow for correlated disturbances. Table 3 shows the median variance decompositions for outpur, hours, real wages and inflation at three horizons: 1-quarter, 1-year, and infinity.[17] 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 but, as Smets and Wouters (2007) emphasized, it is the wage-markup disturbance that dominates at longer horizons. With correlated disturbances, the short-run inference is similar but the long-run one changes dramatically. Wage markups are much less important, and are even close to irrelevant for output. Productivity and government spending now dominate the variance of output. Likewise, for real wages and inflation, wage markups dominate the long-run with independent disturbances, but with correlated disturbances, they are unimportant for wages and less than half as important for inflation. Productivity and especially now government spending replace the wage-markup

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

as dominant source of fluctuations.

A third lesson comes from looking at which of the nuisance parameters in the correlation structure are significantly different from zero.[18] There are two distinct sets for which this is the case. 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. As in the case of the simple RBC model, it shows that also in this more involved MBC model, the model is still missing an important role for fiscal policy rules. 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. Models with financial imperfections seem particularly promising.[19]

To conclude, allowing for correlated disturbances in the MBC model confirmed many of the lessons taken from previous estimations of this model. There are two changes though, that again highlight the need to allow for correlated disturbances to robustify inference and to point the direction of future research. We found that the much debated finding that markup disturbances are important is not robust.[20] The role of wage markup s is much reduced for all variables and becomes insignificant for output and wages beyond a few quarters. As with the simple model, the results 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.

4 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

[18] The full posterior distribution is in the appendix.

[19] For recent DSGE models with financial imperfections see Curdia and Woodford (2009) and Christiano, Trabandt and Wallentin (2009).

[20] See Chari, Kehoe, McGrattan (2009) and Justiniano and Primiceri (2008b) for some of the debate.

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. Moreover, our method pointed to endogenously modelling risk premia and investment-specific productivity disturbances as the most promising avenue to bringing this model closer to the data.

Appendix

Proof of the propositions. The three propositions use standard results in the literature; the novelty in this paper is in combining them to extract relevant conditional conjugates. For the expressions in proposition 1, see for instance Kadiyala and Karlsson (1997) or Geweke (2005).

For proposition 2, the Kalman filter formulae are standard in most time-series textbooks. The second set of recursions, on the disturbance smoother, is less common. It follows from noting that $z_{t|t+1} = E(z_t|\, z_{t+1}, z_{t|t})$ since $z_{t|t}$ captures all of the relevant information about z_t from the data. Since $z_{t+1} = \Gamma_1 z_t + \Gamma_2 e_{t+1}$, using the standard formulae for a linear projection:

$$z_{t|t+1} = z_{t|t} + P_{t|t}\Gamma_1' P_{t+1|t}^{-1}\left(z_{t+1} - \Gamma_1 z_{t|t}\right), \tag{26}$$

which using the definition of M_t leads to the recursion for $z_{t|t}$. Similar steps lead to the recursion for $P_{t|t}$.

Finally, the result in proposition 3 comes from noting that this is a particular block Metropolis algorithm, so the general result of Tierney (1994) applies. See Chib (2001) for a careful exposition.

The monetary-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 the statistical parameters to fit our general setup in section 2. 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:

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

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

Baxter, Marianne and Robert G. King (1993). "Fiscal Policy in General Equilibrium." *American Economic Review*, 83 (3), 315-34.

Basu Susanto, John G. Fernald and Miles S. Kimball (2006). "Are Technology Improvements Contractionary?" *American Economic Review*, 96 (5), 1418-48.

Binsbergen, Jules H. van, Ralph S.J. Koijen, Juan F. Rubio-Ramirez and Jesus Fernandez-Villaverde (2008). "Likelihood Estimation of DSGE Models with Epstein-Zin Preferences." Manuscript.

Blanchard, Olivier J. and Charles M. Kahn (1980). "The Solution of Linear Difference Models under Rational Expectations." *Econometrica*, 48 (5), 1305-11.

Chang, Yongsung and Sun-Bin Kim (2007). "Heterogeneity and Aggregation: Implications for Labor-Market Fluctuations." *American Economic Review*, 97 (5), 1939-56.

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. 3569-649.

Chib, Siddartha and Srikanth Ramamurthy (2008). "MCMC Methods for Bayesian Estimation of DSGE Models." Manuscript.

Christiano, Lawrence J. and Martin Eichenbaum (1992). "Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations." *American Economic Review*, 82 (3), 430-50.

Christiano, Lawrence J., Mathias Trabandt and Karl Walentin (2009). "Introducing Financial Frictions and Unemployment in a Small Open Economy Model." Working paper.

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.

Curdia, 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), 512-49.

Del Negro, Marco and Frank Schorfheide, forthcoming. "Monetary Policy Analysis with Potentially Misspecified Models. " *American Economic Review*.

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.

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), 1379-99.

Gali, Jordi (1999). "Technology, Employment, and the Business Cycle: Do Technology Shocks Explain Aggregate Fluctuations?" *American Economic Review*, 89 (1), 249-71.

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 (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 Econometrcis 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), 342-62.

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), S223-50.

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 (2008). "Can Structural Small Open Economy Models Account for the Influence of Foreign Shocks?" NBER working paper 14,547.

Justiniano, Alejandro and Giorgio E. Primiceri (2008a). "The Time-Varying Volatility of Macroeconomic Fluctuations." *American Economic Review*, 98 (3), 604-41.

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

Lubik, Thomas A. and Frank Schorfheide (2007). "Do Central Banks Respond to Exchange Rate Movements? A Structural Investigation." *Journal of Monetary Economics*, 54 (4), 1069-1087.

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

Kim, Chang-Jin and Cahrles R. Nelson (1999). *State-Space Models with Regime Switching.* MIT Press: Cambridge, MA.

Newey, Whitney K. and Kenneth D. West (1987). "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica*, 55 (3), 703-08.

Parkin, Michael (1988). "A Method for Determining Whether Parameters in Aggregative Models Are Structural." *Carnegie-Rochester Conference Series on Public Policy*, 29, 212-52.

Rabanal, Pau, Juan F. Rubio-Ramirez and Vicente Tuesta (2008). "Cointegrated TFP Shocks and International Business Cycles." Manuscript.

Ramey, Valerie A. (2008). "Identifying Government Spending Shocks: It's All in the Timing." Manuscript.

Ramirez, Juan F. Rubio and Jesus Fernandez-Villaverde (2007). "Estimating Macroeconomic Models: a Likelihood Approach." *Review of Economic Studies*, 74, 1059-87.

Reis, Ricardo (2008). "Using VARs to Identify Models of Fiscal Policy: A Comment on Perotti." *NBER Macoeconomics Annual 2007*, 22, 227-36.

Rogerson, Richard and Johanna Wallenius (2007). "Micro and Macro Elasticities in a Life cycle Model with Taxes." NBER working paper 13017.

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), 1009-44.

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

Sims, Christopher A. (1980). "Macroeconomics and Reality." *Econometrica*, 48 (1), 1-48.

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), 949-68.

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. and Mark W. Watson (2007) *Econometrics.* Addison-Wesley.

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), 361-71.

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 Tests for Aggregation Bias." *Journal of the American Statistical Association*, 57, 348-68.

Table 1. Prior Distributions for RBC model

Parameter	*Density*[a]	*Mode*	*Percentile* 5	50	95
Economic					
γ	*G*	0.6667	0.2327	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	.00066	.00015	.00050
Ω_G	IG^2	0.2297	0.1452	0.4002	1.7917
Panel B. Unrestricted VAR(1)					
Φ_{AA}	*N*	.00008	.00005	.00007	.00009
Φ_{AG}	*N*	0.0000	-0.0041	0.0000	0.0041
Φ_{GA}	*N*	0.0000	-13.960	-0.0600	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	0.1012	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

			Percentile		
Parameter	*Mean*	*Mode*	*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	0.4229	0.4304	0.2860	0.4147	0.5789
θ	4.9287	4.3184	2.0964	4.6568	8.6639
Statistical					
Φ_{AA}	0.9391	0.9355	0.9056	0.9410	0.9664
Φ_{AG}	0.0048	0.0048	0.0041	0.0049	0.0054
Φ_{GA}	-8.62	-8.26	-11.33	-8.49	-6.23
Φ_{GG}	0.8803	0.8828	0.8371	0.8806	0.9228
Ω_{AA}	.00013	.00013	.00011	.00013	.00016
Ω_{AG}	0.0085	0.0071	0.0045	0.0080	0.0141
Ω_{GG}	2.1402	1.3527	0.7858	1.7179	4.9552

Table 3. Variance Decompositions in MBC 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	0.116	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
1-year ahead							
Output	0.101	0.150	0.266	0.217	0.116	0.072	0.053
Hours	0.252	0.120	0.230	0.164	0.088	0.039	0.084
Real wage	0.041	0.021	0.000	0.032	0.041	0.276	0.573
Inflation	0.049	0.007	0.002	0.019	0.040	0.524	0.345
Unconditional							
Output	0.133	0.014	0.207	0.043	0.019	0.020	0.490
Hours	0.046	0.015	0.257	0.041	0.019	0.015	0.558
Real wage	0.379	0.010	0.001	0.073	0.040	0.173	0.279
Inflation	0.044	0.006	0.006	0.019	0.037	0.278	0.594
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	0.201
1-year ahead							
Output	0.033	0.358	0.122	0.127	0.023	0.127	0.160
Hours	0.298	0.234	0.116	0.121	0.009	0.043	0.137
Real wage	0.154	0.021	0.0161	0.007	0.025	0.303	0.421
Inflation	0.067	0.017	0.053	0.050	0.052	0.345	0.356
Unconditional							
Output	0.382	0.044	0.287	0.129	0.005	0.011	0.094
Hours	0.178	0.090	0.095	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

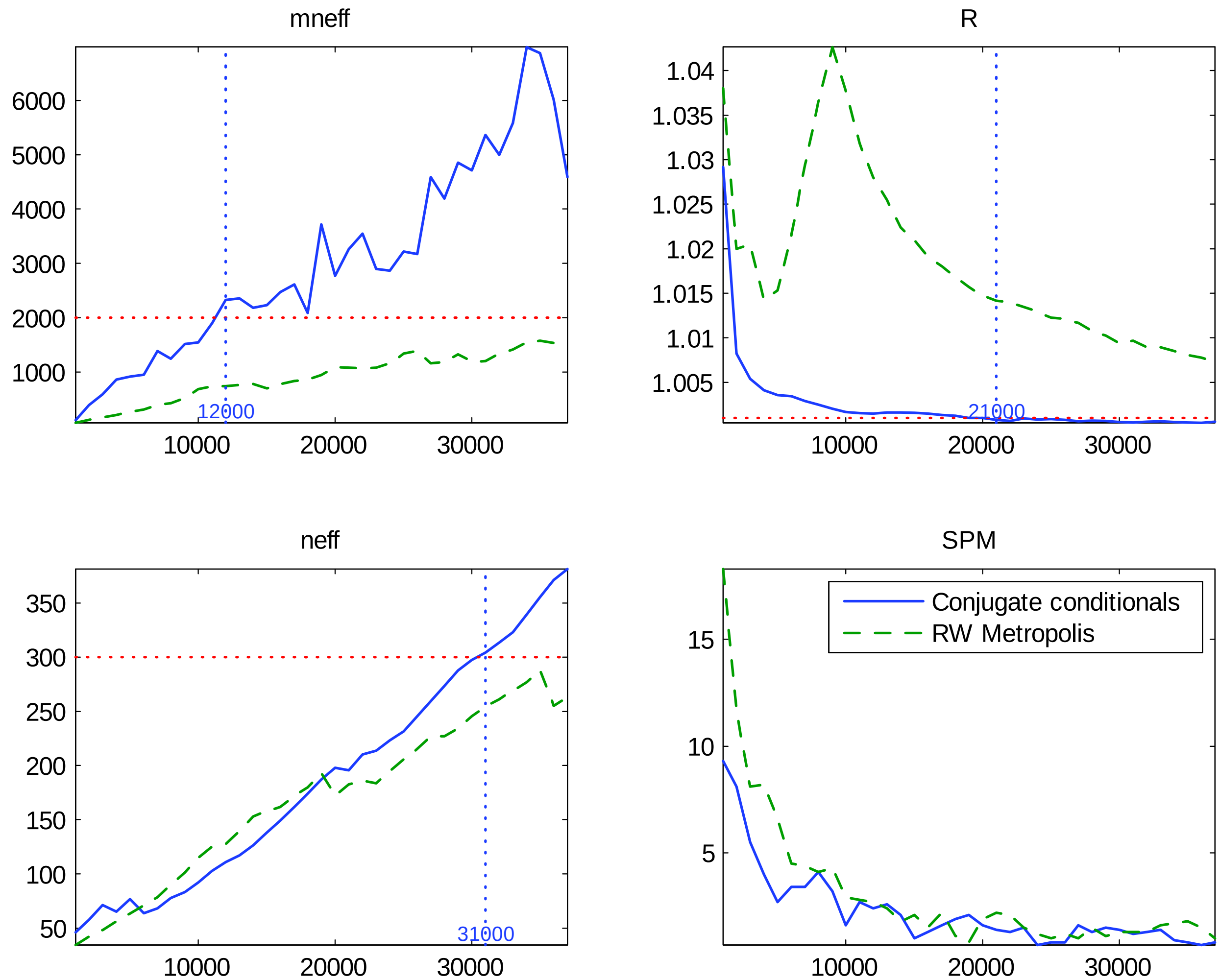
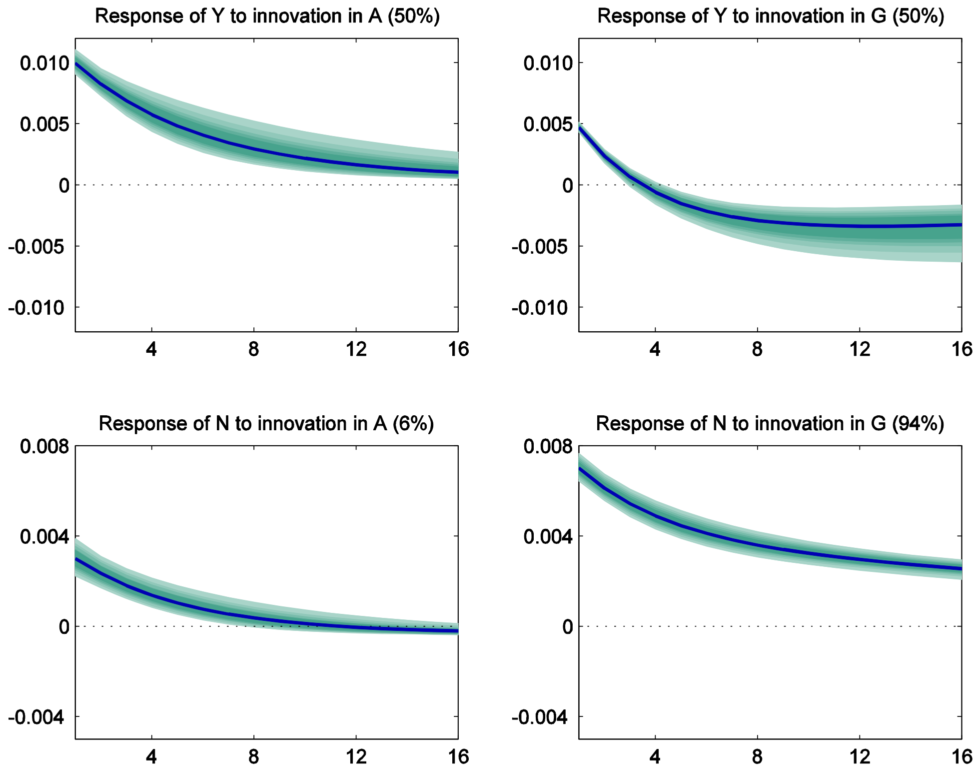


Figure 2. Impulse response functions in RBC model, median and distributions

Panel A. Independent AR(1)s case



Panel B. Unrestricted VAR(1)

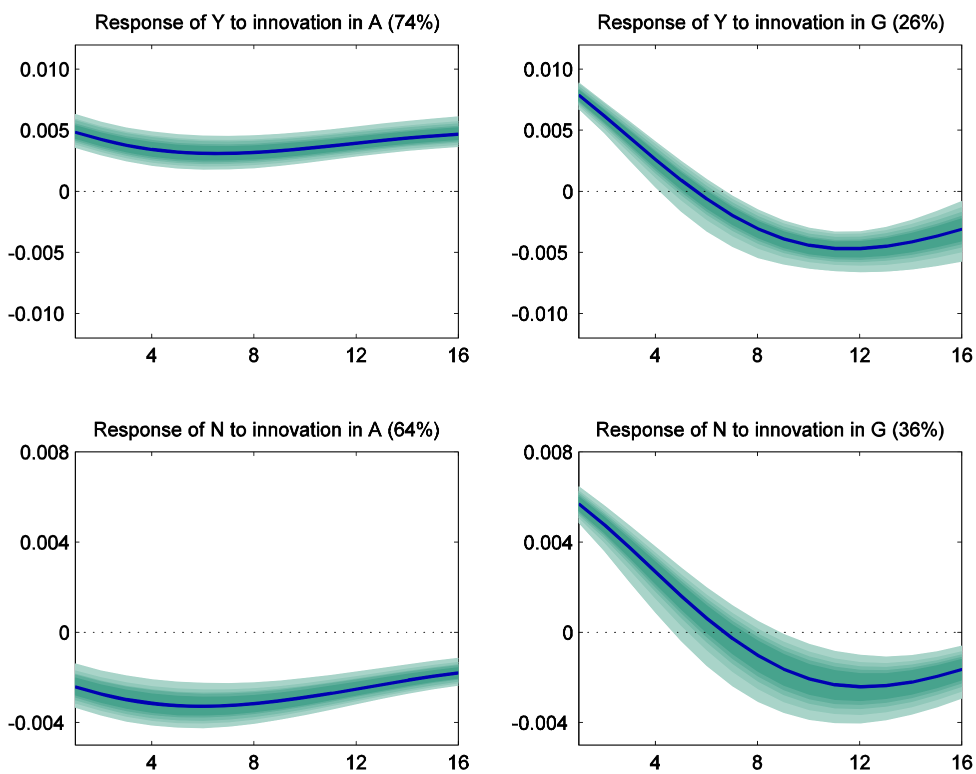


Figure 3. Median impulse response functions in MBC model, with independent and correlated disturbances

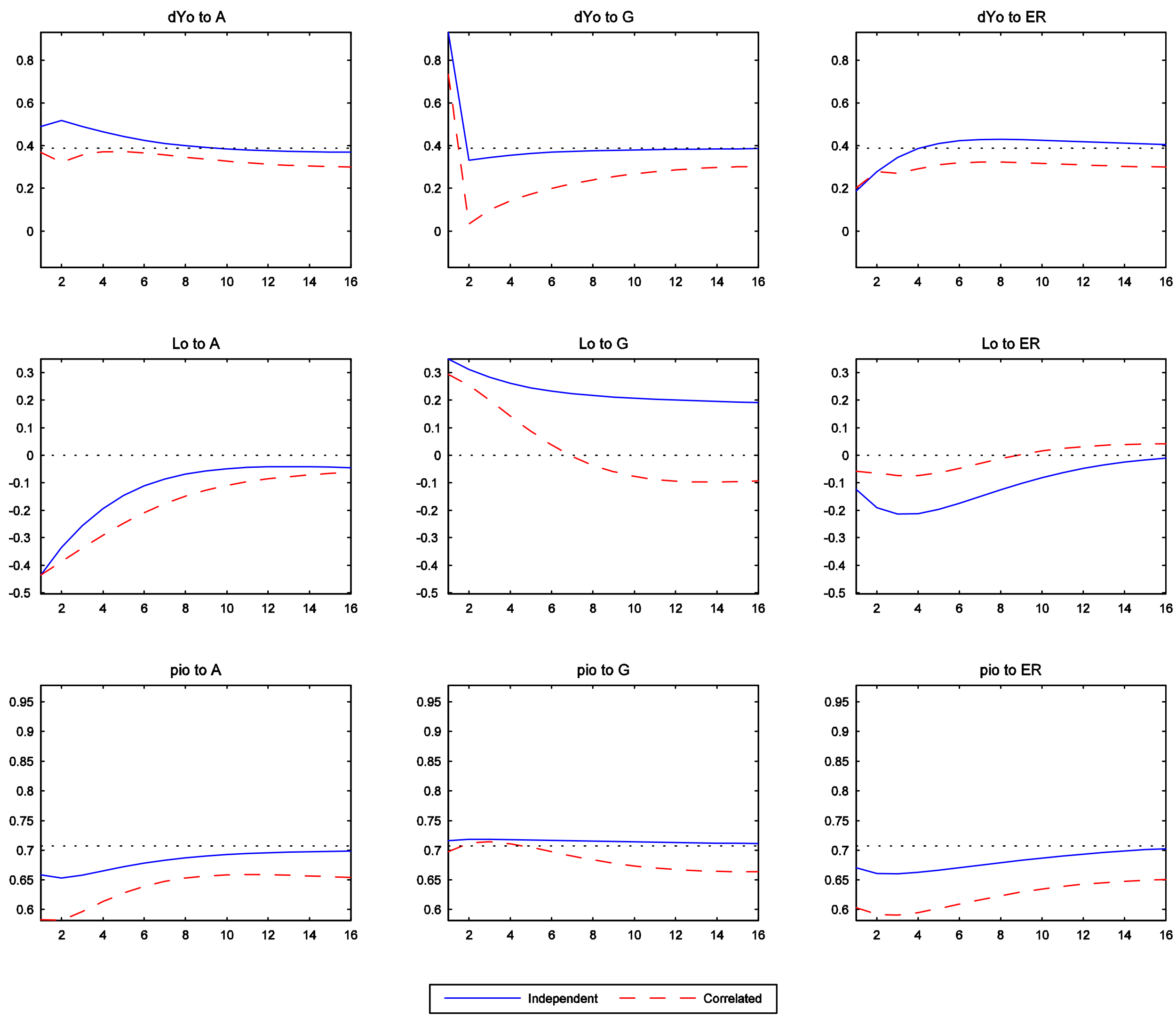


Table A.1. Prior and posterior distribution for MBC model, independent AR(1) disturbances

	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

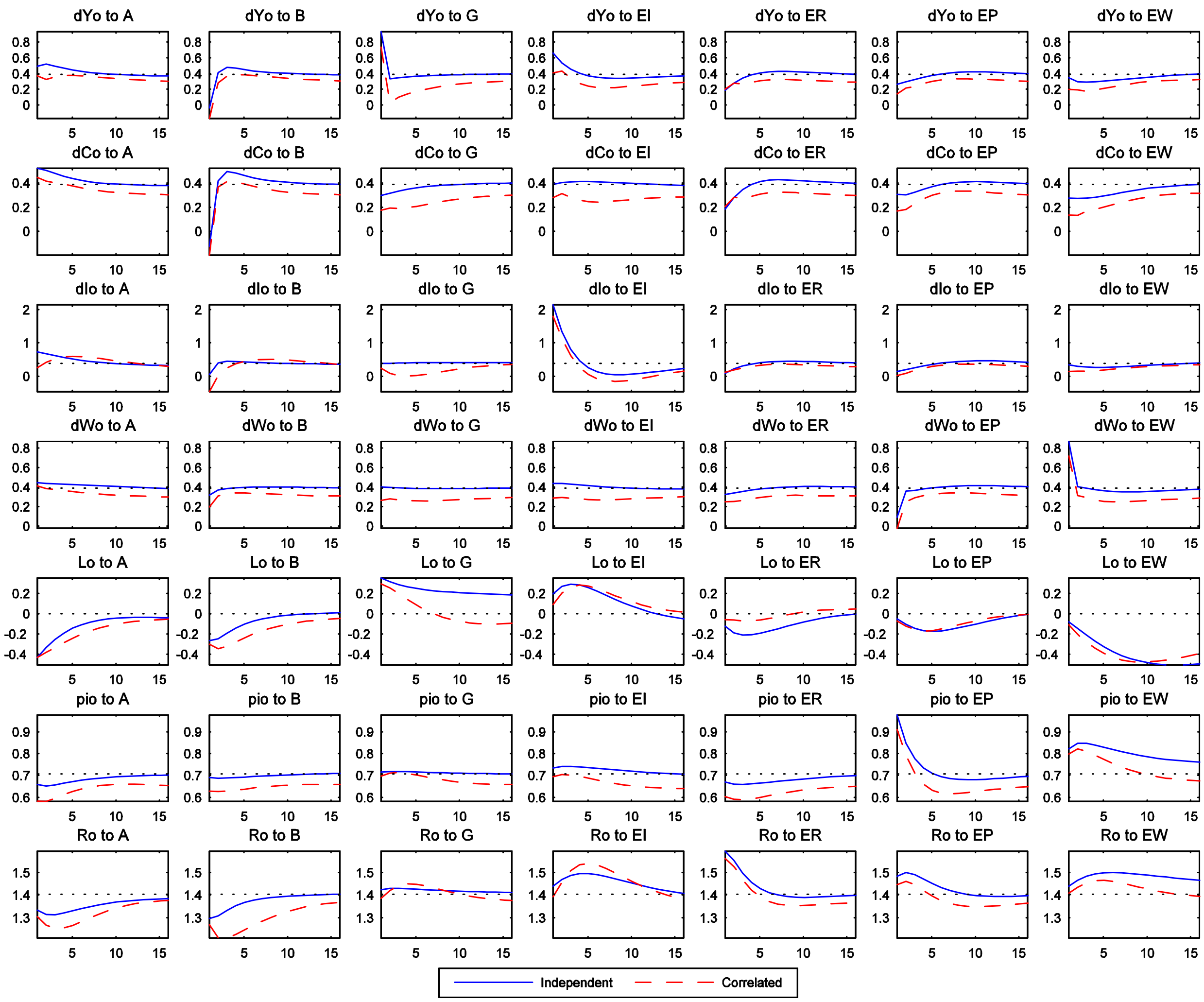
	Prior				Posterior					
	Dist	5%	Median	95%	Mode	Mean	SE	5%	Median	95%
γ^*	N	0.2355	0.4000	0.5645	0.2753	0.2964	0.0201	0.2623	0.2976	0.3271
l^*	N	-0.4935	0.0000	0.4935	-0.0000	-0.0001	0.2976	-0.4911	-0.0003	0.4878
π^*	G	0.4652	0.6146	0.7931	0.6090	0.6559	0.1024	0.4937	0.6525	0.8295
$100(\beta^{-1}-1))$	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 in MBC model, 5% and 95% points 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	.398, .545	.072, .168	.043, .100	.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, .340	.584, .757
Inflation	.011, .051	.001, .012	.000, .003	.001, .023	.005, .033	.679, .907	.069, .222
1-year ahead							
Output	.054, .169	.102, .242	.189, .350	.133, .307	.074, .172	.050, .103	.015, .127
Hours	.176, .330	.08, .212	.181, .284	.110, .230	.055, .135	.023, .063	.037, .165
Real wage	.010, .111	.006, .054	.000, .002	.014, .057	.017, .079	.200, .373	.442, .682
Inflation	.026, .082	.002, .025	.001, .006	.004, .055	.018, .076	.382, .678	.237, .457
Unconditional							
Output	.048, .278	.005, .031	.051, .642	.012, .106	.005, .051	.006, .048	.201, .761
Hours	.018, .088	.005, .035	.076, .674	.012, .097	.006, .048	.005, .035	.226, .829
Real wage	.130, .708	.004, .024	.000, .005	.025, .157	.013, .084	.084, .305	.108, .511
Inflation	.018, .081	.002, .018	.002, .014	.004, .062	.015, .075	.146, .405	.429, .787
Panel B. Dynamic VAR(1) disturbances							
1-quarter ahead							
Output	.000, .068	.354, .561	.258, .488	.003, .076	.001, .060	.021, .114	.005, .073
Hours	.384, .547	.172, .303	.152, .275	.003, .049	.000, .031	.003, .042	.014, .075
Real wage	.014, .102	.006, .093	.000, .026	.000, .015	.000, .032	.219, .440	.420, .695
Inflation	.021, .104	.000, .066	.001, .062	.001, .075	.001, .087	.482, .790	.111, .308
1-year ahead							
Output	.003, .130	.204, .529	.077, .180	.014, .269	.001, .105	.043, .249	.062, .309
Hours	.182, .407	.119, .380	.053, .198	.044, .226	.001, .060	.006, .117	.057, .291
Real wage	.060, .284	.005, .109	.000, .091	.000, .057	.001, .109	.159, .467	.218, .619
Inflation	.024, .132	.001, .125	.006, .141	.002, .155	.003, .166	.215, .514	.225, .497
Unconditional							
Output	.188, .552	.010, .126	.162, .432	.030, .254	.001, .031	.002, .048	.022, .309
Hours	.076, .392	.028, .225	.030, .251	.037, .252	.003, .085	.007, .104	.155, .635
Real wage	.309, .608	.005, .134	.135, .398	.038, .264	.000, .028	.004, .071	.008, .125
Inflation	.073, .432	.008, .115	.059, .321	.033, .203	.006, .116	.047, .267	.079, .436

Figure A.1. Median impulse response functions in MBC 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).