Generalized Exogenous Processes in DSGE: A Bayesian Approach [*]

Alexander Meyer-Gohde[§] Daniel Neuhoff [†]

This Version: May 14, 2014

Abstract

We introduce a Reversible Jump Markov Chain Monte Carlo (RJMCMC) method to Bayesian DSGE estimation. The method enables us to sample from a posterior distribution spanning nonnested models with parameter spaces of different dimensionality. We use the method to jointly sample from an ARMA process of unknown order along with the associated parameters. We apply the method to a canonical neoclassical growth model using post war US GDP data and find that the posterior decisively rejects the standard AR(1) assumption for the mean reverting technology process in favor of higher order processes. While the posterior contains significant uncertainty regarding the exact order, it provides tight posterior credible sets over the impulse responses. At the same time, the responses are hump-shaped. A negative response of hours to a positive technology shock is within the posterior credible set when non-invertible MA representations are considered.

JEL classification: C11; C32; C51; C52

Keywords: Bayesian analysis; Dynamic stochastic general equilibrium model; Model evaluation; ARMA; Reversible Jump Markov Chain Monte Carlo

[*]We are grateful to Michael Burda, as well as to participants of research seminars at the HU Berlin for useful comments, suggestions, and discussions. This research was supported by the DFG through the SFB 649 "Economic Risk." Any and all errors are entirely our own.

[§]Humboldt-Universität zu Berlin, Institut für Wirtschaftstheorie II, Spandauer Straße 1, 10178 Berlin, Germany; Tel.: +49-30-2093 5720; E-Mail: alexander.meyer-gohde@wiwi.hu-berlin.de

[†]Humboldt-Universität zu Berlin, Institut für Wirtschaftstheorie II, Spandauer Straße 1, 10178 Berlin, Germany; Tel.: +49-30-2093 5667; E-Mail:daniel.neuhoff@gmail.com

1 Introduction

Despite recent advances in improving the fit of DSGE models to the data, misspecification remains. In his Nobel Prize Lecture, Sims (2012) observes that "DSGEs could be made to fit better by adding parameters allowing more dynamics in the disturbances." Likewise, Del Negro and Schorfheide (2009) identify three approaches to deal with misspecification in rational expectations models: ignore it, generalize the stochastic driving forces, or relax the cross-equation restrictions. Apart from more or less ad hoc approaches, such as Smets and Wouters (2007) who have the price-markup disturbance follow an ARMA(1,1) process and Del Negro and Schorfheide (2009) who let government expenditures follow an AR(2) instead of an AR(1) process, the literature has not yet provided a systematic, Bayesian framework to address this approach to misspecification. We fill this gap by providing a fully Bayesian approach to estimating the order as well as the parameters of generalized (ARMA) representations of exogenous driving forces within DSGE models.

To accomplish the task, we adapt the Reversible Jump Markov Chain Monte Carlo (RJMCMC) methodology as pioneered by Green (1995).[1] RJMCMC enables the sampling from a posterior distribution spanning several, not necessarily nested, models with parameter spaces of different dimensionality. In our case, each model is identified by a specific set of orders for the lag polynomials of the autoregressive and moving average components of the disturbances, each leading to a different parameter space. This approach provides a framework for the systematic exploration of the fit of DSGE models using different structures for the shock processes which provides a computationally feasible alternative to estimating all different possible combinations of shock orders individually. Additionally, it allows us to quantify posterior model uncertainty and its consequences for impulse responses and correlation structures while being agnostic regarding the order of the underlying shock processes. Finally, if multiple shocks are kept independent while generalizing their individual autocorrelation patterns, the resulting estimations admit a structural interpretation of the shocks that can guide the researcher in identifying those dimensions along which the model requires the most additional internal propagation.[2]

[1] Markov Chain Monte Carlo methods have become increasingly popular for the estimation of DSGE models in recent years. See An and Schorfheide (2007) for a methodological review.

[2] It may furthermore be possible to construct model selection criteria based on the comparison of the spectrum of variables of interest derived from estimates of the posterior with the spectrum using only pure white noise shocks giving a measure of how much structure has to be added to the model outside of economic theory, an idea along the lines of Watson (1993).

The RJMCMC method rests on modifying the proposal ratios in the acceptance probability to circumvent the dimensionality mismatch induced by sampling for ARMA processes of different orders. We find that the RJMCMC provides point estimates of the ARMA orders with a reliability comparable to traditional order selection criteria such as AIC, AICC, and BIC. Yet, RJMCMC is of primary interest for its posterior distribution over different ARMA orders and not for its point estimates of the orders. After confirming that RJMCMC would correctly identify the ARMA order using synthetic data generated from the AR(1) technology process traditionally assumed in the specification of the neoclassical growth model, we turn to US post war GDP data and estimate the order and parameters of the technology process. We find that the data prefers higher order—at the mode, ARMA(3,0)—exogenous processes. The resulting impulse responses are hump-shaped, differing thus qualitatively from the responses to the traditional AR(1) process. From a DSGE likelihood perspective, there is, without a commensurate prior specification, no reason to prefer invertible or "fundamental" representations in the presence of MA terms; in sampling from the covariance equivalent representations for draws of the order with nonzero MA order, we find a downward shift in the amplitude of the impulse responses as well as an overall increase in the posterior uncertainty regarding the impulse responses of endogenous variables to a technology shock. Interestingly, we cannot exclude the possibility of a negative response of hours to a positive technology shock.

This paper is organized as follows: We first introduce our methodology and shortly illustrate the method by constructing a sampler for a univariate autoregressive model of unknown order. Afterwards, we present the results of a small Monte Carlo study designed to gauge the power of the method for identifying univariate autoregressive moving-average models using synthetic data. We then lay out the solution method for DSGE models with shock processes of unknown order and briefly discuss our approach to likelihood evaluation. Lastly, we apply the method to post war US data and the neoclassical growth model.

2 Reversible Jump MCMC for ARMA Processes

2.1 Reversible Jump Markov Chain Monte Carlo

In this paper, we adapt and apply the Reversible Jump Markov Chain Monte Carlo (RJMCMC) methodology pioneered by Green (1995). RJMCMC generalizes the Metropolis-Hastings algorithm (Hastings 1970) to allow for moves between parameter spaces of varying dimensionality while maintaining detailed balance. [3] This transdimensionality allows for inference on a posterior distribution spanning several, not necessarily nested, models. In the following, we will illustrate the mechanics of RJMCMC starting with a short description of conventional Metropolis-Hastings samplers to fix ideas before turning to the construction of a sampler for univariate autoregressive models of unknown order using an RJMCMC approach.[4]

2.2 Conventional Metropolis-Hastings Samplers

Markov Chain Monte Carlo (MCMC) methods in general provide samples from some probability distribution of interest by constructing a Markov chain whose stationary distribution is this distribution of interest. A Markov chain with the sequence of states $\varsigma_1, \varsigma_2, \ldots$ is specified in terms of the distribution for the initial state ς_1 and the transition kernel $K(\cdot)$ that provides the conditional distribution of a state ς_{i+1} given the current state ς_i. That is, the probability that ς_{i+1} is in some set $\mathcal{A} \subseteq \mathbb{R}^d$ given that the current state of the chain is ς_i is given by

$$K(\varsigma, \mathcal{A}) = P(\varsigma_{i+1} \in \mathcal{A} | \varsigma_i = \varsigma) \tag{1}$$

[3] A more extensive treatment of Metropolis-Hastings samplers can be found in Chib and Greenberg (1995). See also Tierney (1998) for a comparison of RJMCMC and conventional Metropolis-Hastings kernels. Another popular MCMC method is the Gibbs sampler which is a special case of Metropolis-Hastings samplers and ultimately RJMCMC samplers. See Gelfand and Smith (1990) for a review and comparison of Gibbs samplers as well as importance samplers and stochastic substition. A Gibbs sampler has been applied to the estimation of autoregressive models in Troughton and Godsill (1998). For applications of Bayesian methods to the estimation of DSGE models see e.g. Fernández-Villaverde and Rubio-Ramírez (2004) or the survey in An and Schorfheide (2007). Geweke (1998) provides an overview over Bayesian methods and their applications in economics.

[4] Several authors have applied RJMCMC to the problem of estimating univariate autoregressive (moving average) models, e.g., Brooks, Giudici, and Roberts (2003), Brooks and Ehlers (2004), and Ehlers and Brooks (2008).Relatedly, different approaches to statistical models of varying dimensionality have emerged; such as birth-death Markov Chain Monte Carlo, based on continuous time birth-death processes, as initiated by Stephens (2000) and applied to the analysis of autoregressive moving-average models by Philippe (2006). A summary and comparison of these methods can be found in Cappè, Robert, and Rydèn (2003).

A distribution π is invariant for some Markov chain if the transition kernel of the chain satisfies

$$\int K(\varsigma, \mathcal{A})\pi(\varsigma)d\varsigma = \int_{\mathcal{A}} \pi(\varsigma)d\varsigma \tag{2}$$

for all subsets $\mathcal{A}$ of the state space. The task in MCMC methods is to construct a kernel such that the distribution of interest π is invariant with respect to the Markov chain defined by $K()$. The expression in (2), however, is not practically useful for the construction of an appropriate kernel, as the verification of (2) would involve integration over the unknown distribution π being sought with the MCMC method.

One widely used approach to overcome this apparent hurdle are so-called Metropolis-Hastings samplers, as laid out first in Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller (1953) and generalized in Hastings (1970). Metropolis-Hastings samplers are accept-reject samplers for which proposals for a new state of the chain are drawn from some distribution γ to be chosen by the researcher and then accepted with an appropriately derived probability α. With Metropolis-Hastings samplers, the stronger condition of reversibility or detailed balance is imposed, which guarantees that π is invariant for the Markov chain. This condition holds if a sequence of two states $(\varsigma_i, \varsigma_{i+1})$ has the same distribution as the reversed subchain $(\varsigma_{i+1}, \varsigma_i)$ whenever $\varsigma_i, \varsigma_{i+1} \sim \pi$, i.e. if

$$\int_{\mathcal{A}} \pi(\varsigma) K(\varsigma, \mathcal{B}) d\varsigma = \int_{\mathcal{B}} \pi(\varsigma') K(\varsigma', \mathcal{A}) d\varsigma' \tag{3}$$

for all subsets $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^d$. Condition (3) is more easily verified and can thus provide a starting point for the construction of a sampler.

A general Metropolis-Hastings algorithm can be written as follows: Let again ς denote a state of the Markov chain, in the case of Bayesian inference in the context of model estimation, the state is just the vector of model parameters and the distribution of interest is the posterior distribution

$$\pi(\varsigma) \propto \mathcal{L}(\varsigma)\rho(\varsigma) \tag{4}$$

where ς denotes the vector of model parameters, $\mathcal{L}$ is the likelihood of the data given the model and its parameters and ρ is the prior over the model parameters. To obtain N samples from the posterior distribution, the following algorithm is run:

1. Set the (arbitrary) initial state ς_0 of the Markov chain
2. For $i = 1$ to N
   (a) Set $\varsigma = \varsigma_{i-1}$
   (b) Propose a new state from some proposal distribution $\gamma(\varsigma'|\varsigma)$

(c) Accept draw with probability

$$\alpha(\varsigma, \varsigma') = \min(1, \chi)$$

with

$$\chi = \underbrace{\frac{\mathcal{L}(\varsigma')}{\mathcal{L}(\varsigma)}}_{\text{Likelihood Ratio}} \times \underbrace{\frac{\rho(\varsigma')}{\rho(\varsigma)}}_{\text{Prior Ratio}} \times \underbrace{\frac{\gamma(\varsigma|\varsigma')}{\gamma(\varsigma'|\varsigma)}}_{\text{Proposal Ratio}}$$

(d) If the draw is accepted set $\varsigma_i = \varsigma'$. If the draw is rejected set $\varsigma_i = \varsigma$

This algorithm defines a transition kernel such that the Markov chain has the desired invariant distribution. The sequence of states of the chain is then a sample from this distribution of interest. The acceptance probability α corrects for differences between the proposal distribution γ and the distribution of interest.[5]

The kernel in the above is given by

$$K(\varsigma, \mathcal{B}) = \underbrace{\int_{\mathcal{B}} \gamma(\varsigma'|\varsigma)\alpha(\varsigma, \varsigma')d\varsigma'}_{\text{Probability of moving to set } \mathcal{B}} + \underbrace{\left[1 - \int_{\mathcal{B}} \gamma(\varsigma'|\varsigma)\alpha(\varsigma, \varsigma')d\varsigma'\right] \mathbb{1}_{\varsigma}}_{\text{Probability of rejecting the move and } \varsigma \in \mathcal{B}} \tag{5}$$

where $\mathbb{1}_{\varsigma} = 1$ if $\varsigma \in \mathcal{B}$ and zero otherwise giving the probability of moving to some subset $\mathcal{B}$ of the parameter space conditional on the chain currently being at ς. The crux when constructing the kernel is to define the appropriate acceptance probability α and the proposal distribution γ so as to satisfy the detailed balance condition and thereby guarantee the convergence of the Markov chain to the desired probability distribution. Indeed, plugging in the formulation of the kernel from (5) into (3) gives an expression from which, given the proposal distribution γ the appropriate acceptance probability α can be readily derived using Peskun's (1973) recipe.

2.3 Reversible Jump MCMC: AR(p) Order and Parameter Sampling

To fix ideas, we will derive our transdimensional random walk sampler implementation of the RJMCMC for a univariate zero-mean normally distributed AR(p) models of unknown order. Our derivation follows the exposition of Waagepetersen and Sorensen (2001). Such an AR(p) model is defined as

$$y_t = P_1^p y_{t-1} + P_2^p y_{t-2} + \ldots + P_p^p y_{t-p} + \epsilon_t, \quad \epsilon_t \sim \mathcal{N}\left(0, \sigma^2\right) \tag{6}$$

[5]Note, that in the case of a standard random walk Metropolis-Hastings sampler with symmetric proposals, i.e. a Metropolis sampler, the proposal ratio reduces to one.

P_i^p are the coefficients of the lag polynomial of order p associated with the i'th lag and ϵ_t is a zero-mean stochastic disturbance. Denote by $P^p \doteq \left\{P_1^p, P_2^p, \ldots, P_p^p\right\}$ the vector of parameters of the AR(p) model.[6] We would like to construct a posterior distribution over the orders, p, and associated parameters, P^p, given observations on y_t.

It is sensible to interpret the order of the lag polynomial p as a model indicator. We will use the terms model indicator and polynomial or lag order interchangeably. The aim is now to construct a sampler for the joint posterior distribution over the different models indexed by p and their parameters. The strategy closely resembles that for Metropolis-Hastings samplers. Indeed, Metropolis-Hastings samplers are a special case in the RJMCMC framework. For expositional purposes, it is convenient to express the state of the Markov chain that we will construct in the following manner

$$\varsigma = (p, P^p) \tag{7}$$

explicitly including the order of the autoregressive polynomial p in the state.

The detailed balance condition poses the main obstacle to the transdimensional sampling's construction of a joint posterior distribution over potentially nonnested models with parameter spaces of varying dimensionality. Recall the detailed balance condition (3),

$$\int_{\mathcal{A}} \pi(\varsigma) K(\varsigma, \mathcal{B}) \, d\varsigma = \int_{\mathcal{B}} \pi(\varsigma') K(\varsigma', \mathcal{A}) \, d\varsigma' \tag{8}$$

Unlike in the foregoing section, the dimension of ς can change. I.e., the state space of the Markov chain spans parameter spaces with differing dimensionality—for a sampler for AR(p) models of unknown order, when p changes so does the number of parameters. Here, the usual strategy for the derivation of the acceptance probability will fail. Green (1995) proposed modifying the proposals in such a way that the integrals on both sides of the detailed balance condition are over spaces of the same dimensionality by introducing an auxiliary proposal variable u together with a mapping $g_{pp'}$ that maps the auxiliary proposal u and the current state of the chain to the new proposed state. The mapping $g_{pp'}$ is chosen such that the dimensionality of the integrals on both sides of the equation is inflated to some higher common dimensionality.

In order to be able to easily verify adherence to detailed balance for a move from a state (p, P^p) to $(p', P^{p'})$ the vectors of Markov chain states and the random auxiliary proposal variables

[6]The standard deviation of the disturbance ϵ_t, σ, will be taken as given in the exposition of this section, to maintain the focus on the order, p, and the associated part of the parameter vectors is left implicit.

(P^p, u) and $(P^{p'}, u')$ must be of equal dimension. This dimension matching condition ensures that $\pi(P^p|p)\gamma_{pp'}(P^p, u)$ and $\pi(P^{p'}|p')\gamma_{p'p}(P^{p'}, u')$ are "joint densities on spaces of equal dimension," (Waagepetersen and Sorensen 2001) allowing an application of a change of variables in the detailed balance equation to facilitate the construction of the transition kernel of the Markov chain. Here, $\gamma_{pp'}(P^p, u)$ is the proposal density for the auxiliary variable u going from an AR model of order p to one with order p' which may also depend on the current parameter vector P^p. The proposed new order p' is drawn from some $\gamma_p(p'|p)$ and the joint proposal density is $\gamma(\varsigma) = \gamma_{pp'}(P^p, u)\gamma_p(p'|p)$.

In our implementation of the method, we use the following differentiable bijection for $g_{pp'}$

$$\begin{bmatrix} P^{p'} \\ u' \end{bmatrix} = g_{pp'}(P^p, u) = \begin{bmatrix} A(p, p')_{p' \times p} & I_{p' \times p'} \\ I_{p \times p} & 0_{p \times p'} \end{bmatrix} \begin{bmatrix} P^p \\ u \end{bmatrix} \tag{9}$$

where

$$A(p, p') = \begin{cases} \begin{bmatrix} I_{p \times p} \\ 0_{(p'-p) \times p} \end{bmatrix} & \text{if } p' > p \\ \begin{bmatrix} I_{p' \times p'} 0_{p' \times (p-p')} \end{bmatrix} & \text{if } p' < p \\ \begin{bmatrix} I_{p' \times p'} \end{bmatrix} & \text{if } p' = p \end{cases} \tag{10}$$

This mapping leads to the transdimensional analog of a full-site updating random walk sampler. Proposals for "newly born" parameters, i.e., those $P_i^{p'}$ for $i = p + 1, \ldots, p'$, are centered around zero. If $p' < p$ the parameter vector is truncated and proposals for these parameters are centered around their previous values. For $p' = p$ this mapping gives a standard random walk sampler.

The detailed balance condition holds if[7]

$$\int_{\mathcal{A}_p} \pi(\varsigma) Q\left(\varsigma, \mathcal{B}_{p'}\right) dP^p = \int_{\mathcal{B}_{p'}} \pi(\varsigma') Q\left(\varsigma', \mathcal{A}_p\right) dP^{p'} \tag{11}$$

for all subsets $\mathcal{A}_p$ and $\mathcal{B}_{p'}$ of the parameter spaces associated with autoregressive polynomials of order p and p' respectively and where

$$Q\left(\varsigma, \mathcal{B}_{p'}\right) = \int_{\mathcal{B}_{p'}} \gamma(\varsigma'|p, P^p)\alpha_{pp'}(\varsigma, \varsigma')d\varsigma'$$

is the first part of the kernel in (5), i.e. the part of the conditional distribution of ς' associated with acceptance of the proposal.

Implementing the change of variables with the mapping $g_{pp'}$, the detailed balance condition is satisfied if

$$\pi(\varsigma)\gamma_p(p'|p)\alpha_{pp'}\gamma_{pp'}(P^p, u) = \pi(\varsigma')\gamma_p(p|p')\alpha_{p'p}\gamma_{p'p}(g_{pp'}(P^p, u)) \tag{12}$$

where the details of the derivation can be found in the appendix.

[7] See also Waagepetersen and Sorensen (2001).

Following Peskun (1973), we set the acceptance probability, $\alpha_{pp'}$, as large as possible,[8]

$$\alpha_{pp'} = \min\left(1, \chi_{pp'}(\varsigma, \varsigma')\right) \tag{13}$$

with

$$\chi_{pp'}(\varsigma, \varsigma') = \underbrace{\frac{\mathcal{L}(\varsigma')}{\mathcal{L}(\varsigma)}}_{\text{Likelihood Ratio}} \underbrace{\frac{\rho(\varsigma')}{\rho(\varsigma)}}_{\text{Prior Ratio}} \underbrace{\frac{\gamma_p(p|p')\gamma_{p'p}(g_{pp'}(P^p, u))}{\gamma_p(p'|p)\gamma_{pp'}(P^p, u)}}_{\text{Proposal Ratio}} \tag{14}$$

Having chosen an appropriate acceptance probability to maintain detailed balanced, we can now implement the procedure.

The RJMCMC algorithm for our AR(p) model with unknown order p is then as follows

1. Set the initial state ς_0 of the Markov chain
2. For $i = 1$ to N
   (a) set $\varsigma = \varsigma_{i-1}$
   (b) Propose a visit to model p' with probability $\gamma_p(p'|p)$
   (c) Sample u from $\gamma_{pp'}(P^p, u)$
   (d) Set $(P', u') = g_{pp'}(P^p, u)$
   (e) Accept draw with probability
   $$\alpha = \min\left(1, \chi_{pp'}(\varsigma, \varsigma')\right)$$
   $\chi_{pp'}$ is defined as in (14)
   (f) If the draw is accepted set $\varsigma_i = \varsigma'$. If the draw is rejected set $\varsigma_i = \varsigma$

The sequence of states of the resulting Markov chain then approximates the joint posterior over all models indexed by their order p and the corresponding parameter vectors.

The application to moving average models follows by analogy and the extension to autoregressive moving average (ARMA) models is straightforward. In such an extension, one simply defines the model indicator as a two-element vector, proposing not only visits to some model with autoregressive order p' but also for a new order for the MA-polynomial q'.

2.4 Imposing Stationarity and Invertibility on ARMA(p,q) Sampling

For many applications, it is desirable to restrict the parameter spaces of ARMA processes to ensure stationarity and/or invertibility.[9] In order to constrain sampling to these invertible and stationary

[8] Which, as noted by Green (1995), is "optimal in the sense of reducing the autocorrelation of the chain."

[9] For the DSGE application in sections 5 and 6, we will require stationarity of the exogenous driving forces. In section 6, we will examine the consequences of imposing or not imposing invertibility on MA components, should

regions of the parameters spaces of each model, we reparametrize the AR (and MA) polynomial in terms of its (inverse) partial autocorrelations (PACs) following Barndorff-Nielsen and Schou (1973), Monahan (1984) and Jones (1987). If the (inverse) partial autocorrelations are between -1 and 1 the process is (invertible) stationary.

First, we generalize the AR(p) model to an ARMA(p,q) as follows

$$y_t = P_1^{p,q} y_{t-1} + P_2^{p,q} y_{t-2} + \ldots + P_p^{p,q} y_{t-p} + \epsilon_t + Q_1^{p,q} \epsilon_{t-1} + \ldots + Q_q^{p,q} \epsilon_{t-q}, \quad \epsilon_t \sim \mathcal{N}\left(0, \sigma^2\right) \tag{15}$$

In order to recover the coefficients of the AR polynomials, the following algorithm is run:

1. Introduce $p^k = \left(p_1^{(k)}, \ldots, p_k^{(k)}\right), k = 1, \ldots, p$
2. Draw $r = r_1, \ldots, r_p,$ for $r_i \in (0, 1)$ partial autocorrelations
3. Set $p_1^{(1)} = r_1$
4. Run the recursion

   $$p_i^{(k)} = p_i^{(k-1)} - r_k p_{k-i}^{(k-1)}, i = 1, \text{ for } \ldots, k-1$$

   with $p_k^{(k)} = r_k$ for $k = 2, \ldots, p$
5. Set $P^p = p^{(p)}$

The MA coefficients are recovered analogously, where the inverse partial autocorrelations substitute for the partial autocorrelations, r_i, in the foregoing. Ultimately, instead of proposing AR(MA) parameters directly, (inverse) partial autocorrelations are proposed in their place from which the parameters are then recovered. This will obviously necessitate the formulation of priors over (inverse) partial autocorrelations instead of parameters.

3 RJMCMC ARMA Order and Parameter Estimation: Monte Carlo Evidence

We examine the performance of the RJMCMC method for ARMA processes of unknown order introduced in the foregoing section by carrying out a Monte Carlo experiment. We generated 100 time series with 100 observations each from the following ARMA(3,2) process:

$$y_t = -0.75 y_{t-3} + \epsilon_t - 1.5\epsilon_{t-1} + 0.5625\epsilon_{t-2}, \quad \epsilon_t \sim N(0, 1.5^2)$$

they exist, on impulse responses.

For each time series, 1,500,000 draws were generated with the RJMCMC procedure. The first 1,000,000 draws were discarded as burn-in. The first state of the chain was set to white noise with unit standard deviation, i.e. $p = q = 0$ where p denotes the autoregressive order, q the moving average order, and $\sigma = 1$. Our metric for model choice is in accordance with a $1 - 0$ loss function, selecting the model at the mode of the posterior distribution over (p, q). It should be noted that one of the strengths of our method, of course, is the ability to quantify posterior uncertainty over models directly, such that model uncertainty can be incorporated in the calculation of posterior credible sets over impulse responses, correlations structures, or the like, providing more than just a point estimate of the model order.

We compare the model choice of our method with the choices that follow from using the Akaike Information Criterion (AIC), the corrected Akaike Information Criterion (AICC), and the Bayesian Information Criterion (BIC).[10] These are defined as

$$\text{AIC} = 2k - 2\ln(\hat{\mathcal{L}}), \quad \text{AICC} = AIC + \frac{2k(k+1)}{n-k-1}, \quad \text{BIC} = -2\ln(\hat{\mathcal{L}}) + k\ln(n)$$

with k being the number of model parameters and n the number of observations. $\hat{\mathcal{L}}$ denotes the maximized likelihood value of a model, i.e., for given ARMA orders p and q.

All three standard information criteria penalize for the number of parameters in the model. This feature is also present in the posterior of our RJMCMC method with proper priors over the (inverse) partial autocorrelations. Increasing the order of, say, an autoregressive model and setting the new parameter to zero gives a model identical to the previous one with lower order; hence, does not change the likelihood. Yet, the posterior with the additional parameter is penalized as the prior probability assigned to the value of the new parameter is smaller than one, yielding a posterior probability lower than with the original, lower order.

3.1 Priors and Proposals

Table 1 summarizes the priors and proposals used in the Monte Carlo study.

[Table 1 about here.]

We choose a uniform prior over the AR and MA orders, restricting the highest allowed order to 10 for both the AR and MA polynomials. Proposals for the AR and MA orders are taken to follow a

[10] Calculations for the three standard measures were carried out using the R package auto.arima.

discretized Laplace distribution, LaplaceD(μ, b), with location parameter, μ, and shape parameter, b, such that

$$\gamma_p(p'|p) \propto exp(-b|p - p'|) \text{ with } p', p \in [0, 1, \ldots, 10] \tag{16}$$

$$\gamma_q(q'|q) \propto exp(-b|q - q'|) \text{ with } q', q \in [0, 1, \ldots, 10] \tag{17}$$

For the (inverse) partial autocorrelations, our prior is a truncated normal distribution, TN($\mu, \sigma, -1, 1$), with location parameter, μ, and dispersion σ, and truncations at 1 and -1, imposing invertibility and stationarity. For the proposal distributions, we center the (inverse) partial autocorrelations around their previous values and new the (inverse) partial autocorrelations are centered around zero, sampling from a truncated normal distribution restricted to the invertibility and stationarity regions.

3.2 Likelihood

For the ARMA (p, q) model introduced in (15), we employ the Kalman filter to evaluated the log likelihood, $\ln \mathcal{L}\left(\{y_t\}_{t=1}^T ; \varsigma\right)$, as a sequence of conditional log likelihoods

$$\ln \mathcal{L}\left(\{y_t\}_{t=1}^T ; \varsigma\right) = \sum_{t=1}^{T} \ln \mathcal{L}\left(y_t | \left\{y_j\right\}_{j=1}^{t-1} ; \varsigma\right) = -\frac{1}{2} \sum_{t=1}^{T} \left[\ln \omega_t + \frac{\upsilon_t^2}{\omega_t} + \ln(2\pi)\right] \tag{18}$$

where the last equality follows from the assumption of normality; the sample size is $T = 100$; υ_t is the innovation in the current observation, $\upsilon_t \doteq y_t - E\left[y_t | \left\{y_j\right\}_{j=1}^{t-1}\right]$; and ω_t the conditional variance of this innovation, $\omega_t \doteq E\left[\upsilon_t^2 | \left\{y_j\right\}_{j=1}^{t-1}\right]$.

The innovation and its conditional variance are recovered from the Kalman filter recursion[11] where we follow Harvey (1993, p. 96) in setting up the recursion for ARMA(p,q) processes.[12] The state equation is

$$w_{t+1} = Aw_t + R\epsilon_t, \ \epsilon_t \sim N(0, \sigma^2) \tag{19}$$

and the observation equation is given by

$$y_t = Zw_t \tag{20}$$

where

(21)

$$Z = \begin{bmatrix} 1 & 0_{1\times m-1} \end{bmatrix}, \quad A = \begin{bmatrix} \mathbf{P}_{m-1}^{p,q} & I_{m-1} \\ P_m^{p,q} & 0_{1\times m-1} \end{bmatrix}, \quad \mathbf{P}_{m-1}^{p,q} = \begin{bmatrix} P_1^{p,q} & \ldots & P_{m-1}^{p,q} \end{bmatrix}', \quad R = \begin{bmatrix} 1 & Q_1^{p,q} & \ldots & Q_m^{p,q} \end{bmatrix}'$$

for $m = \max(p, q + 1)$.

[11] See, e.g., Anderson and Moore (1979).

[12] See de Jong and Penzer (2004) for an overview of alternate state space formulations of ARMA models.

3.3 Results

We report the proportion of correctly identified models in table 2. The performance of the RJMCMC method fits comfortably in the set of traditional information criteria. An increase in the number of the draws from the posterior could further improve the performance of our implementation.

[Table 2 about here.]

Additionally, as the RJMCMC method provides a complete characterization of the posterior distribution over models, we can explore different measures of posterior uncertainty over models. For example when the true model was not selected, figure 1 reports a histogram of the proportion of posterior mass assigned to the true model relative to the posterior mass at the mode of the posterior over (p, q) for all cases in which the wrong model was selected. In a few cases no posterior mass at all was assigned to the true model. This might either be due to non-convergence or "bad luck" in the sense that the realization of the observations exhibited statistical properties far from those of the true model.

[Figure 1 about here.]

Likewise, for when the true model was selected, we can explore the strength of the evidence in favor of the mode selected. In figure 2, we show a typical posterior over (p, q) randomly selected from the set of chains for which the posterior mode is at the true model. While there is some posterior uncertainty regarding the orders of the polynomials, the method clearly rejects lower orders than the true ones.

[Figure 2 about here.]

Of course, the ability of the method to estimate the parameters of the model along with the order of the model is of importance. Figure 3 reports the recursive means of the parameter draws of the model parameters conditional on $p = 3$ and $q = 2$. These values clearly converge close to the values underlying the data generating process.

[Figure 3 about here.]

In conclusion, our method exhibits roughly the same performance as classical methods concerning order identification while providing a complete posterior distribution over parameters and model orders that can be used for the posterior analysis of statistics of interest. We are interested in posterior statistics of DSGE models such as impulse responses and correlation structures and will now turn to a DSGE setting and apply the RJMCMC method there.

4 DSGE Models with Generalized Exogenous Processes

4.1 Class of Linear DSGE Models

We will consider linear(ized) DSGE models that can be expressed compactly as

$$0 = E_t\left[\underset{n_x\times 1}{AX_{t+1}} + BX_t + CX_{t-1} + D\underset{n_z\times 1}{Z_t}\right] \tag{22}$$

where the vector X_t collects the endogenous variables and the vector Z_t the exogenous variables. Instead of the standard assumption of independent AR(1) processes for the elements of the vector Z_t,[13] we shall allow each element in Z_t to be driven by an independent ARMA(p,q) process, whose orders p and q along with whose parameters we shall estimate using the RJMCMC algorithm developed in section 2.

4.2 Multiple ARMA(p,q) Processes

The method laid out in section 2 extends straigthforwardly to multiple autoregressive moving averages of finite order.[14] Specifically, we assume that each exogenous process can be represented as a finite order ARMA[15]

$$z_{i,t} = \rho_{i,1}z_{t-1} + \rho_{i,2}z_{i,t-2}\ldots + \rho_{i,p_i}z_{i,t-p_i} + \gamma_{i,0}\epsilon_{i,t} + \gamma_{i,1}\epsilon_{i,t-1}\ldots + \gamma_{i,q_i}\epsilon_{i,t-q_i}, \quad \epsilon_{i,t} \sim \mathcal{N}\left(0,\sigma_i^2\right) \tag{23}$$

[13] Notable exceptions are Cúrdia and Reis (2010) and Chari, Kehoe, and McGrattan (2007), who let their vector of disturbances follow a vector AR(1) process, and Del Negro and Schorfheide (2009) and Smets and Wouters (2007), who let two of their seven disturbances follow ARMA(1,1).

[14] We will examine multiple ARMA processes instead of VARMA (vector autoregressive moving averages) both to maintain the structural interpretation of the shock and to avoid the proliferation of parameters and reparameterizations, see Monahan (1984), needed to guarantee stationarity in vector processes.

[15] We adopt the convention that sums that terminate with an index smaller than that with which they began are empty sets. For example, if $p_i = 0$ in (23) for some i; $\sum_{j=1}^{-0}\rho_{i,j}z_{t-j} = \emptyset$ such that $z_{i,t}$ in this case would be

$$z_{i,t} = \gamma_{i,0}\epsilon_{i,t} + \gamma_{i,1}\epsilon_{i,t-1}\ldots + \gamma_{i,q_i}\epsilon_{i,t-q_i}$$

We assume that the processes in (23) are stationary and invertible, as we summarize in the following

Assumption 4.1. *The roots of the polynomial*

$$\rho_i(\lambda) \doteq \lambda^{p_i} - \rho_{i,1}\lambda^{p_i-1} + \rho_{i,2}\lambda^{p_i-2} \ldots + \rho_{i,p_i} \tag{24}$$

are all inside the unit circle. That is, there exists no λ such that $\rho_i(\lambda) = 0$ where $|\lambda| \geq 1$.

Expressed in vector form, the exogenous processes can be collected as

$$Z_t = P_1 Z_{t-1} + P_2 Z_{t-2} \ldots + P_p Z_{t-p} + I_{n_z \times 1} \epsilon_t + Q_1 \epsilon_{t-1} \ldots + Q_q \epsilon_{t-q}, \quad \epsilon_t \sim \mathcal{N}(0, \Sigma) \tag{25}$$

where p is the highest autoregressive order ($p = \max(\{p_i\})$) and q the highest moving average order ($q = \max(\{q_i\})$) among the exogenous processes. The covariance matrix Σ is diagonal, collecting the variances of the individual processes along the diagonal—$\Sigma \doteq \operatorname{diag}\left(\sigma_1^2, \sigma_2^2, \ldots, \sigma_{n_z}^2\right)$. The stationarity and invertibility of the individual processes in assumption 4.1 transfers to the vector process (25), as we state formally as

Lemma 4.2. *The latent roots of the λ matrix*

$$I_{n_z}\lambda^p - P_1\lambda^{p-1} + P_2\lambda^{p-2} \ldots + P_p \tag{26}$$

That is, there exists no λ such that $\det(P(\lambda)) = 0$ *where $|\lambda| \geq 1$.*

Proof. Follows directly from assumption 4.1. □

4.3 Recursive Solution

We will solve for a recursive solution for the endogenous variables in the model (22) using a method of undetermined coefficients approach. Given (25) and (22), the state variables of the model are

$$\left\{X_{t-1}, Z_t, Z_{t-1}, \ldots, Z_{t-(\tilde{p}-1)}, \epsilon_t, \epsilon_{t-1}, \ldots, \epsilon_{t-(q-1)}\right\} \tag{27}$$

where $\tilde{p} = \max(p, 1)$,[16]

While we could redefine the model (22) to include the entire state vector (27) as endogenous variables to bring the model into the canonical form of, say, Sims (2001) or Klein (2000), doing

[16]This follows directly from (25) expressed in first order vector form

$$\begin{bmatrix} Z_t' & Z_{t-1}' & \ldots & Z_{t-(\tilde{p}-1)}' & \epsilon_t' & \epsilon_{t-1}' & \ldots & \epsilon_{t-(q-1)}' \end{bmatrix}' = PP \begin{bmatrix} Z_{t-1}' & Z_{t-2}' & \ldots & Z_{t-p}' & \epsilon_{t-1}' & \epsilon_{t-1}' & \ldots & \epsilon_{t-q}' \end{bmatrix}' + QQ\epsilon_t$$

for appropriate PP and QQ matrices. The left hand side of the foregoing is then the current exogenous state vector. The case $p = 0$ is permitted through $\tilde{p}$, which ensures Z_t remains on the left hand side of the foregoing despite the indexing convention laid out in footnote 15.

so would significantly increase the computation costs involved in the QZ decomposition for the state transition and the Sylvester equation for the impact matrix of shocks. The solution for the endogenous variables is, accordingly, given by

$$X_t = \Lambda X_{t-1} + \Phi_0 Z_t + \Phi_1 Z_{t-1} \ldots + \Phi_{\tilde{p}-1} Z_{t-(\tilde{p}-1)} + \Theta_0 \epsilon_t + \Theta_1 \epsilon_{t-1} \ldots + \Theta_{q-1} \epsilon_{t-(q-1)} \tag{28}$$

where

$$\left\{\Lambda, \Phi_0, \Phi_1, \ldots, \Phi_{\tilde{p}-1}, \Theta_0, \Theta_1, \ldots, \Theta_{q-1}\right\} \tag{29}$$

are the unknown coefficients that we solve for.

We will make the following two assumptions that correspond to the Blanchard and Kahn's (1980) order and rank conditions to guarantee a unique stable solution. The order condition assumes a full set of latent roots with half inside and half outside the unit circle

Assumption 4.3. *Order*

There exist $2n_x$ *latent roots of* $A\lambda^2+B\lambda+C$*—that is,* $n_x+\text{rank}\,(A)$ *finite* $\lambda \in \mathbb{R} : \det A\lambda^2 + B\lambda + C = 0$ *as well as* $n_x - \text{rank}A$ *infinite* λ*—of which* n_x *lie inside the unit circle and* n_x *outside.*

We then assume that a solution, or solvent, can be constructed containing these stable roots

Assumption 4.4. *Rank*

There exists an $\Lambda \in \mathbb{R}^{n_x \times n_x}$ *such that* $A\Lambda^2 + B\Lambda + C = 0$ *and* $|eig(\Lambda)| < 1$.

Thus, Λ is the unique solution to the matrix quadratic equation $A\Lambda^2 + B\Lambda + C = 0$ whose eigenvalues coincide with the stable latent roots of the quadratic λ matrix $A\lambda^2 + B\lambda + C$.[17]

Under the order and rank assumptions, as well as the stationarity assumption on the exogenous processes, the model (22) has a unique, stable solution, as we summarize in the following proposition

Proposition 4.5. *Let assumptions 4.3, 4.4, and 4.2 hold. There exists a unique, stable solution (28) to (22). The coefficient* Λ *in (28) is the solvent of assumption 4.4, the coefficients* $\left\{\Theta_0, \Theta_1, \ldots, \Theta_{q-1}\right\}$

[17] See Lancaster (1966), Dennis, Jr., Traub, and Weber (1976), and Higham and Kim (2000) for detailed analysis of matrix polynomials and λ matrices, as well as Lan and Meyer-Gohde (2012) for an application to DSGE models.

for $q > 0$ solve

$$\begin{aligned}
\underset{n_x \times n_z}{0} &= A\left(\Lambda\Theta_0 + \Phi_0 Q_1 + \Theta_1\right) + B\Theta_0 \\
\underset{n_x \times n_z}{0} &= A\left(\Lambda\Theta_1 + \Phi_0 Q_2 + \Theta_2\right) + B\Theta_1 \\
&\vdots \\
\underset{n_x \times n_z}{0} &= A\left(\Lambda\Theta_{q-2} + \Phi_0 Q_{q-1} + \Theta_{q-1}\right) + B\Theta_{q-2} \\
\underset{n_x \times n_z}{0} &= A\left(\Lambda\Theta_{q-1} + \Phi_0 Q_q\right) + B\Theta_{q-1} \qquad (30)
\end{aligned}$$

and the coefficients $\left\{\Phi_0, \Phi_1, \ldots, \Phi_{p-1}\right\}$ solve

$$\begin{aligned}
\underset{n_x \times n_z}{0} &= A\left(\Lambda\Phi_0 + \Phi_0 P_1 + \Phi_1\right) + B\Phi_0 + D \\
\underset{n_x \times n_z}{0} &= A\left(\Lambda\Phi_1 + \Phi_0 P_2 + \Phi_2\right) + B\Phi_1 \\
&\vdots \\
\underset{n_x \times n_z}{0} &= A\left(\Lambda\Phi_{p-2} + \Phi_0 P_{p-1} + \Phi_{p-1}\right) + B\Phi_{p-2} \\
\underset{n_x \times n_z}{0} &= A\left(\Lambda\Phi_{p-1} + \Phi_0 P_p\right) + B\Phi_{p-1} \qquad (31)
\end{aligned}$$

for $p > 0$ and Φ_0 solves

$$\underset{n_x \times n_z}{0} = A\Lambda\Phi_0 + B\Phi_0 + D$$

otherwise.

Proof. Insert the solution (28) for X_t once and for X_{t+1} twice in (22), substitute (25) lagged forward once for the Z_{t+1} that arises when X_{t+1} is replaced with (28), and then collect coefficients on the state variables (27). As the solution (28) must hold for all values of the state variables, the coefficients just collected must all be zero. The resulting equations are those stated in the proposition. □

We can also calculate an infinite moving average representation for the solution, which will prove useful in the estimation exercise, allowing us to calculate the likelihood spectrally and to apply the closed form frequency domain representation of the HP filter (Hodrick and Prescott 1997) to treat the model with the filter while estimating. Taking the unique stable solution derived above as given, we define the following λ matrices for the exogenous processes

$$P(\lambda) \doteq I_{n_z} - P_1\lambda - P_2\lambda^2 \ldots - P_p\lambda^p \qquad (32)$$

$$Q(\lambda) \doteq I + Q_1\lambda \ldots + Q_q\lambda^q \qquad (33)$$

and for the endogenous transfer function

$$\Phi(\lambda) \doteq \Phi_0 + \Phi_1\lambda \ldots + \Phi_{\tilde{p}-1}\lambda^{\tilde{p}-1} \tag{34}$$

$$\Theta(\lambda) \doteq \Theta_0 + \Theta_1\lambda \ldots + \Theta_{q-1}\lambda^{q-1} \tag{35}$$

Replacing λ with the lag or backshift operator L,[18] we can express X_t as an infinite moving average, as we summarize in the following proposition

Proposition 4.6. *Let assumptions 4.3, 4.4, and 4.2 hold. The unique, stable solution (28) to (22) for X_t in proposition 4.5 has a unique infinite moving average representation given by*

$$X_t = \left(\underset{n_x \times n_x}{I} - \Lambda L\right)^{-1} \left[\Phi(L) P(L)^{-1} Q(L) + \Theta(L)\right] \epsilon_t \tag{36}$$

Proof. Invertibility of $\left(\underset{n_x \times n_x}{I} - \Lambda L\right)$ follows from proposition 4.5 and that of $P(L)$ from lemma 4.2. Uniqueness follows from the uniqueness of the homogenous representation from assumptions 4.3 and 4.4 and of the uniqueness of the inhomogenous representation from proposition 4.5. □

4.4 DSGE Likelihood

One difficulty in implementing likelihood methods lies in the evaluation of the likelihood function. As we will consider applying the HP filter to the model when it was applied to the data, the Kalman filter is less desirous here due to the availability of a closed form frequency domain representation for the HP filter, see King and Rebelo (1993). We follow an alternative approach based on the Toeplitz structure of the covariance of stationary time series that uses the iterative method of Meyer-Gohde (2010) for evaluating the likelihood function by treating the sample as a single draw from a multivariate normal distribution,[19] where the derivation of the sequence of autocovariances is done spectrally to enable us to apply the HP filter to the model while evaluating the likelihood function.

Consider now a linear combination of elements of X_t. I.e., the observables, given by

$$Y_t = \underset{n_y \times n_x}{\Upsilon^X} X_t \tag{37}$$

To evaluate the likelihood function, we will need to calculate the sequence of autocovariance matrices associated with the observables, Y_t,

$$\Gamma_0 \doteq E\left[Y_t Y_t'\right],\ \Gamma_1 \doteq E\left[Y_t Y_{t-1}'\right],\ \ldots \Gamma_n \doteq E\left[Y_t Y_{t-n}'\right] \tag{38}$$

[18] See, e.g., Sargent (1987)

[19] Similarly to Leeper and Sims (1994) and Schmitt-Grohé and Uribe (2010).

Using the moving average representation of the observables

$$Y_t = \Upsilon^X \left(\underset{n_x \times n_x}{I} - \Lambda L \right)^{-1} \left[\Phi(L) P(L)^{-1} Q(L) + \Theta(L) \right] \epsilon_t \tag{39}$$

The autocovariances can be recovered, see, e.g., Sargent (1987), Hamilton (1994), and Uhlig (1999), through

$$\Gamma_n = \int_{-\pi}^{\pi} G(\omega) e^{i\omega n} d\omega \tag{40}$$

the inverse Fourier transformation of the spectral density of Y_t, $G(\omega)$ given by

$$\begin{aligned} G(\omega) = &\left[\Upsilon^X \left(\underset{n_x \times n_x}{I} - \Lambda e^{-i\omega} \right)^{-1} \left[\Phi\left(e^{-i\omega}\right) P\left(e^{-i\omega}\right)^{-1} Q\left(e^{-i\omega}\right) + \Theta\left(e^{-i\omega}\right) \right] \right] \\ &\times \Sigma \left[\Upsilon^X \left(\underset{n_x \times n_x}{I} - \Lambda e^{i\omega} \right)^{-1} \left[\Phi\left(e^{i\omega}\right) P\left(e^{i\omega}\right)^{-1} Q\left(e^{i\omega}\right) + \Theta\left(e^{i\omega}\right) \right] \right]' \end{aligned} \tag{41}$$

As we will also consider applying the HP filter to the model as well as to the data, we can use closed form representation of the HP filter in the frequency domain, see King and Rebelo (1993), given as

$$HP(\lambda, \omega) = \frac{4\lambda (1 - \cos(\omega))^2}{1 + 4\lambda (1 - \cos(\omega))^2} \tag{42}$$

where λ is the HP smoothing parameter and ω a frequency. In this case, the autocovariances of the HP filtered observables can be recovered through

$$\Gamma_n = \int_{-\pi}^{\pi} HP(\lambda, \omega)^2 G(\omega) e^{i\omega n} d\omega \tag{43}$$

Given the assumptions of linearity and stationarity behind proposition 4.5 and that of the normality of the innovations ϵ_t, T observations on Y_t are normally distributed with mean zero and non-singular block Toeplitz covariance matrix

$$\Psi = \begin{bmatrix} \Gamma_0 & \Gamma_1' & \dots & \Gamma_{T-2}' & \Gamma_{T-1}' \\ \Gamma_1 & \Gamma_0 & \dots & \Gamma_{T-3}' & \Gamma_{T-2}' \\ \vdots & & \ddots & & \vdots \\ \Gamma_{T-2} & \Gamma_{T-3} & \dots & \Gamma_0 & \Gamma_1' \\ \Gamma_{T-1} & \Gamma_{T-2} & \dots & \Gamma_1 & \Gamma_0 \end{bmatrix} \tag{44}$$

with the autocovariance matrices, Γ_n, given by (40) or (43) depending on whether the HP filter was used and the log-likelihood of a vector of parameters ς given the data is thus

$$\mathcal{L}(\varsigma | Y) = -0.5pT ln(2\pi) - 0.5 ln(det(\Psi(\vartheta))) - 0.5 Y' \Psi(\vartheta)^{-1} Y \tag{45}$$

where $X = [Y_1' Y_2' \dots Y_T']'$.

Given (44), only two potentially challenging quantities need to be calculated: $ln(det(\Psi(\vartheta)))$ and $X'\Psi(\vartheta)^{-1}X$, which we calculate using the recursive block-Levinson type algorithm of Meyer-Gohde (2010).

5 Neoclassical Growth Model

As a baseline model to examine how the RJMCMC model can be applied to a DSGE model, we consider Hansen (1985) specification of the neoclassical growth model. In this simple model, the social planner's problem is to maximize the discounted lifetime expected utility of a representative household given by

$$E_0 \sum_{t=0}^{\infty} \beta^t \left[ln\left(c_t\right) + \psi ln\left(1 - l_t\right) \right], \ 0 < \beta < 1 \tag{46}$$

with c_t representing consumption and l_t hours; $\beta \in (0, 1)$ is the subjective discount factor of the household and ψ weights the utility of leisure, $1 - l_t$, in the household's utility function. The social planner faces the resource constraint

$$c_t + i_t = y_t \tag{47}$$

where investment, i_t, contributes to the accumulation of capital, k_t, through

$$k_t = \left(1 - \delta\right) k_{t-1} + i_t \tag{48}$$

with the depreciation rate, δ, and where production, y_t is neoclassical and given by

$$y_t = e^{z_t} k_{t-1}^{\alpha} l_t^{1-\alpha} \tag{49}$$

with z_t being stationary stochastic productivity. Hansen (1985) assumed a highly autocorrelated AR(1) process—with the autoregressive parameter set to 0.95— following Kydland and Prescott (1982). Relaxing this assumption will be the focus of our investigation.

The first order conditions of the social planner's problem are given by

$$\frac{1}{c_t} = \beta E_t \left[\frac{1}{c_{t+1}} \left(1 - \delta + \alpha e^{z_{t+1}} \left(\frac{l_{t+1}}{k_t} \right)^{1-\alpha} \right) \right] \tag{50}$$

$$\frac{\psi}{1 - l_t} = \frac{1}{c_t} \left(1 - \alpha\right) e^{z_t} \left(\frac{k_{t-1}}{l_t} \right)^{\alpha} \tag{51}$$

An equilibrium is defined by the equations (47) through (51) along with a specification for the stochastic productivity process, z_t.

In this exercise, we will take the parameters of Hansen's (1985) calibration of all parameters outside the specification of the stochastic productivity process, z_t, as given. This will allow us to concentrate on the contribution of the RJMCMC algorithm in estimating the order and parameters of the exogenous process.The calibrated model parameters are reported in table 3.

[Table 3 about here.]

The calibration delivers standard values for parameters, reflecting empirical observations, e.g., that about one third of agents' time endowment is spent in employment activities, capital contributes a little more than one third to production.

6 Estimation Results for the Neoclassical Growth Model Model

We carry out two exercises using the neoclassical growth model model as presented above. First, in order to check whether the method could pick up the correct underlying process for a technology shock in this model, we generated 250 observations of synthetic data using the AR(1) process as reported by Hansen (1985) in his original study. Second, we estimate order and parameters of the technology shock process for the model using US GDP data, treated with the HP filter as in Hansen's (1985) original study.

6.1 Priors and Proposals

The priors and proposals for the shock process orders and parameters are reported in table 4.

[Table 4 about here.]

The priors remain the same as in the Monte Carlo study, while the dispersion parameters of the proposals were tuned on the basis of short pilot runs to increase the efficiency of the RJMCMC algorithm.

6.2 Synthetic AR(1) Data

For this exercise we generated 250 realizations for the technology shock according to the AR(1) specification and calibration in Hansen (1985)

$$z_t = 0.95 z_{t-1} + \epsilon_t \tag{52}$$

We then fed the resulting series for z_t into the linearized RBC model and applied our method generating 650.000 draws discarding the first 100.000 draws as burning to the resulting synthetic data on output, y_t. Standard visual measures over the chains indicated convergence. Figure 4 shows the posterior distribution over the orders for the disturbance. The method places an overwhelming

majority of the posterior weight on the correct, AR(1), model order—obviously correctly identifying the AR(1) data generating process for the productivity process with observations on output, y_t.

[Figure 4 about here.]

This result gives us further confidence that, if the real world process for the productivity shock were AR(1), it would be correctly identified by the RJMCMC method we propose.

6.3 US GDP Data: Estimates

We now address what US postwar GDP data can reveal about the productivity shock in Hansen's (1985) model. Following Hansen's (1985), [20] we estimated the productivity shock process using HP-filtered quarterly US GDP per capita as in Hansen (1985) taking his original calibration and value of 1600 for the smoothing parameter in the HP filter as given. In applying the RJMCMC method introduced in section 2, we generated 4.000.000 draws discarding the first 1.000.000 draws as burn in. The HP filter was applied to the DSGE model when evaluating the likelihood, thus treating the data and the model with the same filter.[21]

[Figure 5 about here.]

Figure 5 shows the posterior over (p, q) for this exercise. The model at the mode is ARMA(3,0) and the baseline AR(1) specification of Hansen (1985) is clearly rejected. There is much more substantial uncertainty regarding the correct shock process than in the Monte Carlo exercises above. The prior posterior plots are indicative that our results are not being overly driven by our choice of priors.

[Figure 6 about here.]

Figure 6 reports recursive means of the first AR parameter for three chains with differing initial states for the orders of the ARMA polynomial for the technology shock, calculated both conditional

[20] We take 1948:1-2013:3 real GDP from the NIPA tables, expressed on a per capita basis using the BLS series on the civilian noninstitutional population. Both data sets were downloaded from the St. Louis Federal Reserve's FRED database.

[21] See section 4.4 for details.

on the model at the mode of the posterior as well as unconditional means. Inspection suggests that all three chains have converged. It is not clear, however, whether these standard graphical or other formal measures of convergence, e.g., Brooks and Gelman (1998), apply without adaptation in transdimensional analyses, see e.g. Fan and Sisson (2011). In any case, the posterior statistics, such as impulse responses, that we will examine are indicative of a lack of convergence.

[Figure 7 about here.]

Table 5 reports point estimates for the shock process parameters taken from the posterior distribution conditional on $(p, q) = (3, 0)$. Additionaly, the first two autocorrelations of the exogenous process, z_t, implied by these point estimates are given. The first autocorrelation is high, consistent with the choice of Hansen (1985) following Kydland and Prescott (1982) to model the technology process with a near unit root.

[Table 5 about here.]

6.4 US GDP Data: Impulse Responses

With a posterior distribution over both models—i.e., orders p and q—and their parameters for the ARMA technology process, we plot impulse responses taking posterior uncertainty about the model into account. In the presence of MA components, this requires us to take a stand on which covariance equivalent representation we choose.[22] We will first examine the invertible or fundamental impulse responses associated with the posterior distribution. Then, we will allow the possibility of nonfundamental representations by sampling with a noninformative prior from the admissible (i.e., real valued) covariance equivalent representations and examine the resulting impulse responses.

For an invertible or fundamental moving average representation, the roots, λ_{q_i}, of the MA polynomial

$$\gamma_i(1/\lambda) \doteq \lambda^{q_i} + \gamma_{i,1}\lambda^{q_i-1} \ldots + \gamma_{i,q} \tag{53}$$

must all be contained within the unit circle. That is, there exists no λ such that $\gamma_i(\lambda) = 0$ where $|\lambda| \geq 1$.[23] In figure 8, we plot the impulse responses to a one standard deviation technology shock.

[22] See Lippi and Reichlin (1994), Fernández-Villaverde, Rubio-Ramírez, Sargent, and Watson (2007), and Alessi, Barigozzi, and Capasso (2011) for more on different MA representations in macroeconomic modeling.

[23] See, e.g., Hamilton (1994).

We plot the invertible impulse associated with the model at the posterior mode of the ARMA order and parameter space against the pointwise posteriors (mode and 80% credible set) over all impulse responses weighted by posterior probabilities. Additionally, we plot the impulse response with Hansen's (1985) AR(1) technology assumption. The data driven selection of the specification of the shock process implies a different dynamic behavior of the model compared to Hansen's calibration. Our RJMCMC procedure identifies hump-shaped impulse responses, a salient feature of the data identified in many empirical studies; e.g., Cogley and Nason (1995) identify a hump shaped response of output to transitory technology shocks using both an SVAR and a VEC model.

[Figure 8 about here.]

In admitting nonfundamental or noninvertible MA representations, we acknowledge that the covariance structure associated with our posterior distribution potentially implies several possible different structural representations. We follow Lippi and Reichlin (1994) and engage in a root-flipping procedure to construct admissible covariance equivalent representations. We proceed as follows. Given a draw of order q for the MA component of the exogenous process, we factor the MA polynomial as

$$1 + \gamma_{i,1}L \ldots + \gamma_{i,q}L^{q_i} = (1 - \lambda_1 L)(1 - \lambda_2 L)\ldots\left(1 - \lambda_{q_i}L\right) \tag{54}$$

we then enumerate the number of different admissible root flips[24] and draw an integer $n \in \{0, 1, \ldots, \tilde{n}\}$ from a uniform distribution, where $\tilde{n}$ is the number of admissible equivalent representations. A draw of 0 gives the invertible or fundamental representation and a positive draw returns the non-fundamental representation associated with that integer during the enumeration. For example, if $n = 10$ is drawn and the number 10 was associated with flipping roots λ_2 and λ_3, the MA polynomial for calculating impulse responses becomes

$$\gamma_i(L) = (-\lambda_2)(-\lambda_3)\left(1 - \frac{1}{\lambda_2}L\right)\left(1 - \frac{1}{\lambda_3}L\right)(1 - \lambda_1 L)(1 - \lambda_4 L)\ldots\left(1 - \lambda_{q_i}L\right) \tag{55}$$

Drawing the covariance equivalent representation from a uniform distribution over all admissible covariance equivalent representations puts equal weight on each admissible representation, reflecting our flat prior over the different representations over which DSGE theory is noninformative.

Figure 9 contains the pointwise posteriors (mode and 80% credible set) over all impulse responses weighted by posterior probabilities and drawn, potentially, from nonfundamental covari-

[24] "Admissible" because conjugate pairs of complex roots must be flipped together, as the flipping of a single complex root would result in a complex-valued impulse response that we rule out on economic grounds.

ance equivalent representations as outlined above. We plot these pointwise posteriors against the invertible representation of the model at the posterior mode over ARMA orders and their parameter values and against the impulse response with Hansen's (1985) AR(1) technology assumption. The admission of non-fundamental representations increases our uncertainty over the dynamic response of variables to a technology innovation, spreading the bounds of the 80% credible sets apart. Most of this spread is downward so that the number of periods for which the 80% credible set covers exclusively positive responses to a technology shock is greatly reduced.

Interestingly, when allowing for non-fundamental moving average representation a negative response of hours to a positive technology shock is contained in the credible set. It is therefore possible even in a simple stochastic growth model with an estimated technology shock process allowing for news shocks, that the responses are in line with the findings of Gali (1999) and Francis and Ramey (2005). One cannot, therefore, conclude, that the stochastic growth model is unable to generate this kind of response to a positive technology shock unless one has a strong prior against the existence of noninvertible moving average representations, e.g., news shocks and policy announcement shocks. At the same time, however, the majority of posterior mass lies in a region where the response of hours to technology is conventional, in line with the results in V.V. Chari and McGrattan (2008) or Uhlig (2004).

[Figure 9 about here.]

In sum, the posterior mode model and the posterior distribution over impulse responses, both fundamental and admitting the possibility of non-fundamental moving average representations, as markedly different than those implied by the AR(1) assumption in Hansen's (1985) original study. The data clearly favors hump-shaped impulse responses and cannot rule out a drop in hours in response to a positive technology shock.

6.5 US GDP Data: Correlation Structure

We now examine the variance and correlation structures implied by our posteriors and compare these with the data and the statistics implied by our baseline AR(1) model implied by Hansen (1985).[25] The posterior matches the structure of the second moments of output quite well. As we

[25] Following Hansen (1985), we calculate the second moments for his model using an HP filtered (with the smoothing parameter, λ, set to 1600) version of model.

estimated with real per capita GDP data, this is reassuring and indicates that the procedure does indeed provide a substantial improvement in fit.

[Table 6 about here.]

The standard deviations of output are in table 6. Both the standard deviation of model at the posterior mode of the ARMA order and parameter space and the posterior mode of the standard deviations line up very close to the statistic in the data, whereas the statistic of Hansen (1985) shows greater a difference from the value in the data. The 80% posterior credible set shows the extent of posterior uncertainty, which here is great enough to encompass all the point values reported. The first six autocorrelations tell a more certain story, however, and can be found in figure 10. Again, both the autocorrelations of the model at the posterior mode of the ARMA order and parameter space and the posterior mode of the autocorrelations match the statistic in the data very closely. The AR(1) structure imposed by Hansen (1985) forces a compromise, with the initial autocorrelation are somewhat lower and the later values somewhat higher than in the data.

[Figure 10 about here.]

The fit as implied by the point estimates of our posterior with respect to our observable series output is reassuring in that our application of the RJMCMC method is successfully doing what it should. With a mean zero normally distributed process, the second moments describe the stochastic properties of the process and our posterior brings the second moments of output from the RBC model closer to the data by selecting appropriate ARMA processes.

7 Conclusion

In this paper we present a novel approach to addressing misspecification in DSGE models. We relax the assumptions usually placed on the structure of exogenous processes, usually taken to follow simple AR(1) processes, and estimate generalized, ARMA(p, q) processes of unknown orders. Since theory provides no guidance on autocorrelation patterns of exogenous variables and the order of the these processes is seldom if ever estimated, the usual choice of the AR(1) structure

on exogenous processes would seem to have little support. From a Bayesian perspective, it does not appear sensible to put a point prior on an AR(1) structure for an exogenous process in a DSGE model, while acknowledging uncertainty regarding other parameters of the model. Our method treats the ARMA orders of shock processes as additional parameters to be estimated, enabling the researcher to identify those shock process structures which bring the model closer to the data.

Our method has the advantage that it will ultimately enable the analysis of a joint posterior over different specifications of the exogenous processes including their parameters as well as parameters of the model, as we are investigating in work in progress. This allows for the quantification of posterior uncertainty regarding the model parameters and all parameters of the exogenous processes including their orders, while maintaining the interpretability of these processes as structural. The impulse responses implied by the estimated ARMA using US GDP data with Hansen's (1985) specification of the canonical stochastic neoclassical growth model are markedly different than those generated under the original calibration. Our posterior clearly identifies hump-shaped impulse responses and cannot rule out a drop in hours in response to a positive technology shock.

In applications directed at providing policy recommendations, our procedure should be brought to bear on a policy-relevant model with the data series of interest and combined with standard MCMC estimation of parameters, giving policy makers estimated models that take the cross-equation restrictions of rational expectations models seriously while accounting for model (in terms of exogenous driving forces) as well as parameter uncertainty. If one interprets the richer shock structure preferred by our method as a means of controlling for misspecification and, insofar as this misspecification is taken to be policy invariant, the generalized shocks should improve the accuracy of policy experiments while at the same time improving the fit of the model, as indicated in Del Negro and Schorfheide (2009).

References

Alessi, L., M. Barigozzi, and M. Capasso (2011): "Non-Fundamentalness in Structural Econometric Models: A Review," *International Statistical Review*, 79(1), 16–47.

An, S., and F. Schorfheide (2007): "Bayesian Analysis of DSGE Models," *Econometric Reviews*, 26(2-4), 113–172.

Anderson, B. D. O., and J. B. Moore (1979): *Optimal Filtering*. Prentice-Hall, Inc.

Barndorff-Nielsen, O., and G. Schou (1973): "On the Parametrization of Autoregressive Models by Partial Autocorrelations," *Journal of Multivariate Analysis*, 3, 408–419.

Blanchard, O. J., and C. M. Kahn (1980): "The Solution of Linear Difference Models under Rational Expectations," *Econometrica*, 48(5), 1305–1311.

Brooks, S. P., and R. Ehlers (2004): "Bayesian Analysis of Order Uncertainty in ARIMA Models," Discussion paper, Federal University of Paran, Brazil; University of Cambridge, UK.

Brooks, S. P., and A. Gelman (1998): "General Methods for Monitoring Convergence of Iterative Simulations," *Journal of Computational and Graphical Statistics*, 7(4), 434–455.

Brooks, S. P., P. Giudici, and G. O. Roberts (2003): "Efficient construction of reversible jump Markov chain Monte Carlo proposal distributions," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(1), 3–39.

Cappè, O., C. P. Robert, and T. Rydèn (2003): "Reversible jump, birth-and-death and more general continuous time Markov chain Monte Carlo samplers," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(3), 679–700.

Chari, V., P. J. Kehoe, and E. R. McGrattan (2007): "Business Cycle Accounting," *Econometrica*, 75(3), 781–836.

Chib, S., and E. Greenberg (1995): "Understanding the Metropolis-Hastings Algorithm," *The American Statistician*, 49(4), 327–335.

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

Cúrdia, V., and R. Reis (2010): "Correlated Disturbances and U.S. Business Cycles," NBER Working Papers 15774, National Bureau of Economic Research, Inc.

de Jong, P., and J. Penzer (2004): "The ARMA Model in State Space Form," *Statistics & Probability Letters*, 70(1), 119 – 125.

Del Negro, M., and F. Schorfheide (2009): "Monetary Policy Analysis with Potentially Misspecified Models," *American Economic Review*, 99(4), 1415–50.

Dennis, Jr., J. E., J. F. Traub, and R. P. Weber (1976): "The Algebraic Theory of Matrix Polynomials," *SIAM Journal on Numerical Analysis*, 13(6), 831–845.

Ehlers, R., and S. P. Brooks (2008): "Adaptive Proposal Construction for Reversible Jump MCMC," *Scandinavian Journal of Statistics*, 35(4), 677–690.

Fan, Y., and S. Sisson (2011): *Handbook of Markov Chain Monte Carlo*chap. 3, pp. 67–87. CRC Press.

Fernández-Villaverde, J., and J. F. Rubio-Ramírez (2004): "Comparing dynamic equilibrium models to data: a Bayesian approach," *Journal of Econometrics*, 123(1), 153 – 187.

Fernández-Villaverde, J., J. F. Rubio-Ramírez, T. J. Sargent, and M. W. Watson (2007): "ABCs (and Ds) of Understanding VARs," *American Economic Review*, 97(3), 1021–1026.

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

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

Gelfand, A. E., and A. F. M. Smith (1990): "Sampling-Based Approaches to Calculating Marginal Densities," *Journal of the American Statistical Association*, 85(410), pp. 398–409.

Geweke, J. (1998): "Using simulation methods for Bayesian econometric models: inference, development, and communication," Staff Report 249, Federal Reserve Bank of Minneapolis.

Golub, G. H., and C. F. V. Loan (1996): *Matrix Computations*. The Johns Hopkins University Press, Baltimore, MD, and London, UK, 3 edn.

Green, P. J. (1995): "Reversible Jump Markov Chain Monte Carlo," *Biometrika*, 82, 711–732.

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

Hansen, G. D. (1985): "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics*, 16(3), 309–327.

Harvey, A. C. (1993): *Time Series Models*. Harvester Wheatsheaf, London, UK, 2 edn.

Hastings, W. (1970): "Monte Carlo Sampling Methods using Markov Chains and Their Applications," *Biometrika*, 57, 97–109.

Higham, N. J., and H.-M. Kim (2000): "Numerical Analysis of a Quadratic Matrix Equation," *IMA Journal of Numerical Analysis*, 20, 499–519.

Hodrick, R. J., and E. C. Prescott (1997): "Postwar U.S. Business Cycles: An Empirical Investigation," *Journal of Money, Credit and Banking*, 29(1), 1–16.

Jones, M. (1987): "Randomly Choosing Parameters from the Stationarity and Invertibility Region of Autoregressive-Moving Average Models," *Journal of the Royal Statistical Society, Series C (Applied Statistics)*, 36, 134–138.

King, R. G., and S. T. Rebelo (1993): "Low frequency filtering and real business cycles," *Journal of Economic Dynamics and Control*, 17(1-2), 207–231.

Klein, P. (2000): "Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model," *Journal of Economic Dynamics and Control*, 24(10), 1405–1423.

Kydland, F. E., and E. C. Prescott (1982): "Time to Build and Aggregate Fluctuations," *Econometrica*, 50(6), 1345–70.

Lan, H., and A. Meyer-Gohde (2012): "Existence and Uniqueness of Perturbation Solutions to DSGE Models," SFB 649 Discussion Papers 2012-015.

Lancaster, P. (1966): *Lambda-Matrices and Vibrating Systems*. Pergamon Press, Oxford.

Leeper, E. M., and C. A. Sims (1994): "Toward a Modern Macroeconomic Model Usable for Policy Analysis," in *NBER Macroeconomics Annual*, ed. by S. Fischer, and J. J. Rotemberg, vol. 9, pp. 81–118. MIT Press Books.

Lippi, M., and L. Reichlin (1994): "VAR analysis, nonfundamental representations, blaschke matrices," *Journal of Econometrics*, 63(1), 307–325.

Metropolis, N., A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller (1953): "Equation of State Calculations by Fast Computing Machines," *The Journal of Chemical Physics*, 21(6), 1087–1092.

Meyer-Gohde, A. (2010): "Linear Rational-Expectations Models with Lagged Expectations: A Synthetic Method," *Journal of Economic Dynamics and Control*, 34(5), 984–1002.

Monahan, J. (1984): "A note on enforcing stationarity in autoregressive-moving average models," *Biometrika*, 71, 403–404.

Peskun, P. (1973): "Optimum Monte-Carlo sampling using Markov chains," *Biometrika*, 60, 607–612.

Philippe, A. (2006): "Bayesian analysis of autoregressive moving average processes with unknown orders," *Computational Statistics & Data Analysis*, 51, 1904–1923.

Sargent, T. J. (1987): *Macroeconomic Theory*. Academic Press, San Diego, CA, 2nd edn.

Schmitt-Grohé, S., and M. Uribe (2010): "Evaluating the Sample Likelihood of Linearized DSGE Models without the Use of the Kalman Filter," *Economics Letters*, 109(3), 142–143.

Sims, C. A. (2001): "Solving Linear Rational Expectations Models," *Computational Economics*, 20(1-2), 1–20.

——— (2012): "Statistical Modeling of Monetary Policy and Its Effects," *American Economic Review*, 102(4), 1187–1205.

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

Stephens, M. (2000): "Bayesian analysis of mixture models with an unknown number of componentsan alternative to reversible jump methods," *The Annals of Statistics*, 28(1), 40–74.

Tierney, L. (1998): "A note on Metropolis-Hastings kernels for general state spaces," *The Annals of Applied Probability*, 8(1), 1–9.

Troughton, P., and S. J. Godsill (1998): "A Reversible Jump Sampler for Autoregressive Time Series, Employing Full Conditionals to Achieve Efficient Model Space Moves," pp. 2257–2260.

Uhlig, H. (1999): "A Toolkit for Analysing Nonlinear Dynamic Stochastic Models Easily," in *Computational Methods for the Study of Dynamic Economies*, ed. by R. Marimon, and A. Scott, chap. 3, pp. 30–61. Oxford University Press.

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

V.V. Chari, P. J. K., and E. R. McGrattan (2008): "Are structural VARs with long-run restrictions useful in developing business cycle theory?," *Journal of Mone*, 55, 1337–1352.

Waagepetersen, R., and D. Sorensen (2001): "A Tutorial on Reversible Jump MCMC with a View toward Applications in QTL-mapping," *International Statistical Review*, 69, 49–61.

Watson, M. (1993): "Measures of Fit for Calibrated Models," *Journal of Political Economy*, 101, 1011–1041.

A Appendices

A.1 Detailed Derivation of Inflated Proposal Mapping

To choose an appropriate mapping $g_{pp'}$, it is useful to break the mapping into two parts according to the desired parameters P_p and the auxiliary parameters u. The mapping $g_{pp'}$ is given by

$$(P^{p'}, u') = g_{pp'}(P^p, u) = (g_{1pp'}(P^p, u), g_{2pp'}(P^p, u)) \tag{A-1}$$

and its inverse

$$(P^p, u) = g^{-1}_{pp'}(P^{p'}, u') = g_{p'p}(P^{p'}, u') = (g_{1p'p}(P^{p'}, u'), g_{2p'p}(P^{p'}, u') \tag{A-2}$$

Start with $g_{1pp'}$. Suppose now that the current state of the Markov chain is at $\varsigma = (p, P^p)$. Now with probability $\gamma_p(p'|p)$, a move to the model with order p' is proposed. Conditional on this proposal, we draw u from some proposal distribution $\gamma_{pp'}(u)$. Then, we introduce a deterministic mapping $g_{1pp'}$ that maps the current state and the auxiliary proposal u to the proposed new state such that $(p', P^{p'}) = (p', g_{1pp'}(P^p, u))$. Note that u is not part of the state of the chain.

Additionally, we have to find $g_{2pp'}$. In order to be able to easily verify adherence to detailed balance for a move from a state (p, P^p) to $(p', P^{p'}) = (p', g_{1pp'}(P^p, u))$ the vectors of Markov chain states and the random auxiliary proposal variables (P^p, u) and $(P^{p'}, u')$ must be of equal dimension and requiring $g_{pp'}$ to be a differentiable bijection lets us use a simple change-of-variables in the detailed balance equation. I.e., the kernel of the chain is now defined in terms of the auxiliary variable u together with the model indicator and the parameter vectors.

Armed with this structure it is now straightforward to derive the appropriate acceptance probability. The detailed balance condition holds if[26]

$$\int_{\mathcal{A}_p} \pi(p|y)\, \pi(P^p|p, y)\, Q\left(\varsigma, \mathcal{B}_{p'}\right) dP^p = \int_{\mathcal{B}_{p'}} \pi(p'|y)\, \pi\left(P^{p'}|p', y\right) Q\left(\varsigma', \mathcal{A}_p\right) dP^{p'} \tag{A-3}$$

for all subsets $\mathcal{A}_p$ and $\mathcal{B}_{p'}$ of the parameter spaces associated with autoregressive polynomials of order p and p' respectively. The posterior distribution $\pi(\varsigma|y)$ is factorized as $\pi(\varsigma|y) = \pi(p|y)\pi(P^p|p, y)$ and

$$\begin{aligned} Q\left(\varsigma, \mathcal{B}_{p'}\right) &= \int_{\mathcal{B}_{p'}} \gamma(\varsigma'|\varsigma)\alpha(\varsigma, \varsigma')d\varsigma' \\ &= \gamma_p(p'|p) \int \mathbb{1}(g_{1pp'}(P^p, u) \in \mathcal{B}_{p'})\alpha_{pp'}(P^p, g_{1pp'}(P^p, u)\gamma_{pp'}(P^p, u)du \end{aligned}$$

[26] See also Waagepetersen and Sorensen (2001).

The left hand side of (A-3) is then

$$\int_{\mathcal{A}_p} \pi(\varsigma|y)\, Q\left(\varsigma, \mathcal{B}_{p'}\right) dP^p = \int\int \mathbb{1}(P^p \in \mathcal{A}_p, g_{1pp'}(P^p,u) \in \mathcal{B}_{p'}) \pi(p|y)\, \pi(P^p|p,y) \times \tag{A-4}$$

$$\gamma_p(p'|p)\alpha_{pp'}(P^p, g_{1pp'}(P^p,u)\gamma_{pp'}(P^p,u) dP^p du \tag{A-5}$$

and the right hand side reads

$$\int_{\mathcal{B}_{p'}} \pi(\varsigma'|y)\, Q\left(\varsigma', \mathcal{A}_p\right) dP^{p'} = \int\int \mathbb{1}(P^{p'} \in \mathcal{B}_{p'}, g_{1p'p}(P^{p'},u') \in \mathcal{A}_p) \pi(p'|y)\, \pi\left(P^{p'}|p',y\right) \times \tag{A-6}$$

$$\gamma_p(p|p')\alpha_{p'p}(P^{p'}, g_{1p'p}(P^{p'},u'))\gamma_{p'p}(P^{p'},u') dP^{p'} du' \tag{A-7}$$

where $\gamma(\varsigma'|\varsigma)$ is again factorized as $\gamma_p(p|p)\gamma_{pp'}(P^p,u)$. The fact that $g_{pp'}$ is a differentiable bijection together with the dimension matching conditions enables a change of variable in (A-6) leading to

$$\int\int 1(g_{1pp'}(P^p,u) \in \mathcal{B}_{p'}, P^p \in \mathcal{A}_p) \pi(p'|y)\, \pi\left(g_{1pp'}(P^p,u)|p',y\right) \gamma_p(p|p')$$

$$\times \alpha_{p'p}(g_{1pp'}(P^p,u), P^p)\gamma_{p'p}(g_{1pp'}(P^p,u), g_{2pp'}(P^p,u))|g'_{pp'}(P^p,u)| dP^p du \tag{A-8}$$

where $dP^{p'} du' = |g'_{pp'}(P^p,u)| dP^p du$ and $|g'_{pp'}(P^p,u)|$ is the determinant of the Jacobian of $g_{pp'}$.

By inspection of (A-4) and (A-8), the reversibility condition (A-3) is satisfied if

$$\pi(p|y)\, \pi(P^p|p,y)\, \gamma_p(p'|p)\alpha_{pp'}(P^p, g_{1pp'}(P^p,u))\gamma_{pp'}(P^p,u) =$$

$$\pi(p'|y)\, \pi\left(g_{1pp'}(P^p,u)|p',y\right) \gamma_p(p|p')\alpha_{p'p}(g_{1pp'}(P^p,u), P^p) \times$$

$$\gamma_{p'p}(g_{1pp'}(P^p,u), g_{2pp'}(P^p,u))|g'_{pp'}(P^p,u)| \tag{A-9}$$

Choosing the acceptance probability as large as possible, we have

$$\alpha_{pp'} = \min\left(1, \chi_{pp'}(\varsigma,\varsigma')\right) \tag{A-10}$$

with

$$\chi_{pp'}(\varsigma,\varsigma') = \underbrace{\frac{\mathcal{L}(\varsigma')}{\mathcal{L}(\varsigma)}}_{\text{Likelihood Ratio}} \underbrace{\frac{\rho(\varsigma')}{\rho(\varsigma)}}_{\text{Prior Ratio}} \underbrace{\frac{\gamma_p(p|p')\gamma_{p'p}(g_{pp'}(P^p,u))}{\gamma_p(p'|p)\gamma_{pp'}(P^p,u)} |g'_{pp'}(P^p,u)|}_{\text{Proposal Ratio}} \tag{A-11}$$

With our mapping $g_{pp'}$, in (9), $|g'_{pp'}(P^p,u)|$ is equal to one and (A-11) reduces to (14).[27]

A.2 Solving for the Coefficients in the Recursive Solution

For the sequence of coefficients $\{\Phi_i\}_{i=0}^{\tilde{p}-1}$ that measure the impact of the exogenous processes in Z_t on X_t we need to solve (31) or (32) if $p = 0$. This set of equations can be rewritten by recursive

[27] The posterior π is here written factorized as the product of likelihood and prior $\mathcal{L}(\varsigma)\rho(\varsigma)$ for correspondence with the general formulation of the detailed balance condition (3).

substitution as[28]

$$\Phi_{p-i} = \begin{cases} \sum_{j=1}^{i} \left(-(B+A\Lambda)^{-1}A\right)^{j} \Phi_0 P_{p-i+j} & \text{for } i = 1, 2, \ldots p-1 \\ \sum_{j=1}^{i} \left(-(B+A\Lambda)^{-1}A\right)^{j} \Phi_0 P_{p-i+j} - (B+A\Lambda)^{-1} D & \text{for } i = p \end{cases} \tag{A-12}$$

where the invertibility of $B + A\Lambda$ follows from assumptions 4.3 and 4.4.[29] Thus, given Φ_0 from the $i = p$ case we can recover the remaining matrices Φ_i.

For $i = p$, (A-12) is

$$\Phi_0 = \sum_{j=1}^{p} \left(-(B+A\Lambda)^{-1}A\right)^{j} \Phi_0 P_j - (B+A\Lambda)^{-1} D \tag{A-13}$$

or

$$\Phi_0 + \sum_{j=1}^{p} \left(-(B+A\Lambda)^{-1}A\right)^{j} \Phi_0 \left(-P_j\right) = -(B+A\Lambda)^{-1} D \tag{A-14}$$

which is linear in Φ_0, being a p'th generalized Sylvester equation of the form

$$x + \beta x \gamma_1 + \beta^2 x \gamma_2 \ldots + \beta^p x \gamma_J = \delta \tag{A-15}$$

where $x \doteq \Phi_0$ and $\beta \doteq -(B+A\Lambda)^{-1}A$.[30]

Proposition A.1. *A generalized Sylvester equation of the form*

$$x + \beta x \gamma_1 + \beta^2 x \gamma_2 \ldots + \beta^J x \gamma_J = \delta \tag{A-16}$$

can be solved recursively for $\underset{n_a \times n_b}{x}$ *as follows*

$$\tilde{x}_{i,\bullet} \left(\sum_{j=0}^{J} \gamma_j U_{i,i}^{j} \right) = \delta_{i,\bullet} - \left(\sum_{k=1}^{n_a - i} \sum_{j=0}^{J} \{U^j\}_{i, n_a+k} \tilde{x}_{n_a+k, \bullet} \gamma_j \right), \quad \text{for } i = n_a, n_a - 1, \ldots, 1 \tag{A-17}$$

where $\tilde{x} \doteq Q^{\dagger} x$, $QUQ^{\dagger} = \beta$ *with* U *upper diagonal and* Q *unitary is the complex Schur decomposition*[31] *of* β, † *indicates conjugate transposition, and* $_{c,d}$ *references the c'th row and d'th column of a matrix.*

Proof. With the Schur decomposition $QUQ^{\dagger} = \beta$, (A-15) can be rewritten as

$$x + QUQ^{\dagger} x \gamma_1 + \left(QUQ^{\dagger}\right)^2 x \gamma_2 \ldots + \left(QUQ^{\dagger}\right)^J x \gamma_J = \delta \tag{A-18}$$

The matrix Q is unitary, so $Q^{\dagger} = Q^{-1}$ reducing the foregoing to

$$x + QUQ^{\dagger} x \gamma_1 + QU^2 Q^{\dagger} x \gamma_2 \ldots + QU^J Q^{\dagger} x \gamma_J = \delta \tag{A-19}$$

multiplying through with $Q^{\dagger}$ and using the definition $\tilde{x} \doteq Q^{\dagger} x$ gives

$$\tilde{x} + U\tilde{x}\gamma_1 + U^2 \tilde{x} \gamma_2 \ldots + U^J \tilde{x} \gamma_J = Q^{\dagger} \delta \tag{A-20}$$

[28] Starting with the last equation of (31). It is already in this form. Then proceed to the second-to-last equation and eliminate Φ_{p-1} in this equation using the last equation. Proceed thusly to the first equation.

[29] See Lan and Meyer-Gohde (2012).

[30] For completeness, $\gamma_j \doteq -P_j$, for $j = 1, 2, \ldots, p$ and $\delta \doteq -(B+A\Lambda)^{-1} D$.

[31] See, e.g., Golub and Loan (1996).

As U is upper diagonal, so is any power of U; thus given all rows of the matrix $\tilde{x}$ after some i, the i'th row of $\tilde{x}$, $\tilde{x}_{i,\bullet}$ solves

$$\sum_{j=0}^{J} U_{i,i}^{j} \tilde{x}_{i,\bullet} \gamma_j = \delta_{i,\bullet} - \left(\sum_{k=1}^{n_a - i} \sum_{j=0}^{J} \{U^j\}_{i,n_a+k} \tilde{x}_{n_a+k,\bullet} \gamma_j \right) \tag{A-21}$$

recognizing that $U_{i,i}$ is a scalar gives (A-17) which can be solved by multiplying on the right by the inverse of $\left(\sum_{j=0}^{J} \gamma_j U_{i,i}^{j} \right)$. □

Given Φ_0, the remaining sequence of coefficients $\{\Phi_i\}_{i=1}^{p-1}$ can be recovered recursively from (31) starting with Φ_{p-1} and working backwards to Φ_1. Likewise, given Φ_0, the sequence of coefficients $\{\Theta_i\}_{i=0}^{q-1}$ an be recovered recursively from (30) starting with Θ_{q-1} and working backwards to Θ_0.

Variable	Prior	Proposal
p	U(0,10)	LaplaceD(p,2)
q	U(0,10)	LaplaceD(q,2)
AR PAC	TN(0,0.25)	TN(PAC,0.0025)
MA inverse PAC	TN(0,0.25)	TN(PAC,0.0025)
σ: Standard Deviation ϵ_t	IG(1,1)	TN(σ,0.0025)

Table 1: Prior and Proposal Distribution for Monte Carlo Experiment

Method	Proportion of correctly identified models
RJMCMC	0.50
AIC	0.36
AICC	0.44
BIC	0.71

Table 2: Proportion of Correctly Identified Models

$\overline{L}$	$\frac{1}{3}$	Steady state employment 1/3 of total time endowment
α	0.36	Capital share
δ	0.025	Depreciation rate for capital
$\overline{R}$	1.01	One percent real interest rate per quarter

Table 3: Model Calibration

Variable	Prior	Proposal
p	U(0,10)	LaplaceD(p,2.2)
q	U(0,10)	LaplaceD(q,2.2)
AR PAC	TN(0,0.25)	TN(PAC,0.0016)
MA PAC	TN(0,0.25)	TN(PAC,0.0016)
σ	IG(1,1)	TN(σ,0.0025)

Table 4: Priors and Proposals for RBC Model Estimation

Parameter	Conditional Mean	Conditional Median	Hansen
AR(1)	1.1689 (0.04)	1.1681	0.95
AR(2)	-0.0732 (0.06)	-0.0725	N/A
AR(3)	-0.1224 (0.04)	-0.1215	N/A
σ	0.5873 (0.08)	0.5733	0.712
Autocorr(1)	0.9804	0.9810	0.95
Autocorr(2)	0.9528	0.9542	0.9025

Table 5: Posterior Point Parameter Estimates Conditional on $(p, q) = (3, 0)$; Standard Errors in Parentheses

Data	Hansen	Posterior Mode Model	Posterior Mode	90% Posterior Credible Set
2.8491	3.2574	2.8332	2.8182	2.1074 — 4.0965

Table 6: Standard Deviation of Output, in %

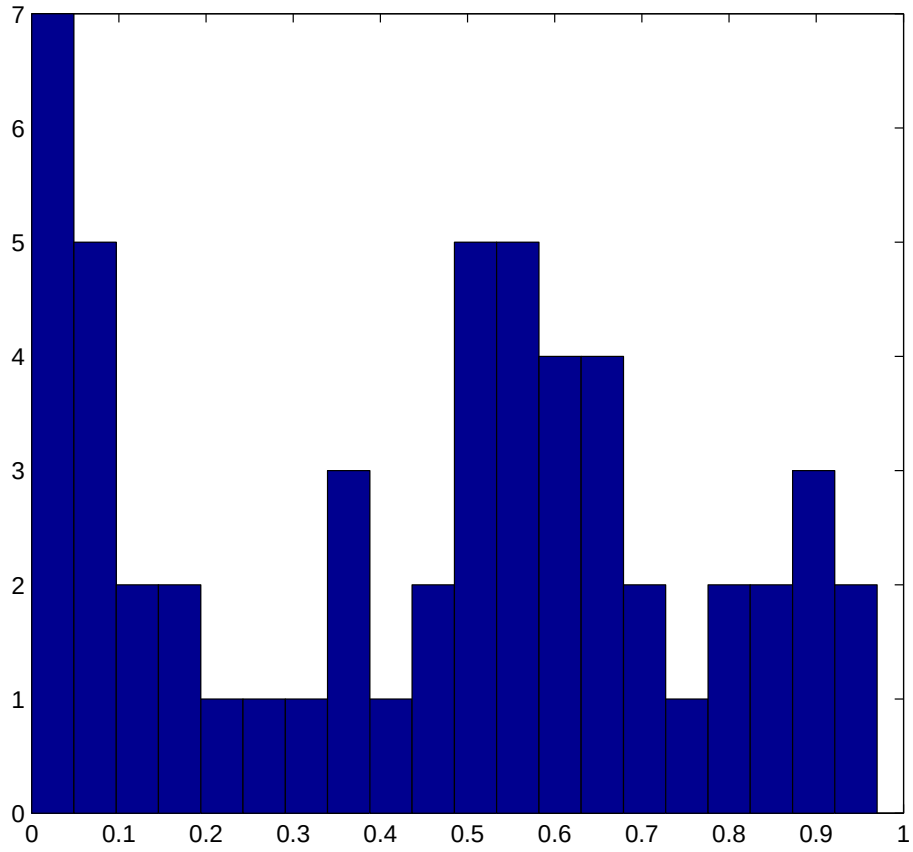


Figure 1: Histogram of Posterior Mass Assigned to True Model Relative to Model at the Mode, Model Incorrectly Identified

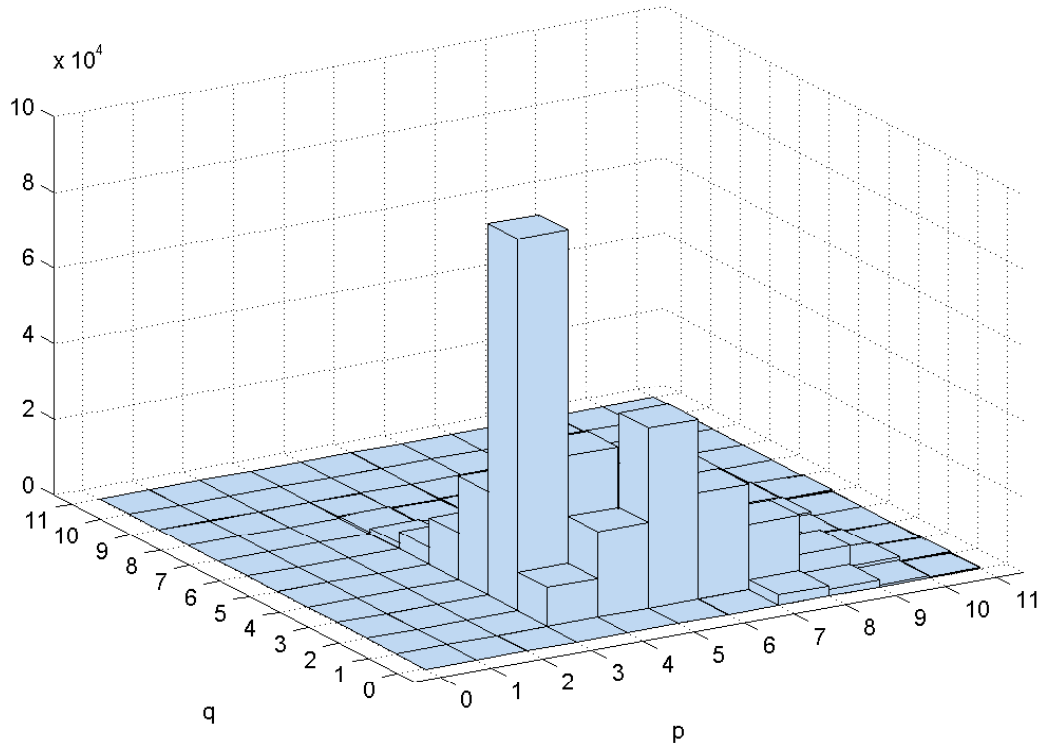


Figure 2: Typical Posterior Distribution, Model Correctly Identified

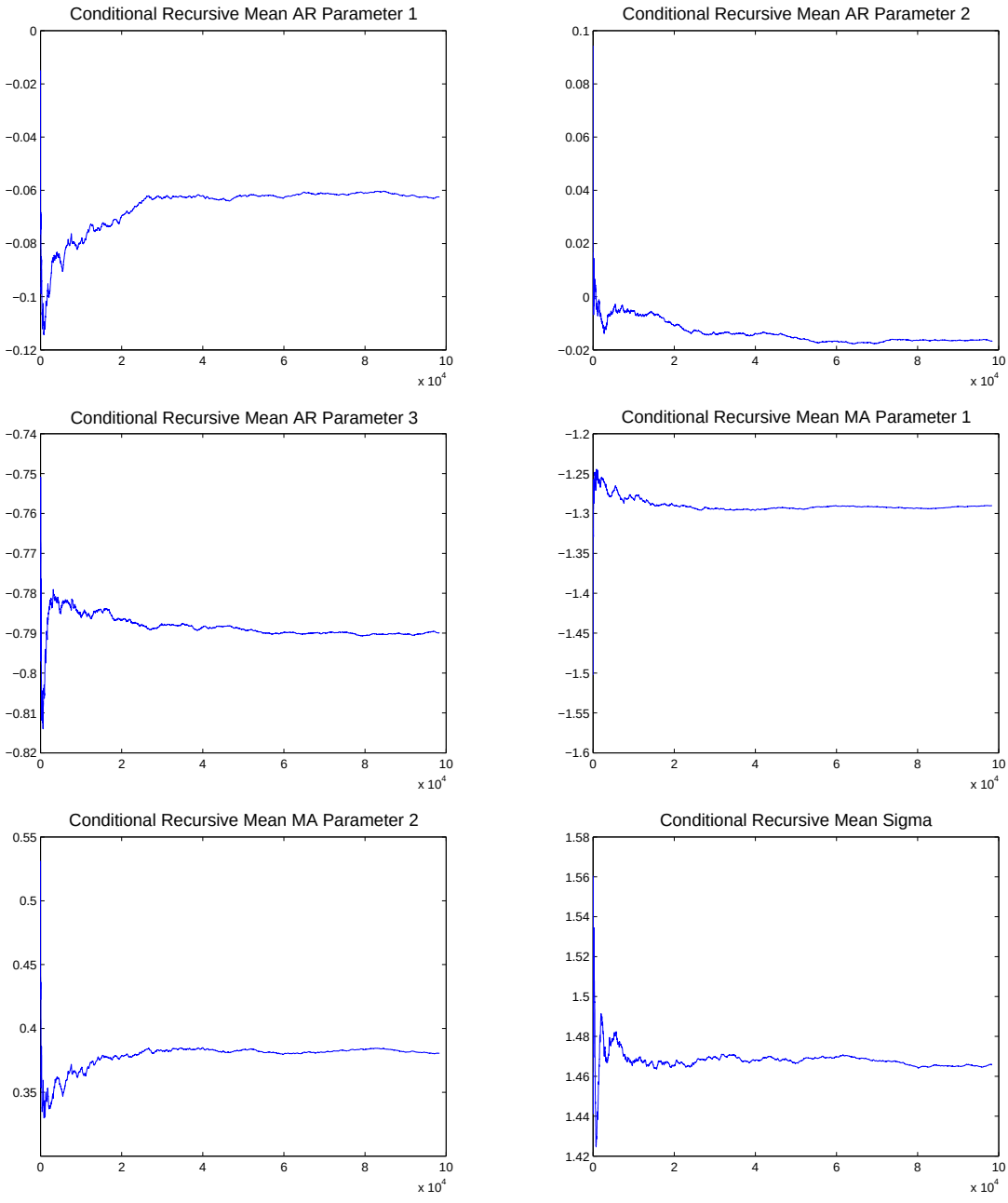


Figure 3: Recursive Parameter Means from the Conditional Posterior

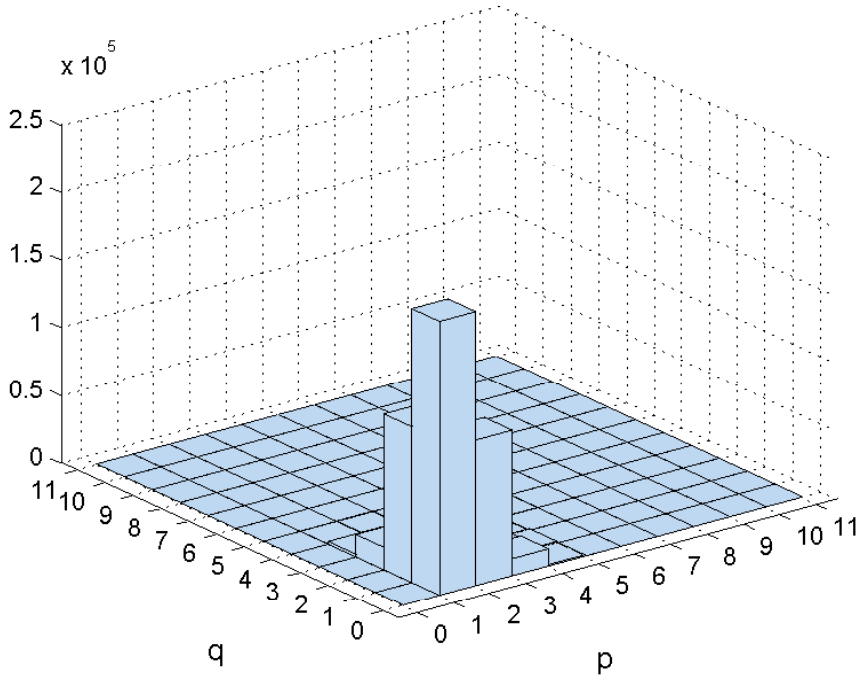


Figure 4: Posterior over the Orders for the Shock Process for Synthetic Data Generated with the Shock Process from Equation (52)

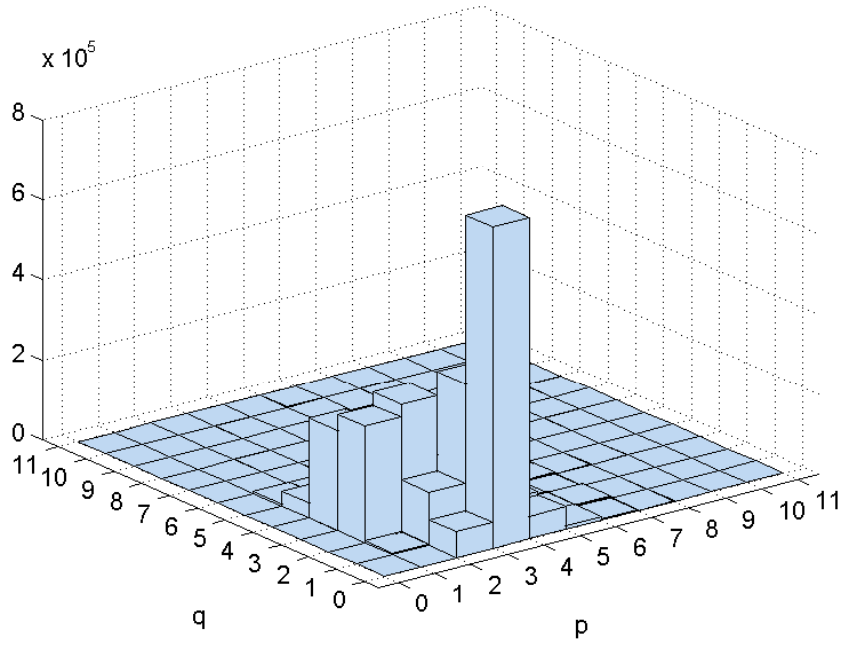
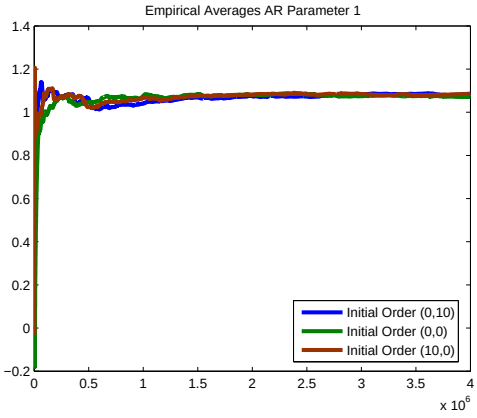


Figure 5: Posterior over the Orders for the Shock Process

Unconditional



Conditional

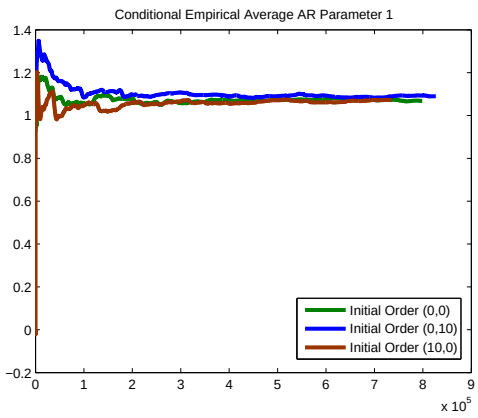


Figure 6: Convergence Diagnostics

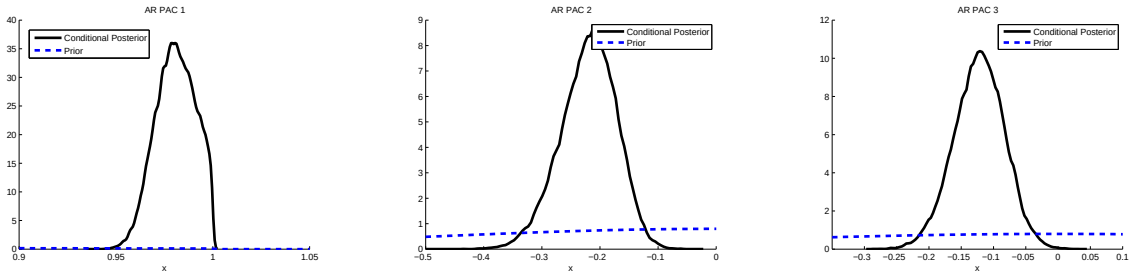


Figure 7: Priors and Posteriors for (Inverse) Partial Autocorrelations

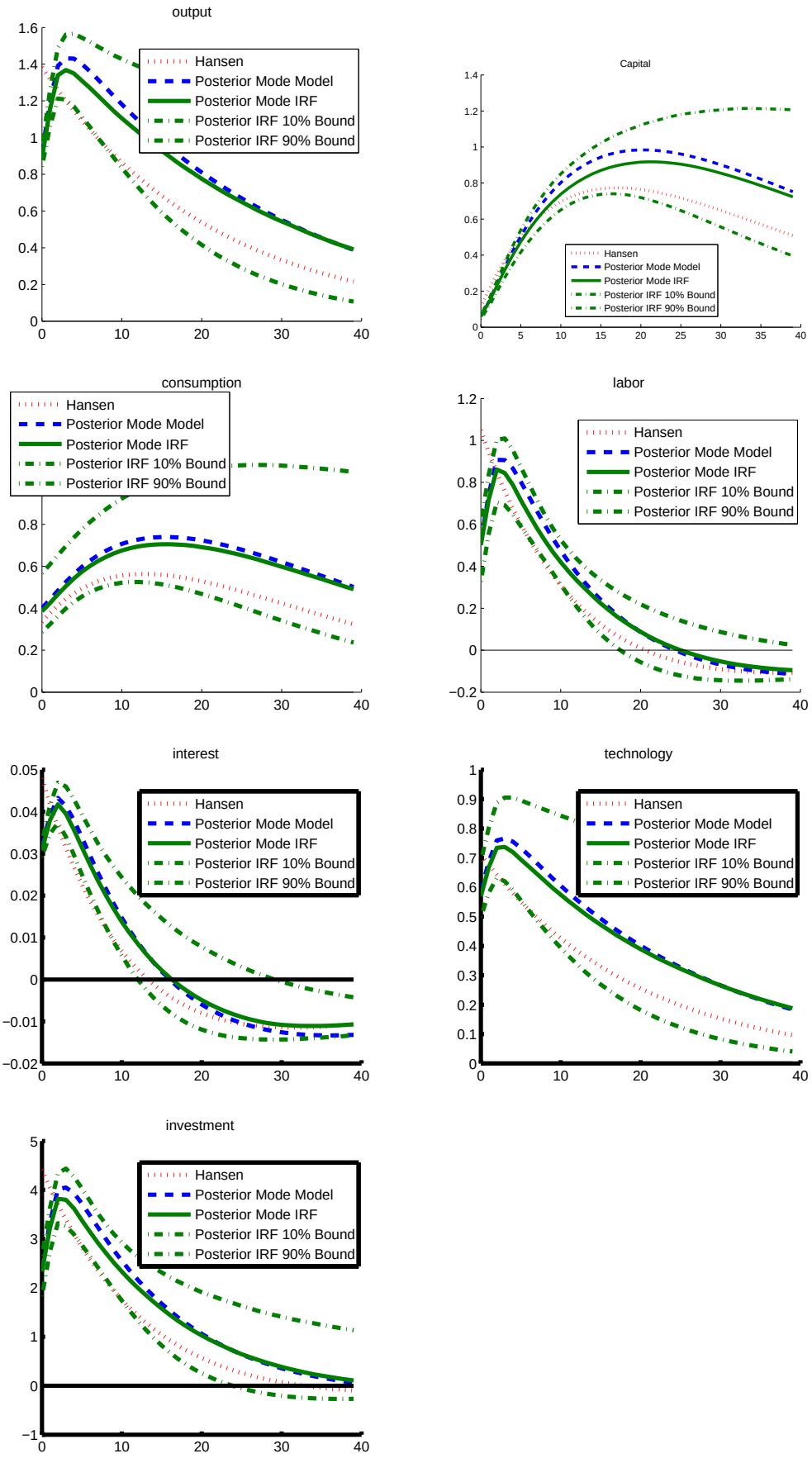


Figure 8: Impulse Responses to a One Standard Deviation Technology Shock
Invertibility of MA Components Imposed

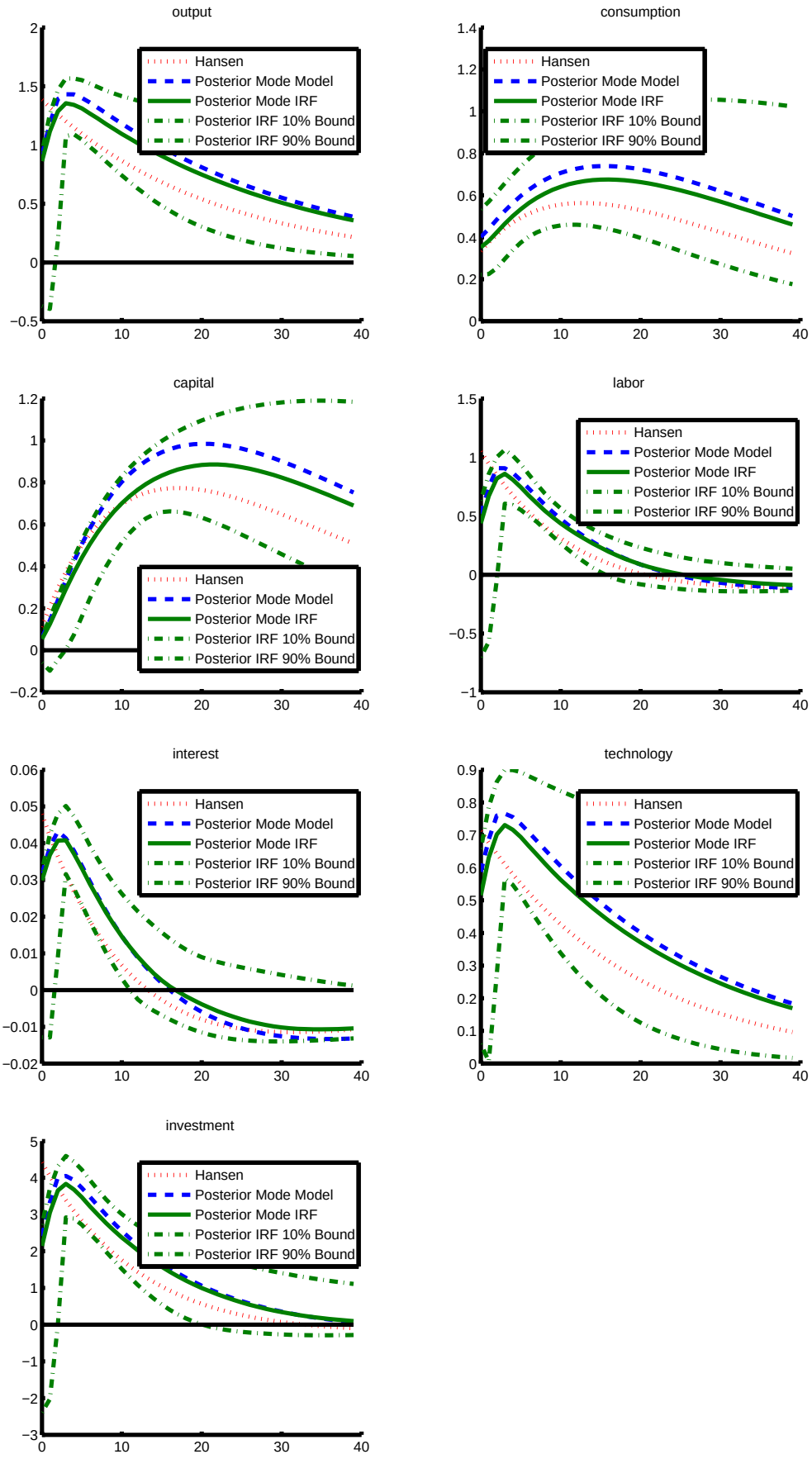


Figure 9: Impulse Responses to a One Standard Deviation Technology Shock
Invertibility of MA Components Not Imposed

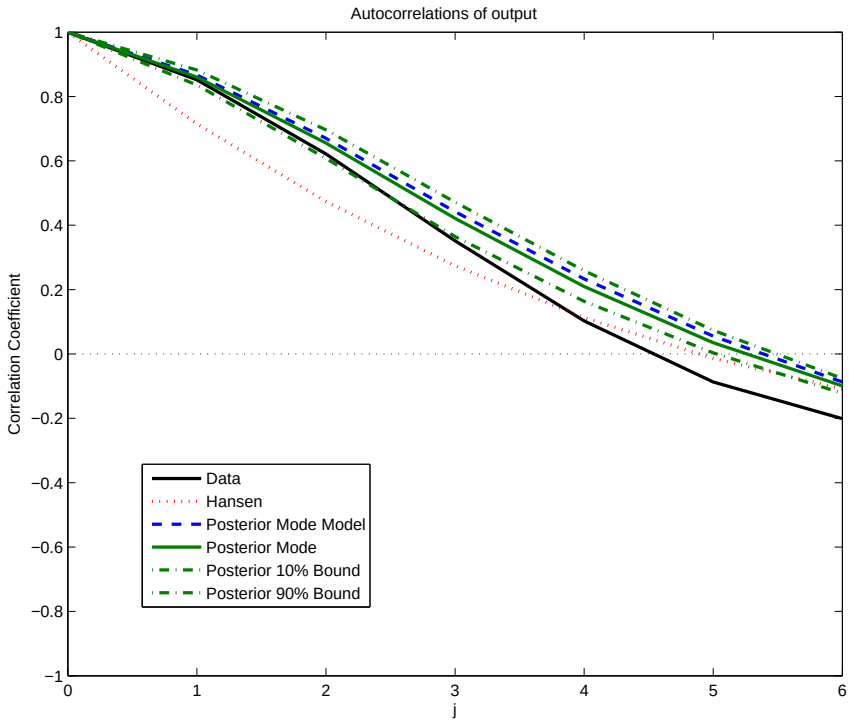


Figure 10: Comparison of Autocorrelations of Output