Approximating time varying structural models with time invariant structures

Fabio Canova, EUI and CEPR
Filippo Ferroni Banque de France and University of Surrey
Christian Matthes Federal Reserve Bank of Richmond *

May 14, 2015

Abstract

The paper studies how parameter variation affects the decision rules of a DSGE model and structural inference. We provide diagnostics to detect parameter variations and to ascertain whether they are exogenous or endogenous. Identification and inferential distortions when a constant parameter model is incorrectly assumed are examined. Likelihood and VAR-based estimates of the structural dynamics when parameter variations are neglected are compared. Time variations in the financial friction of a Gertler and Karadi (2010) model are studied.

Key words: Structural model, time varying coefficients, endogenous variations, misspecification.

JEL Classification: C10, E27, E32.

*We thank Tao Zha, Ferre de Graeve and the participants to seminars at Goethe University, University of Milan, Bank of England, Carlos III Madrid, Humboldt University Berlin, Federal Reserve Board, and the conferences ESSIM 2015; Identification in Macroeconomics, National Bank of Poland; Econometric Methods for business cycle analysis, forecasting and policy simulations, Norge Bank for comments and suggestions. The views presented in this paper do not reflect those of the Banque de France and the Federal Reserve system.

1 Introduction

In macroeconomics, it is standard to study models that are structural in the sense of Hurwicz (1962); that is, models where the parameters characterizing the preference and the constraints of the agents and the technologies to produce goods and services are invariant to changes in the parameters describing government policies. Such a requirement is fundamental to distinguish structural models from reduced form ones, and it is crucial to conduct correctly designed policy counterfactuals in dynamic stochastic general equilibrium (DSGE) models.

Recently, Dueker et al. (2007), Fernandez Villaverde and Rubio Ramirez (2007), Canova (2009), Rios Rull and Santaeularia Llopis (2010), Liu et al. (2011), Galvao, et al. (2014), Vavra (2014), Seoane (2014), Meier and Sprengler (forthcoming), have shown that DSGE parameters are not time invariant and variations are consistent with Stock and Watson' (1996) idea that reduced economic relationships display small but persistent evolutions. Parameter variations can not be taken as direct evidence that DSGE models are not structural. For example, Cogley and Yagihashi (2010), and Chang et al. (2013) showed that parameter variations may result from the misspecification of a model with time invariant structure, while Schmitt Grohe and Uribe (2003) indicated that parameter variations may be needed in certain small open economy models to insure the existence of a stationary equilibrium.

The approach the DSGE literature has taken to model parameter variations follows the VAR literature, (see Cogley and Sargent, 2005 and Primiceri, 2005): parameters are assumed to be exogenously drifting as independent random walks, see e.g. Fernandez and Rubio (2007). Many economic questions, however, hint at the possibility that parameter variations may instead be endogenous. For example, is it reasonable to assume that the Federal Reserve reacts in the same way to inflation in an expansion or in a contraction? Davig and Leeper (2006) analyze a situation where the policy rule could be state dependent and describe how it affects the dynamics induced by different structural shocks. Does the propagation of shocks depend on the state of private and government debt? Do fiscal multipliers depend on inequality, see e.g. Brinca et al. (2014)? Are household as risk averse or as impatient when they are wealthy as when they are poor? Questions of this type are potentially numerous. Clearly, policy analysis derived assuming time invariant parameters or an inappropriate form of time variations may be misleading; comparisons of the welfare costs of business cycles may be biased; growth prescriptions may be invalid; and standard business cycle tools may provide a distorted picture of the dynamics.

This paper has a number of goals. First, we want to characterize the decision rules of a DSGE when parameter variations are either exogenous or endogenous, and in the latter case, when agents internalize or not the effects that their decisions may have on parameter variations. Second, we wish to provide diagnostics to detect misspecifications due to neglected parameter variations. Third, we are interested in studying the consequences of using time invariant models when the DGP features parameter variations in terms of identification, estimation, and inference.

In particular, we wish to compare likelihood-based and SVAR-based estimates of the structural dynamics to shed light on whether approaches that take a less structural approach are competitive when parameter variations are neglected.

The existing literature is generally silent on the issues of interest in this paper. Seoane (2014) is the closest, in the sense that parameter variations are used to gauge potential model misspecifications. Kulish and Pagan (2014) characterize the decision rules of a DSGE model when structural breaks, which are partly predictable, occur. Magnusson and Mavroedis (2014) and Huang (2014) examine how time variations in the certain parameters may affect the identification of other structural parameters, the asymptotic theory of maximum likelihood estimators, and standard break tests. Andreasen (2012) studies how time variations in the variance of the exogenous shocks affect risk premia in models approximated with higher order perturbations and Fernandez et al. (2013) investigate to what extent variations in shock volatility matter for real variables. Ireland (2007) assumes that trend inflation in a standard New Keynesian model is driven by structural shocks and estimates the model by likelihood techniques; Ascari and Sbordone (2014) highlight that trend inflation may be a function of policy decisions.

The next section characterizes the decision rules in a general setup where both exogenous and endogenous variations in the parameters regulating preferences, technologies, and constraints are possible. We consider both first order approximations and higher order perturbed solutions. We present a simple RBC example to illustrate our results and to provide some intuition for the outcomes we obtain.

We show that if parameter variations are exogenous, structural dynamics are the same as in a model with no parameter variations. Thus, if one correctly identifies structural disturbances, she would make no mistakes in characterizing structural impulse responses even if she employes a constant coefficient model. Clearly, variance and historical decompositions exercises will be distorted, since some sources of disturbances will be omitted. If parameter variations are instead endogenous, structural dynamics may be different from the one of a constant coefficient model. The extent of the differences in the two specifications depends on two matrices entering the decision rules. When time variations affect structural dynamics, they do so because they alter income and substitution effects present in the constant coefficient model. The conclusions obtained with a first order approximately solution carry over to a second order approximation, but not to higher order approximations.

Section 3 provides two diagnostics, based on the optimality wedges of Chari et al. (2007) and on forecast errors, which can help to detect misspecification induced by neglecting parameter variations. We also describe a marginal likelihood diagnostics which can be employed to distinguish exogenous vs. endogenous parameter variations.

Section 4 deals with parameter identification. We are interested in measuring the identification repercussions that neglected time variations may have for time invariant parameters. Since the likelihood is constructed using forecast errors, which are generally misspecified when parameter variations are neglected, one expects the likelihood shape to be both flattened and distorted. In the context of the RBC example, we show that indeed both pathologies occur. Overall, our conclusions agree with Huang (2014) in the sense that weakly identified (time invariant) parameters do not become better identified when time variations in other parameters exist.

Section 5 considers the structural estimation of a model with time invariant parameters when the data is generated by models with time varying parameters. We expect distortions because the dynamics assumed by the constant coefficient model are generally incorrect and shocks misaggregation is present. Indeed, we find that important biases in parameter estimates are present, that they are most important in parameters controlling income and substitution effects, and that they do not die away as sample size increase. Estimated impulse responses differ from the true ones both in quantitative and qualitative sense, and technology shocks tend to absorb the variability contribution of missing shocks.

Section 6 studies the performance of less structural time invariant VAR model as far as the dynamics induced by structural shocks are concerned. We show that important qualitative features (impact effect, shape and persistence) of the structural dynamics are well captured. As with likelihood based methods, the performance of SVAR worsens when shocks to the parameters account for a considerable portion of the variability of the endogenous variables even though the deterioration is not as large as with more structural approaches.

Section 7 estimates the parameters of Gertler and Karadi (2010) model of unconventional monetary policy, applies the diagnostics to detect parameter variations, and estimates versions of the model where bank's moral hazard parameter is allowed to vary over time. We find evidence that a fixed coefficient model is misspecified, that making parameter variations endogenous function of net worth is preferable, and that the dynamic effects of capital quality shocks on the spread and on bank net worth can be more persistent than previously thought. Section 8 concludes.

2 The setup

The optimality conditions of a DSGE model can be represented as:

$$E_t\left[f(X_{t+1}, X_t, X_{t-1}, Z_{t+1}, Z_t, \Theta_{t+1}, \Theta_t)\right] = 0 \tag{1}$$

where X_t is an $n_x \times 1$ vector of endogenous variables, Z_t is an $n_z \times 1$ vector of strictly exogenous variables, $\Theta_t = [\Theta_{1t}, \Theta_{2t}]$, is a vector of possibly time varying structural parameters, where Θ_{2t} is a $n_{\theta_1} n_{x_1} \times 1$ vector, $n_x \geq n_{x_1}$, appearing in the case agents internalize the effects that their decisions have on the parameters and Θ_{1t} is an $n_{\theta_1} \times 1$ vector, while f is a continuous function, assumed to be differentiable up to order q, mapping onto a $\mathbb{R}^{n_x}$ space. The law of motion of the Z_{t+1} variables is given by

$$Z_{t+1} = \Psi(Z_t, \sigma \Sigma_\epsilon \epsilon^z_{t+1}) \tag{2}$$

where Ψ is a continuous function, assumed to be differentiable up to order q, mapping onto a $\mathbb{R}^{n_z}$ space; ϵ^z_{t+1} is a $n_e \times 1$ vector of i.i.d. structural disturbances with mean zero and identity covariance matrix, $n_z \geq n_e$; $\sigma \geq 0$ is an auxiliary scalar, Σ_ϵ is a known $n_e \times n_e$ matrix.

The law of motion of the structural parameters Θ_t is given by

$$\Theta_{t+1} = \Phi(\Theta, X_t, U_{t+1}) \tag{3}$$

where Φ is a continuous function, assumed to be differentiable up to order q, mapping onto the $\mathbb{R}^{n_\theta}$ space; U_t is a $n_u \times 1$ vector of exogenous disturbances, $n_\theta = n_{\theta_1}(1+n_{x_1}) \geq n_u$; Θ is a vector of constants. The law of motion of U_{t+1} is given by

$$U_{t+1} = \Omega(U_t, \sigma\Sigma_u\epsilon^u_{t+1}) \tag{4}$$

where Ω is continuous and differentiable up to order q, mapping onto the $\mathbb{R}^{n_u}$ space; ϵ^u_t is a $n_u \times 1$ vector of i.i.d. disturbances, with mean zero and identity covariance matrix, uncorrelated with the ϵ^z_{t+1},and Σ_u is a known $n_u \times n_u$ matrix.

The decision rule of the problem is written as

$$X_t = h(X_{t-1}, Z_t, U_t, \sigma\Sigma\epsilon_{t+1}, \Theta) \tag{5}$$

where h is a continuous function, assumed to be differentiable up to order q, and mapping onto a $\mathbb{R}^{n_x}$ space, $\epsilon_{t+1} = [\epsilon^{z\prime}_{t+1}, \epsilon^{u\prime}_{t+1}]'$, $\Sigma = diag[\Sigma_z, \Sigma_u]$.

Few features of the setup need some discussion. First, Θ_t will be serially correlated if U_t is serially correlated. Second, the vector of structural disturbances ϵ^z_{t+1} may be smaller than the vector of exogenous variables and the dimension of ϵ^u_{t+1} may be smaller than the dimension of the structural parameters. Thus, there may be common patterns of variations in the Z_{t+1} and U_{t+1}. Third, we allow for time variations in the parameters regulating preferences, technologies and constraints but we not consider variations in the auxiliary parameters regulating the law of motion of the Z_t and the U_t, as we are not interested in stochastic volatility, GARCH or rare events phenomena (as in e.g. Andreasan, 2012), nor in time variations driven by evolving persistence of the exogenous processes. Fourth, (1) makes no distinction between states and controls. Thus, (5) has the format of a final form (endogenous variables as a function of the exogenous variables and the parameters) rather than of a state space form (control variables as a function of the states and of the parameters).

2.1 First order approximate decision rule

We start by studying the implications of structural parameters variation for the decision rule when a first order approximate solution is considered. Taking a linear expansion of (1) around the steady states leads to

$$0 = E_t\left[Fx_{t+1} + Gx_t + Hx_{t-1} + Lz_{t+1} + Mz_t + N\theta_{t+1} + O\theta_t\right] \tag{6}$$

where $F = \partial f/\partial X_{t+1}$, $G = \partial f/\partial X_t$, $H = \partial f/\partial X_{t-1}$, $L = \partial f/\partial Z_{t+1}$, $M = \partial f/\partial Z_t$, $N = \partial f/\partial \Theta_{t+1}$ $O = \partial f/\partial \Theta_t$, all evaluated at the steady states values of (X_t, Z_t, Θ_t)

and lower case letters indicate deviations of the variables from the steady states. Taking a linear expansion of (5) leads to

$$x_t = Px_{t-1} + Qz_t + Ru_t \tag{7}$$

where $P = \partial h/\partial X_{t-1}$, $Q = \partial h/\partial Z_t$, $R = \partial h/\partial U_t$, all evaluated at steady state values.

Proposition 2.1. *The matrices P, Q, R satisfy the following:*

- *P solves $FP^2 + (G + N\phi_x)P + (H + O\phi_x) = 0$.*
- *Given P, Q solves $VQ = -vec(L\psi_z + M)$ and $V = \psi_z' \otimes F + I_{n_z} \otimes (FP + G + N\phi_x)$ where vec denotes the columnwise vectorization.*
- *Given P, R solves $WR = -vec(N\phi_u\omega_u + O\phi_u)$ where $W = \omega_u' \otimes F + I_{n_\theta} \otimes (FP + G + N\phi_x)$*

where $\phi_u = \partial\Phi/\partial U_{t+1}, \phi_x = \partial\Phi/\partial X_t, \psi_z = \partial\Psi/\partial Z_t, \omega_u = \partial\Omega/\partial U_t$ and we assume that all the eigenvalues of ψ_z and of ω_u are strictly less than one in absolute value.

Proof. The proof is straightforward. Substituting (7) into (6), we obtain

$$\begin{aligned} 0 =& [FP^2 + (G + N\phi_x)P + (H + O\phi_x)]x_{t-1} + [(FP + G + N\phi_x)Q + FQ\psi_z + L\psi_z + M]z_t \\ &+ [(FP + G + N\phi_x)R + FR\omega_u + N\phi_u\omega_u + O\phi_u]u_t \end{aligned}$$

Since the solution must hold for every realization of x_{t-1}, z_t, u_t, we need to equate their coefficient to zero and the result obtains. □

Corollary 2.2. *If $\phi_x = 0$, the dynamics in response to the structural shocks z_t are identical to those obtained when parameters are time invariant. Variations in the j-th parameter have instantaneous impact on the endogenous variables x_t, if and only if the j^{th} column of $N\phi_u\omega_u + O\phi_u \neq 0$.*

Corollary 2.3. *If $\phi_u = 0$ and the matrices $N\phi_x$ and $O\phi_x$ are zero, variations in the j-th parameter have no dynamic effects on the endogenous variables x_t.*

Proposition 2.1 indicates that the first order approximate decision rule will, as in a constant coefficient setup, be a VARMA(1,1). Note that disturbances to the parameters play the role of additional structural shocks, making a model with m structural shocks and a model with m_1 structural shocks and m_2 disturbances to the parameters, $m = m_1 + m_2$, potentially indistinguishable.

Corollaries 2.2 and 2.3 give conditions under which parameter variations alter the dynamics induced by structural disturbances. If parameter variations are purely exogenous, $\phi_x = 0$, the P and Q matrices will be identical to those of a constant coefficient model. The intuition for the result is simple: as long as the shocks to Z_t and Θ_t are uncorrelated, parameter variations adds variability to the endogenous variables without altering the dynamics produced by structural disturbances. In other words, suppose an economy is perturbed by a technology shocks. Then, the dynamics induced by these shocks do not depend on whether the discount factor is constant or time varying, provided that its innovations are exogenous and unrelated to the innovations in the discount factor.

This implies that if one considers a time invariant version of the model and is able to identify the structural disturbances ϵ_t^z, she would make no mistakes in characterizing the dynamics in response to these shocks. Clearly, variance or historical decompositions exercises will be distorted, since certain sources of variations (the ϵ_t^u disturbances) are omitted. One interesting question is whether standard procedures allow a researcher employing a time invariant model to recover ϵ_t^z from the data when the DGP features time varying structural parameters. When this is not the case, one would like to know which structural disturbance absorbs the missing shocks.

On the other hand, if parameter variations are purely endogenous, $\phi_u = 0$, the dynamics in response to structural shocks may be altered. To know if distortions are present one needs to check whether the columns of the matrices $N\phi_x$ and $O\phi_x$are equal to zero. If they are not, a researcher employing a time invariant version of the model is likely to incorrectly characterize both the structural dynamics and the relative importance of different sources of disturbances for the variability of the endogenous variables.

Finally, note that exogenous variations in the structural parameters typically have an instantaneous effect on the endogenous variables, i.e. R$\neq$ 0, but purely endogenous parameter variations do not. Thus, to avoid identification problems, it seems preferable to have the relationship between parameters and the endogenous variables disturbed by shocks.

2.2 Higher order approximate decision rule

Are the conclusions are affected when higher order approximations are considered? In the second order approximation, the first order term are the same as in the linear approximation. To examine whether quadratic terms will be affected by the presence of time variations let $W_t = [Z_t', U_t']'$ so that, taking into account (5) ,(1) is

$$0 = E_t[F(X_t, W_t, \sigma\Sigma\epsilon_{t+1}, \Theta)] \tag{8}$$

The second order approximation of (8) is

$$\begin{aligned} E_t[(F_x x_{t-1} + F_w w_t + F_\sigma \sigma) + 0.5(F_{xx}(x_{t-1}\otimes x_{t-1}) + F_{ww}(w_t \otimes w_t) + F_{\sigma\sigma}\sigma^2) \quad + \\ F_{xw}(x_{t-1}\otimes w_t) + F_{x\sigma}x_{t-1}\sigma + F_{w\sigma}w_t\sigma] \quad = 0 \end{aligned} \tag{9}$$

Note that $F_\sigma\sigma$, $F_{x\sigma}x_{t-1}\sigma, F_{w\sigma}w_t\sigma$ are all zero, see Schmitt Grohe and Uribe (2004). The second order expansion of (5) is

$$\begin{aligned} x_t &= h_x x_{t-1} + h_w w_t + 0.5(h_{xx}(x_{t-1}\otimes x_{t-1}) + h_{ww}(w_t\otimes w_t) + h_{\sigma\sigma}\sigma^2) \\ &+ h_{xw}(x_{t-1}\otimes w_t) + h_{x\sigma}x_{t-1}\sigma + h_{w\sigma}w_t\sigma \end{aligned} \tag{10}$$

Proposition 2.4. *Time variations in the structural parameters affect h_{yy}, h_{ww} and h_{yw} the if only if they affect h_y and h_w.*

Proof. The proof is straightforward. Collecting terms and requiring, for example, $F_{xx} = 0$ implies that $h_{xx} = -(F_{x'})^{-1}J_1$ where

$$J_1 = 0.5F_{x'x'}(h_x \otimes h_x) + F_{x'x}h_x + F_{x'w'}h_x h_x + 0.5F_{xx} + F_{xw'}h_x + 0.5F_{w'w'}h_x {h_x}' \quad (11)$$

where primes indicate future values. Thus h_{xx} depends only on the first order term h_x, the matrix of derivatives of the optimality conditions with respect to the argument evaluated at the steady state ($F_{x'x'}, F_{w'w'}$, etc.) and h_y. Similarly $F_{ww} = 0$ implies that $g_{ww} = -(F_{x'})^{-1}J_2$ where

$$J_2 = 0.5F_{x'x'}(h_w \otimes h_w) + F_{x'w'}h_x g_w + F_{xw}g_w + 0.5F_{w'w'}h_w {h_w}' + F_{w'w}h_w + 0.5F_{ww} \quad (12)$$

which also depends on first order terms (x_w),the matrix of derivatives of the optimality conditions with respects to the argument evaluated at the steady state ($F_{x'x'}, F_{w'w'}$, etc.) and h_w.The proof for the other terms of the expression is analogous. □

The intuition for proposition 2.4 is simple: since second order terms are function of first order terms, of the parameters of the law of motion of the W$_t$, and of the gradient and of the Hessian of the optimality conditions evaluated at the steady states, they do not feature independent variations. Thus, the second order structural dynamics will be altered by time variations only if the first order dynamics are.

For higher order approximate solutions, the dynamics induced by structural shocks in constant coefficients and time varying coefficients models will generally differ even with exogenous time variations. For example, in a third order approximation, the optimality condition will feature terms in $F_{x\sigma\sigma}$ and $F_{w\sigma\sigma}$, which require a correction of the linear terms to account for uncertainty. Since in the constant coefficients model some shocks are omitted, one should expect the correction terms to differ in constant and time varying coefficient models.

2.3 Discussion

The results derived in this section require parameter variations to be continuous. This is in line with the evidence produced by Stock and Watson (1996) and with the standard practice employed in time varying coefficient SVAR. Note that our framework is flexible and can accommodate once-and-for-all breaks (at a known date) as long as transition between states is smooth. For example, a smooth threshold switching specification such as $\theta_{t+1} = (1-\rho)\theta + \rho\theta_t + a\exp(t - T_0)/(b + exp(t - T_0))$, $t = 1, \ldots, T_0 - 1, T_0, T_0 + 1, \ldots T$, where a and b are vectors is an acceptable form of exogenous time variations. Similarly, endogenous forms of time variations of the type $\theta_{t+1} = (1-\rho)\theta + \rho\theta_t + a\exp(-(X_t - X))/(b + exp(-(X_t - X))$, where X is the steady state value of X_t can also be used. What the framework does not allow for are Markov switching variations which occur at unknown dates or abrupt changes, such as those considered in Davig and Leeper (2006), since the smoothness conditions on the f function may not hold.

Kulish and Pagan (2014) have developed solution and estimation procedures for models with abrupt breaks and learning between the states. Their solution for the pre and post break period is a constant coefficient VAR, while for the learning period is a time varying coefficient VAR. Since (7) is a constant coefficient VAR with an extended set of shocks, a few words distinguishing the two approaches are needed. First, they are interested in characterizing the solution during the learning period, when the structure is unchanged but expectations move, while we are interested in the solution when parameters are continuously varying. Second, their modelling of time variations is abrupt and the solution is designed to deal with that situation. Third, in our setup expectations are varying with the variations of the structure, while Kulish and Pagan have expectations varying only in anticipation of a (foreseeable) break.

One way of thinking about the differences between exogenous and endogenous parameter variations is that in the former each parameter evolves independently and covariations, if they exist, can be modelled by selecting the matrix Σ_u to be of reduced rank. With endogenous variations, there is an observable factor (the X's) which drives the common parameter variations. Thus, Σ_u can be made diagonal and will feature reduced rank if for some parameter variations are purely endogenous.

2.4 An example

To convey some intuition into the mechanics of corollaries 2.2-2.3, we use a simple, closed economy, RBC model. The representative agent maximizes the discounted stream of future utilities given by

$$\max E_0 \sum_{t=1}^{\infty} \beta_t \left(\frac{C_t^{1-\eta}}{1-\eta} - A\frac{N_t^{1+\gamma}}{1+\gamma}\right) \tag{13}$$

subject to the sequence of constraints

$$\begin{aligned} Y_t(1-g_t) &= C_t + K_t - (1-\delta_t)K_{t-1} \\ Y_t &= \zeta_t K_{t-1}^{\alpha} N_t^{1-\alpha} \end{aligned}$$

where Y_t is output, C_t consumption, K_t the stock of capital N_t is hours worked and $g_t = \frac{G_t}{Y_t}$ is the share of government expenditure in output. The system is perturbed by two exogenous structural disturbances: one to the technology $Z_{t,}$ and one to the government spending share, g_t, both assumed to follow time invariant AR(1) processes

$$\begin{aligned} \ln \zeta_t &= (1-\rho_\zeta)\ln\zeta + \rho_\zeta \ln\zeta_{t-1} + e_t^{\zeta} \\ \ln g_t &= (1-\rho_g)\ln g + \rho_g \ln g_{t-1} + e_t^{g} \end{aligned} \tag{14}$$

where variables without time subscript denote steady state quantities. There are 12 parameters in the model: 6 structural ones (α is the capital share, η the risk aversion coefficient, γ the inverse of the Frisch elasticity of labor supply, A the constant in front of labor in utility, β_t the time discount factor and δ_t the depreciation rate), and 6 auxiliary ones (the steady state values of the government expenditure share and of TFP, (ζ, g), their autoregressive parameters, (ρ_ζ, ρ_g), and their standard deviations (σ_ζ, σ_g)). We assume that all parameters but β_t and δ_t are time invariant. Dueker et al. (2007), Liu et al (2011) and Meier and Sprenger (forthcoming) provide evidence that these two parameters are indeed evolving over time. Since Canova and Sala (2009) have shown that they are only weakly identified in this model, the setup can also be used to verify some of the claims of Magnusson and Mavroedis (2014) and Huang (2014) in a likelihood context. The first order approximation to the law of motion of (β_t ,δ_t) is described below.

The optimality conditions of the problem are:

$$\begin{aligned}
AC_t^{\eta}N_t^{\gamma} &= (1-\alpha)(1-g_t)Y_t/N_t && (15)\\
\beta_t C_t^{-\eta} &= E_t\left(\beta_{t+1}C_{t+1}^{-\eta}(\frac{\alpha(1-g_{t+1})Y_{t+1}}{K_{t+1}}+1-\delta_{t+1}\right) \\
&+ E_t\left(\frac{\partial\beta_{t+1}}{\partial K_t}u(C_{t+1},N_{t+1})-\frac{\partial\delta_{t+1}}{\partial K_t}K_t)\right) && (16)\\
(1-g_t)Y_t &= C_t+K_t-(1-\delta_t)K_{t-1} && (17)\\
Y_t &= \zeta_t K_{t-1}^{\alpha}N_t^{1-\alpha} && (18)
\end{aligned}$$

Time variations in β_t and δ_t affect optimal choices in two ways: there is a direct effect in the Euler equation and in the resource constraint when β_t and δ_t are time varying; and if agents take into account that their decisions may affect parameter variations, there will be a second (endogenous) effect due variations in the derivatives of β_{t+1} and δ_{t+1} with respect to the endogenous states - see equation (16).

Note that varying parameters can not be considered wedges in the sense of Chari et al. (2007). The reason is that, in general, there are cross equation restrictions that need to be satisfied. Furthermore, while the rank of the covariance matrix of the wedges is full, this is not necessarily the case in our setup.

We specialize this setup to consider various possibilities.

2.4.1 Model A: Constant coefficients.

As a benchmark, we let $\beta_t = \beta^t$ and $\delta_t = \delta$. The optimality conditions are

$$E_t\left[f(X_{t+1},X_t,X_{t-1},Z_{t+1},Z_t,\Theta)\right] =$$

$$E_t\begin{pmatrix} AC_t^{\eta}N_t^{\gamma+1}-(1-\alpha)(1-g_t)Y_t \\ C_t^{-\eta}-E_t\beta C_{t+1}^{-\eta}\left(\alpha(1-g_{t+1})Y_{t+1}/K_t+1-\delta\right) \\ (1-g_t)Y_t-C_t+K_t-(1-\delta)K_{t-1} \\ Y_t-\zeta_t K_{t-1}^{\alpha}N_t^{1-\alpha}\end{pmatrix}=0 \qquad (19)$$

$X_t = (K_t, Y_t, C_t, N_t)'$, $Z_t = (\zeta_t, g_t)'$.In the steady state, we have:

$$\frac{K}{Y} = \frac{\alpha(1-g)}{\delta - 1 + 1/\beta}; \ \frac{C}{Y} = 1 - \delta\frac{K}{Y} - \frac{g}{Y}; \ \frac{N}{Y} = \zeta^{\frac{1}{1-\alpha}}\left(\frac{K}{Y}\right)^{\frac{\alpha}{\alpha-1}}; \ Y = \left[\frac{A}{(1-\alpha)(1-g)}\left(\frac{C}{Y}\right)^{\eta}\left(\frac{N}{Y}\right)^{1+\gamma}\right]^{-\frac{1}{\eta+\gamma}}. \tag{20}$$

2.4.2 Model B: Exogenous parameter variations

Set $d_t = \beta_{t+1}/\beta_t$. We let $\Theta_{t+1} - \Theta \equiv (d_{t+1} - (1-\rho_\beta)\beta, \delta_{t+1} - (1-\rho_\delta)\delta)' = U_{t+1}$ and postulate

$$\begin{aligned} u_{d,t+1} &= \rho_d u_{d,t} + e_{d,t+1} && (21) \\ u_{\delta,t+1} &= \rho_\delta u_{\delta,t} + e_{\delta,t+1} && (22) \end{aligned}$$

Since Θ_{t+1} is exogenous, $\partial\beta_{t+1}/\partial K_t = \partial\delta_{t+1}/\partial K_t = 0$ and the f function becomes

$$E_t\left[f(X_{t+1}, X_t, X_{t-1}, Z_{t+1}, Z_t, \Theta_{t+1}, \Theta_t)\right] =$$

$$E_t\begin{pmatrix} AC_t^{\eta}N_t^{\gamma+1} - (1-\alpha)(1-g_t)Y_t \\ 1 - d_t C_{t+1}^{-\eta}/C_t^{-\eta}\left(\alpha(1-g_{t+1})Y_{t+1}/K_t + 1 - \delta_{t+1}\right) \\ (1-g_t)Y_t - C_t - K_t + (1-\delta_t)K_{t-1} \\ Y_t - \zeta_t K_{t-1}^{\alpha}N_t^{1-\alpha} \end{pmatrix} = 0 \tag{23}$$

where $X_t = (K_t, Y_t, C_t, N_t)'$, $Z_t = (\zeta_t, g_t)'$ and $\Theta_t = \Theta_{1t}$.

It is easy to verify that the steady state values of $(\frac{K}{Y}, \frac{C}{Y}, \frac{N}{Y}, Y)$ coincide with those of the constant coefficient model. In addition, since $\phi_x = 0$, variations in (d_{t+1}, δ_{t+1}) leave the decision rule matrices P and Q as in model A. Thus, as far as structural dynamics are concerned, models A and B are observationally equivalent.

To examine whether variations in Θ_t have an instantaneous impact on X_t, we need to check the columns of $N\phi_u\omega_u + O\phi_u$.

$$N\phi_u\rho_u + O\phi_u = \begin{pmatrix} 0 & 0 \\ -1/\beta & -\rho_\delta/\beta \\ 0 & -K \\ 0 & 0 \end{pmatrix} \neq 0 \tag{24}$$

Note that if d_t were a fast moving variable, the impact effect on X_t depends on the persistence of shocks to the growth rate of the discount factor. For example, if $\rho_d = 0$, shocks to the growth rate of the time discount factor have no effects on X_t. Thus, if only the discount factor is time varying and variations in its growth rate are i.i.d., the decision rules of the models A and B will be identical.

2.4.3 Model C: State dependent parameter variations, no internalization

Assume that the time variations in the growth rate of the discount factor and in the depreciation rate are driven by the aggregate capital stock. We specify

$$\Theta_{t+1} = [\Theta_u - (\Theta_u - \Theta_l)e^{-\phi_a(K_t - K)}] + [\Theta_u - (\Theta_u - \Theta_l)e^{\phi_b(K_t - K)}] + U_{\theta,t+1} \tag{25}$$

where $\phi_a, \phi_b, \Theta_u, \Theta_l$ are vectors of parameters and $\mathrm{U}_{\theta,t+1}$ is a zero mean, i.i.d. vector of shocks. This specification is flexible: depending on the choice of $\phi's$,we can accommodate linear or quadratic relationships, which are symmetric or asymmetric. To insure that models C and A have the same steady states, we set $\Theta_l = (\beta/2,\ \delta/2)$.

We assume that agents treat the capital stock appearing in (25) as an aggregate variable. This assumption is similar to the 'small k -big k' situation encountered in standard rational expectations models or to the distinction between internal and external habit formation. Thus, agents' first order conditions do not take into account the fact that their optimal capital choice changes d_t and δ_t and $\partial\beta_{t+1}/\partial K_t = \partial\delta_{t+1}/\partial K_t = 0$. Hence, the equilibrium conditions are then as in (23).Since the f function is the same as in model B, the matrices N and O are unchanged.

First, examine whether parameter variations affect the matrices regulating structural dynamics. We have,

$$N\phi_x = \begin{pmatrix} 0 & 0 \\ 0 & 1/\beta \\ 0 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} (d_u - \beta/2)(\phi_1 - \phi_2) & 0 & 0 & 0 \\ (\delta_u - \delta/2)(\phi_3 - \phi_4) & 0 & 0 & 0 \end{pmatrix} \tag{26}$$

$$O\phi_x = \begin{pmatrix} 0 & 0 \\ -1/\beta & 0 \\ 0 & -K \\ 0 & 0 \end{pmatrix} \begin{pmatrix} (d_u - \beta/2)(\phi_1 - \phi_2) & 0 & 0 & 0 \\ (\delta_u - \delta/2)(\phi_3 - \phi_4) & 0 & 0 & 0 \end{pmatrix} \tag{27}$$

Thus, unless $\phi_1 \neq \phi_2$ and/or $\phi_3 \neq \phi_4$, endogenous variations in d_t, δ_t leave P and Q unaffected, i.e. asymmetries in the law of motion of d_t, δ_t are needed to produce structural dynamics which are different from those of a constant coefficient model. To verify whether parameter variations have an impact effect on X_t, we need to check the columns of $N\phi_u\omega_u + O\phi_u$. We have:

$$N\phi_u\omega_u + O\phi_u = \begin{pmatrix} 0 & 0 \\ 1/\beta(d_u - \beta/2)(-\phi_1 + \phi_2) & 0 \\ 0 & K(\delta_u - \delta/2)(-\phi_3 + \phi_4) \\ 0 & 0 \end{pmatrix} \neq 0 \tag{28}$$

as long as $\phi_1 \neq \phi_2$or $\phi_3 \neq \phi_4$ and regardless of whether shocks to the parameters are i.i.d. or persistent.

2.4.4 Model D: State dependent parameter variations, internalization.

We still assume that time variations in the discount factor and in the depreciation rate are driven by the aggregate capital stock and by an exogenous shock, as in equation (25). Contrary to case C, we assume that agent internalize the effects that their capital decisions have on parameter variations. The relevant derivatives are

$$d'_{t+1} \equiv \partial d_{t+1}/\partial K_t = -(\beta_u - \beta/2)[-\phi_1 e^{-\phi_1(K_t-K)} + \phi_2 e^{\phi_2(K_t-K)}] \quad (29)$$

$$\delta'_{t+1} \equiv \partial \delta_{t+1}/\partial K_t = -(\delta_u - \delta/2)[-\phi_3 e^{-\phi_3(K_t-K)} + \phi_4 e^{\phi_4(K_t-K)}] \quad (30)$$

In order for the steady states of model D to equal to those of model A, we restrict $\phi_1 = \phi_2 = \phi_1$, $\phi_3 = \phi_4 = \phi_3$. The optimality conditions now are:

$$0 = E_t\left[f(X_{t+1}, X_t, X_{t-1}, Z_{t+1}, Z_t, \Theta_{t+1}, \Theta_t)\right] =$$

$$E_t \begin{pmatrix} AC_t^{\eta} N_t^{\gamma+1} - (1-\alpha)(1-g_t)Y_t \\ 1 - d'_t\ u(C_{t+1}, N_{t+1})/C_t^{-\eta} - d_t\ C_{t+1}^{-\eta}/C_t^{-\eta}(\alpha(1-g_{t+1})Y_{t+1}/K_{t+1} + 1 - \delta_{t+1} + \delta'_{t+1}K_t) \\ (1-g_t)Y_t - C_t - K_t + (1-\delta_t)K_{t-1} \\ Y_t - \zeta_t K_{t-1}^{\alpha} N_t^{1-\alpha} \end{pmatrix} \quad (31)$$

where as before $X_t = (K_t, Y_t, C_t, N_t)'$, $Z_t = (\zeta_t, g_t)'$ but now $\Theta_t = (d_t, \delta_t, d'_t, \delta'_t)'$ and its law of motion is

$$\begin{pmatrix} d_{t+1} \\ \delta_{t+1} \\ d'_{t+1} \\ \delta'_{t+1} \end{pmatrix} = \Phi(\Theta, K_t, U_{t+1}) = \begin{pmatrix} 2d_u - (d_u - \beta/2)[e^{-\phi_1(K_t-K)} + e^{\phi_1(K_t-K)}] + U_{\beta,t+1} \\ 2\delta_u - (\delta_u - \delta/2)[e^{-\phi_3(K_t-K)} + e^{\phi_3(K_t-K)}] + U_{\delta,t+1} \\ -(d_u - \beta/2)\phi[-e^{-\phi_1(K_t-K)} + e^{\phi_1(K_t-K)}] \\ -(\delta_u - \delta/2)\phi[-e^{-\phi_3(K_t-K)} + e^{\phi_3(K_t-K)}] \end{pmatrix} \quad (32)$$

The relevant matrices of derivatives evaluated at the steady states are $\omega_u = 0_{2\times 2}$

$$N = \frac{\partial f}{\partial \Theta_{t+1}} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1/\beta & -u(C,N)/C^{-\eta} & -\beta K \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad O = \frac{\partial f}{\partial \Theta_t} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1/\beta & 0 & 0 & 0 \\ 0 & -K & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\phi_x = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -2(\beta_u - \beta/2)\phi_1^2 & 0 & 0 & 0 \\ -2(\delta_u - \delta/2)\phi_3^2 & 0 & 0 & 0 \end{pmatrix}, \quad \phi_u = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ -2(\beta_u - \beta/2)\phi_1^2 & 0 \\ 0 & -2(\delta_u - \delta/2)\phi_3^2 \end{pmatrix}$$

Clearly, $N\phi_x \neq 0$, $O\phi_x = 0$ and $N\phi_u\omega_u + O\phi_u = 0$. Thus, a shock to the law of motion of the parameters alters the dynamics produced by structural shocks, even when the relationship between parameters and states is symmetric.

In sum, parameter variations matter for the structural dynamics either if the relationship between parameters and the states is asymmetric; or if agents internalize the consequences their decisions have on parameter variations, or both.

2.4.5 Impulse responses

Why are structural dynamics in models C and D different from those in model A? To understand what drive the economic differences, we compute impulse responses. For the parameters common to all models we choose $\alpha = 0.30$, $\beta = 0.99$, $\delta = 0.025$,

$\gamma = 2$, $\eta = 2$, $A = 4.50$, ζ=1;$\rho_\zeta = 0.90$, $\sigma_\zeta = 0.00712$, $g = 0.18$, $\rho_g = 0.50$ and $\sigma_g = 0.01$. For the parameter specific to the time varying parameters models we choose:

- Model B: $\rho_\beta = 0.985, \rho_\delta = 0.95$ and $\sigma_\beta = 0.002$ $\sigma_\delta = 0.07$.
- Model C : $\phi_{1\beta} = 0.01, \phi_{2\beta} = 0.03, \phi_{1\delta} = 0.2, \phi_{2\delta} = 0.1$, $\sigma_d = \sigma_\delta = 0.5$, $\beta_u = 0.999, \delta_u = 0.025$.
- Model D : $\phi_{1\beta} = 0.0001, \phi_{2\beta} = 0.016, \phi_{1\delta} = 0.2, \phi_{2\delta} = 0.1$, $\sigma_d = 0.0001$;$\sigma_\delta = 0.1$, $\beta_u = 0.999, \delta_u = 0.025$.

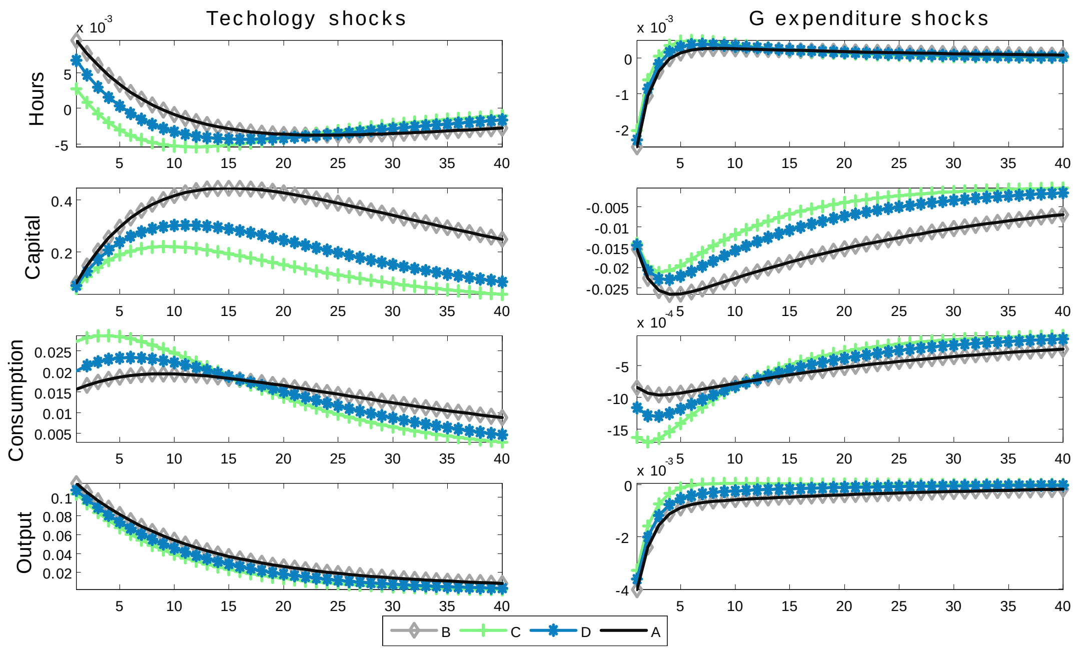

Figure 1: Impulse responses, first order approximation

Figure 1 reports the responses of hours, capital, consumption and output to the two structural shocks in the four models: the first column has the responses to technology shocks; the second the responses to government expenditure shocks [1].

The responses of models C and D differ from those of models A and B primarily in the shape and the persistence of consumption and capital responses. Differences occur because income and substitution effects are different from those in model A. For

[1]Since the responses of hours and output to government expenditure shocks are different from what the conventional wisdom indicates, a few words of explanations are needed. In a standard RBC in response to government expenditure shocks hours and output typically increase because of a wealth effect. However, here the shock affects the share of government expenditure in GDP. Thus, the positive wealth effect on labor supply is absent because government expenditure increase in exactly the same proportion as output, thus disincentivizing agents to try to increase private output.

instance, in responses to technology shocks, agents work less, save less, and consume more in models C and D than in the constant coefficients model, while in response to government expenditure shocks consumption falls more and capital falls less relative to the constant coefficients case. Thus, the uncertainty present in the parameters play a role and influence the way agents react to transitory structural shocks.

3 Characterizing time varying misspecification

Because the decision rules of constant coefficient models are generally misspecified when the data generating process (DGP) features parameter variations, it is important to have diagnostics which may detect potential problems. This section describes two diagnostics, one based on "wedges" (see Chari et al., 2008) and one based on forecast errors, which can be useful for the purpose.

Consider the optimality conditions of a constant coefficient model

$$E_t\left[F(X_{t-1}, W_t, \sigma\Sigma\epsilon^z_{t+1}, \Theta)\right] = 0 \tag{33}$$

obtained using the resulting decision rule:

$$X_t = h(X_{t-1}, W_t, \sigma\Sigma\epsilon^z_{t+1}, \Theta) \tag{34}$$

When the data X_{t-1} for has been generated by the constant coefficient model, F is a martingale difference and there is no wedge. When instead X_{t-1} has been generated by a time varying coefficient model with

$$X^*_t = h^*(X^*_{t-1}, W_t, \sigma\Sigma\epsilon_{t+1}, \Theta) \tag{35}$$

E$[F(X^*_{t-1}, W_t, \sigma\Sigma\epsilon^z_{t+1}, \Theta)] \neq 0$, since $\sigma\Sigma\epsilon^z_{t+1} \neq \sigma\Sigma\epsilon_{t+1}$ and h$\neq h^*$. Furthermore, $F(X^*_{t-1}, W_t, \sigma\Sigma\epsilon^z_{t+1}, \Theta)$ will be predictable using past values X^*_{t-1}. To understand why this is the case, consider the first order approximate optimality conditions. In this system of equations, the wedge is

$$\begin{aligned} (F(P^*-P)^2 + G(P^*-P))x_{t-1} \quad &+ \\ (F(Q^*-Q)\psi_z + G(Q^*-Q) + F(P^*-P)(G^*-G))z_t \quad &+ \\ (F(P^*-P)R^* + GR^* + FR^*\omega_u)u_t \quad & \end{aligned} \tag{36}$$

When $P^* = P, Q^* = Q$, as in the exogenously varying model, the wedge reduces to

$$(GR^* + FR^*\omega_u)u_t \tag{37}$$

which is different from zero if $R^* \neq 0$ and predictable if $\omega_u \neq 0$. When, as in the endogenously varying model, $P^* \neq P, Q^* \neq Q$, the wedge will be different from zero, even when $R = 0$, and predictable using past x_{t-1}, even when $\omega_u = 0$.

Hence, to detect time varying misspecification one can compute wedges and regress them on the lags of the observables. If they are significant, the martingale difference condition is violated, and there is evidence of time varying parameters. Note that the diagnostic uses the assumption that the model is correctly specified up to parameter variations. If it is incorrect, lags of the observables may be significant, even without time varying coefficients.

The logic of the forecast error diagnostic is similar. The linearized decision rule in a constant coefficients model is $x_t = Px_{t-1}+Qz_t$, while in a time varying coefficients model is $x^*_t = P^*x^*_{t-1}+Q^*z_t+R^*u_t$. Let v^*_t be the forecast error in predicting x^*_t using the decision rules of the constant coefficient model and the data generated from the time varying coefficient model. The forecast error can be decomposed as

$$v^*_t = x^{*j}_t - Px^*_{t-1} = Q^*z_t + R^*u_t + (P^* - P)x^*_{t-1} \tag{38}$$

Thus, forecast errors are functions of the lags of the observables x^*_{t-1} when $P^* \neq P$. However, even if $P^* = P$, forecasts error linearly depend on the lags of the observables if u_t is serially correlated. Hence, an alternative way to check for model misspecification involves regressing the forecast errors v^*_t on lagged values of the observables and checking the significance of the regression coefficients.

We apply the two diagnostics to data produced by the RBC example previously considered. Using a long realization (T=1000) of (Z_t, Θ_t), we find that the Euler wedge is equal to 0.012 (DGP model B), to 0.008 (DGP model C) and to 0.011 (DGP model D) and that the consumption growth component of the wedge $[(\frac{c^*_{t+1}}{c^*_t})(\frac{c_t}{c_{t+1}})]^{-\eta}$ largely account for the deviations of the wedge from zero. Regressing the Euler wedge on the observables, we find that the coefficients on the first lag of the real rate are 1.06 (DGP model B) 1.81 (DGP model C) and -0.59 (DGP model D) all significantly different from zero, while the coefficients on the first lag of consumption growth are -0.0008 (DGP model B), 0.06 (DGP model C) and -0.05 (DGP model D) and the second is significantly different from zero [2].

When we consider the forecast errors in the hour equation, we find that lagged values of the endogenous variables are significant regardless of the process generating time variations (see table 1) and that an F-test strongly rejects the null hypothesis that all the coefficient are jointly zero.

3.1 Exogenous vs. endogenous parameter variations

If the wedge and the forecast error diagnostics indicate the presence of parameter variations, one may interested in knowing whether they are of exogenous or of endogenous type. One way to distinguish the two options is to use the DGSE-VAR methodology of Del Negro and Schorfheide (2004). In a DSGE-VARs one uses the DSGE model as a prior for the VAR of the observable data and employs

[2] To check what would happen to our diagnostic when the model is incorrectly specified, we simulate data from an RBC model with constant coefficients and one period time to build and consider our baseline model with no time to build and constant coefficients. We find that the coefficients on the first lag of the real rate and of consumption growth are -0.03 and 0.06, both insignificantly different from zero.

DGP	n_{t-1}	k_{t-1}	y_{t-1}	c_{t-1}	Ftest, P-value
B	0.08	-0.40	0.05	0.51	0.00
	(0.004)	(0.006)	(0.007)	(0.002)	
C	0.08	-0.28	-0.15	0.43	0.00
	(0.002)	(0.007)	(0.003)	(0.29)	
D	0.27	0.09	0.33	-1.93	0.00
	(0.06)	(0.01)	(0.02)	(0.21)	

Table 1: Regression coefficients, hours equation; the dependent variable is the forecast error obtained using the decisions rules of model A and data from the model in the first column; the indipendent variables are in the first row. In parenthesis standard errors.

the marginal likelihood to measure the value of the additional information the DSGE provides. Intuitively, a DSGE prior can be thought as a set additional observations added to the actual data. If the additional observations come from the DGP, the quality of the estimates improves (standard errors will be reduced), and the marginal likelihood, which measures the fit of the specification, increases. On the other hand, if the additional observations come from a DGP different from the one generating the data, biases may be introduced, noise added, and the precision of the estimates and the fit of the model reduced. Thus, for a given data set, a researcher comparing the marginal likelihood produced by adding data from the exogenous and the endogenous specifications, should detect whether the observable sample is more likely to be generated by one of the two models.

Let $L(\alpha|y)$ be the likelihood of the VAR model for data y and let $g_j(\alpha|\gamma_j, M_j)$ be the prior induced by the DSGE model M_j using parameters γ_j on the VAR parameters α .The marginal likelihood is $h_j(y|\gamma_j, M_j) = \int L(\alpha|y)g_j(\alpha|\gamma_j, M_j)d\alpha$ which, for given y,is a function of M_j. Since $L(\alpha|y)$ is fixed, $h_j(y|\gamma_j, M_j)$ reflects the plausibility of $g_j(\alpha|\gamma_j, M_j)$ in the data. Thus, if g_1 and g_2 are two DSGE-based priors and $h_1(y|\gamma_1, M_1) > h_2(y|\gamma_2, M_2)$, there is better support for in the data for g_1.

	T_1=150			T_1=750		
DGP	Model B	Model C	Model D	Model B	Model C	Model D
Simulated from B	1586	-6709	-5108	9714	-3478	-12597
Simulated from C	1421	2005	-855	7480	4828	-409
Simulated from D	697	-2649	1864	6083	622	11397

Table 2: Log marginal likelihood obtained using T data points produced by the models listed in the first row and T_1 simulated data from the model listed in the first column.

Table 2 reports results using this technology in the RBC example. The sample size of the data is $T = 150$, and the log marginal likelihood is computed when

$T_1 = 150, 750$ simulated data from the DSGE listed in the first row are added to the actual data. The marginal likelihood picks the correct DGP in all experiments and the differences within columns are quite large, even when $T_1 = 150$.

4 Parameter identification

Since forecast errors are typically used to construct the likelihood function via the Kalman filter, one should expect the misspecification present in the forecast errors to spread to the likelihood function. In this section we examine whether time invariant parameters can be identified from a potentially misspecified likelihood function. Canova and Sala (2009) have shown that standard DSGE models feature several population identification problems, intrinsic to the models and to the solution method employed. In particular, in many models several structural parameters are weakly or partially identified and others are close to be underidentified. The issue we are concerned with here is whether parameters which could be identified if the correct likelihood is employed became weakly or partially identified when the wrong likelihood is used. In other words, we ask whether identification problems in time invariant parameters may emerge as a by-product of neglecting variations in other parameters. Magnusson and Mavroedis (2014) have shown that when GMM is used, time variations in certain parameters help the identification of time invariant parameters. Huang (2014) qualifies the result by showing that time variations in weakly identified parameters have no effect on the asymptotic distribution of strongly identified parameters.

Figures 2 and 3 plot the likelihood function of the RBC model in the risk aversion coefficient γ and the share parameter η; and in the labor share α and the autoregressive parameter of the technology ρ_ζ, when the forecast errors of the correct model (top row) and of the constant coefficient model (bottom row) are used to construct the likelihood function. The first column considers data generated by the model B, the second and the third data generated by models C and D.

While the likelihood curvature in the correct model is not large, it is easy to verify that the maximum occurs at $\gamma = 2$, $\eta = 2, \alpha = 0.30$, $\rho_\zeta = 0.9$ for all three specifications. When the decision rules of the constant coefficients model are used to construct the likelihood function and the true DGP is model B, the likelihood is flattened and the risk aversion coefficient γ become very weakly identified. When the true model features endogenous time variations, distortions are larger: likelihood function become locally convex in ρ_ζ; γ and α become very weakly identified, and the maximum in the ρ_ζ is shifted away from the true value.

These observations are confirmed by the Koop et al. (2013) statistic, see table 4. Koop et al. show that asymptotically the precision matrix grows at the rate T for identified parameters and at rate less than T for underidentified parameters.

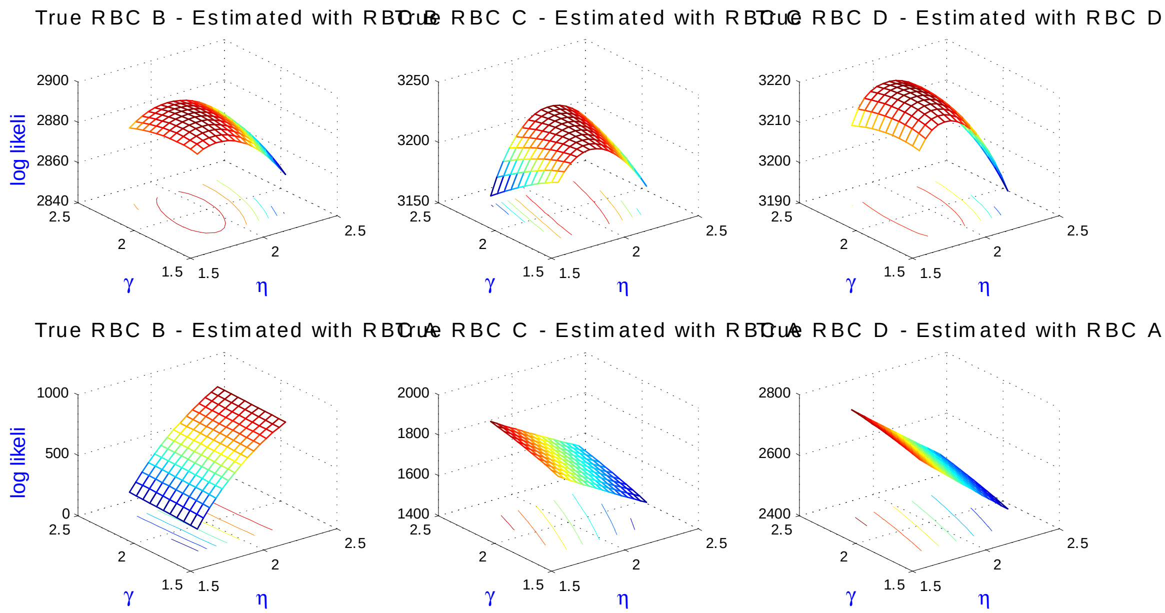

Figure 2: Likelihood surfaces

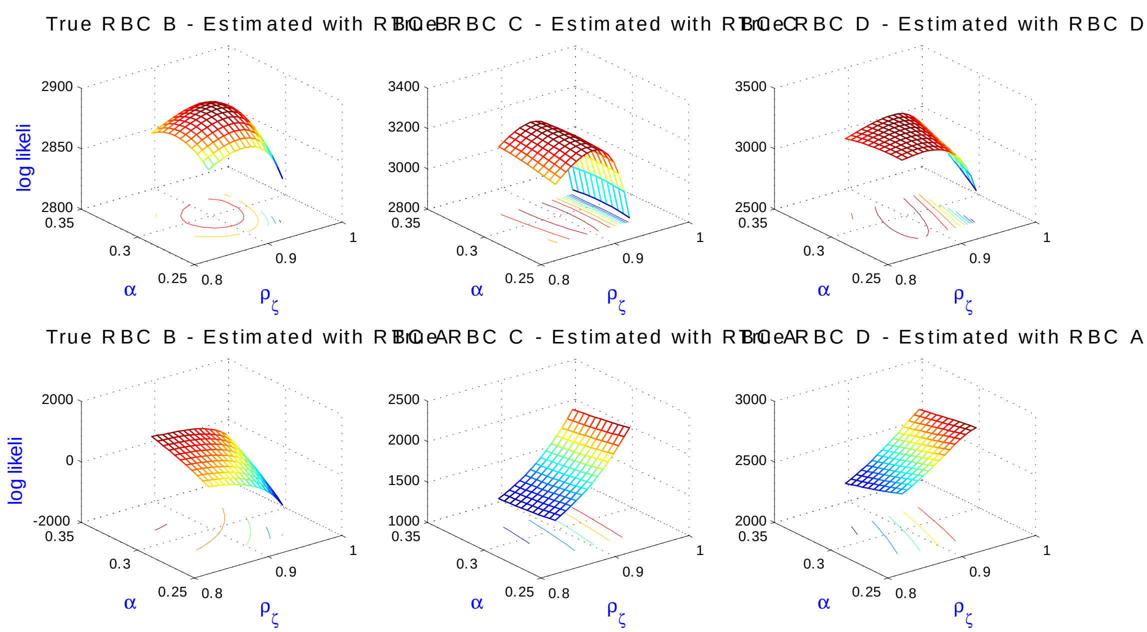

Figure 3: Likelihood surfaces

Thus, the precision of the estimates, scaled by the sample size, converges to a constant for identified parameters and to zero for underidentified parameters. Furthermore, the magnitude of the constant measures identification strength: a large value indicates a strongly identified parameter; a small value a weakly identified one.

Koop, Pesaran, Smith diagnostic							
Parameter	T=150	T=300	T=500	T=750	T=1000	T=1500	T=2500
	DGP Model B, Estimated model A						
η	15.9	17.8	17.2	18.8	18.4	19.3	17.9
γ	28.5	45.7	108.4	81.4	93.6	104.2	90.17
ρ_z	1.8e+4	2.6e+4	4.2e+4	4.2e+4	4.5e+4	4.9e+4	4.37e+4
ρ_g	209.2	655.5	2741	2190	2860	3417	2802
δ	927.3	973.8	1.7e+4	1.7e+4	2.4e+4	2.3e+4	2.5e+4
α	140.2	156.2	264.2	215.5	239.1	252.1	229.3
A	28.42	30.67	7.99	10.99	9.15	7.83	9.83
	DGP Model C, Estimated model A						
η	822	1033	743	785	759	746	752
γ	2261	3147	2682	2809	2720	2579	2566
ρ_z	3073	2673	2952	2909	2799	2806	2877
ρ_g	1.74	2.23	2.44	2.96	3.17	2.82	2.90
δ	4.6e+5	4.4e+5	4.3e+5	4.0e+5	3.8e+5	4.4e+5	4.3e+5
α	1.8e+4	1.1e+4	1.4e+4	1.2e+4	1.1e+4	1.6e+4	1.5e+4
A	351	493	441	505	500	449	444
	DGP Model D, Estimated model A						
η	550	575	592	610	545	542	494
γ	3577	2442	2660	2870	2564	2711	2430
ρ_z	1613	1243	1120	1162	1068	1189	1074
ρ_g	1.22	1.28	1.44	1.53	1.60	1.62	1.67
δ	5.2e+5	6.7e+5	6.5e+5	6.0e+5	5.7e+5	5.8e+5	5.7e+5
α	1.1e+4	2.5e+4	2.4e+4	1.9e+4	2.1e+4	2.0e+4	2.1e+4
A	488	276	340	382	349	395	334

Table 3: Koop et al diagnostic. Different sample sizes

When the DGP is model B and a fixed coefficients model is considered, all parameters are identified, even though some weak identification issues for A and η exist. When the DGP are models C and D, all parameters but ρ_g seem identifiable. Interestingly, in models C and D, ρ_g is weakly identified, even when the correct likelihood is used. Thus, time variations in the weakly identified β_t and δ_t do not help in the identification of ρ_g, in line with Huang (2014).

5 Likelihood estimation with a misspecified model

To study the properties of likelihood based estimates of a misspecified constant coefficients model, we conduct a Monte Carlo exercise. We generate either 150 or 1000 data points from versions B, C, D of the RBC model previously considered, estimate the structural parameters using the likelihood function constructed with the decision rules of the time invariant model A, and repeat the exercise 150 times using different realizations of the shocks. We also estimate the structural parameters using the likelihood constructed with the decision rules of the correct model (i.e. model B rules if the data has been generated by model B, and so on), for benchmarking the estimation distortions.

We consider two setups: one where parameter variations are small (2-5 percent of the variance of output is explained by shocks to the parameters, henceforth DGP1) and one where parameter variations are substantial (around 20 percent of the variance of output is explained by shocks to the parameters, henceforth DGP2). Table 4 has the results for DGP1: it reports the fixed parameters used to generate the data (column 1), the mean posterior estimate (across replications) obtained when the likelihood is constructed with the correct decision rules (column 2), and the mean posterior estimate, the 5th and the 95 percentile of the distribution of estimates obtained when the likelihood function is constructed with the decision rules of the time invariant model, when T=150 (columns 3-5) and when T=1000 (columns 6-8). Table A1 in the appendix has the results for DGP2. When the model is correctly specified, the distribution of estimates should collapse around the true value. Thus, if the mean is away from the true parameter value and/or the spread of the distribution is large, likelihood based methods have difficulties in recovering the constant parameters of the data generating process. Figures A1 and A2 in the appendix plot the distributions of estimates for the two DGPs: the vertical line represents the true parameter value; in solid black lines are distributions obtained with the correct model; in solid blue (red) lines are the distributions obtained with the incorrect constant coefficient model when T=150 (T=1000). . Figure 4 presents the impulse responses for DGP1: in the first two columns are the responses to technology shocks and government expenditure shocks in model B, in the next two the responses in model C, and in the last two the responses in model D. In each box we report the response obtained using mean value of the correct distribution of estimates, and the 16th and 84th percentiles of the distribution of responses obtained using the estimated distribution of parameters produced by the time invariant model. Figure A3 in the appendix has the same information for DGP2. Finally, table 5 presents the long run variance decomposition for DGP1 (table A2 has the information for DGP2) when T=150 and the mean posterior estimate is used in the computations: in the first two columns we have the contribution of technology and government spending shocks in the correct model; the last two columns the contribution when the incorrect constant coefficient model is used.

For the two time varying parameters, we set $d_t = \beta_{t+1}/\beta_t$, and assume that in model B, $\Theta_{t+1} - \Theta \equiv (d_{t+1}(1-\rho_\beta)\beta, \delta_{t+1} - (1-\rho_\delta)\delta)' = U_{t+1}$, where $\beta = 0.99$,the components of $U_{t+1} = (u_{d,t+1}, u_{\delta,t+1})'$ are independent AR(1) process with persistence $\rho_d = 0.9, \rho_\delta = 0.8$, and standard deviations $\sigma_d = 0.002, \sigma_\delta = 0.07$. For models C and D, the law of motion of the time varying parameters is $\Theta_{t+1} = [\Theta_u - (\Theta_u - \Theta_l)e^{-\phi_a(K_t - K)}] + [\Theta_u - (\Theta_u - \Theta_l)e^{\phi_b(K_t-K)}] + U_{t+1}$, where $\Theta_u' = (0.9999, 0.03)$, $\phi_a' = (0.03, 0.2)$, $\phi_b' = (0.031, 0.1)$, U_{t+1} is iid with Σ_u diagonal and $\sigma_d = 0.03, \sigma_\delta = 0.008$.

True Parameter	Correct Mean T=150	Time invariant T=150 Mean	5th percentile	95th percentile	Time invariant T=1000 Mean	5th percentile	95th percentile
DGP Model B							
$\eta = 2.0$	2.00	2.03	1.47	2.88	2.32	1.55	3.37
$\gamma = 2.0$	2.02	1.23	-0.14	2.07	0.96	-0.38	2.04
$\rho_z = 0.98$	0.97	0.99	0.97	1.00	0.99	0.96	1.00
$\rho_g = 0.5$	0.47	0.74	0.60	0.96	0.87	0.77	0.98
$\delta = 0.025$	0.03	0.01	0.01	0.02	0.01	0.01	0.05
$\alpha = 0.3$	0.30	0.19	0.11	0.28	0.23	0.15	0.40
$A = 4.5$	4.55	2.79	1.33	4.12	2.68	1.23	4.06
DGP Model C							
$\eta = 2.0$	2.00	2.42	1.63	3.85	2.85	1.73	6.14
$\gamma = 2.0$	2.00	0.64	-0.26	1.77	0.60	-0.50	1.79
$\rho_z = 0.98$	0.98	0.99	0.97	1.00	0.97	0.85	1.00
$\rho_g = 0.5$	0.48	0.43	-0.10	0.96	0.65	0.27	0.98
$\delta = 0.025$	0.03	0.01	0.01	0.02	0.02	0.01	0.09
$\alpha = 0.3$	0.30	0.22	0.13	0.34	0.29	0.18	0.47
$A = 4.5$	4.49	2.14	1.18	3.47	2.37	1.18	3.66
DGP Model D							
$\eta = 2.0$	2.00	2.58	1.69	3.34	2.40	1.74	3.26
$\gamma = 2.0$	2.01	0.29	-0.28	1.54	1.09	-0.30	1.99
$\rho_z = 0.97$	0.96	0.99	0.94	1.00	0.96	0.91	1.00
$\rho_g = 0.5$	0.48	0.51	-0.26	0.96	0.66	0.39	0.98
$\delta = 0.025$	0.02	0.01	0.01	0.03	0.01	0.01	0.02
$\alpha = 0.3$	0.30	0.22	0.14	0.35	0.22	0.15	0.30
$A = 4.5$	4.52	2.32	1.42	3.68	3.45	1.37	4.51

Table 4: Distributions of posterior estimates, DGP1.

A few features of the results are worth discussing. First, when the correct model is employed, estimation is successful even when T=150, regardless of the DGP and of whether time variations are exogenous or endogenous. Thus, numerical distortions seem not to be present. Second, with DGP1, a number of distortions occur when a time invariant model is used in estimation. For example, when exogenous variations are present, the persistence of government spending shock is poorly estimated (mean persistence is about 50 percent larger than the true one), while estimates of δ, α and A are severely biased downward. The distortions are smaller when the time variations are endogenous (models C and D): nevertheless significant downward bias exists in the inverse of the Frisch elasticity γ, in δ and α. Notice that the performance of the time invariant model is roughly independent of whether the data features external or internal endogenous time variations. Third, the performance of the time invariant model does not improve when T=1000 for all three specifications: as sample size increases, failure to converge to the true DGP becomes more obvious.

When parameter variations explain a significant portion of output variability, all features become more striking. For example, when parameter variations are exogenous, estimating a time invariant model leads to an overestimation of the persistence of the structural shocks. In fact, the only way a time invariant model can accommodate the additional dynamics and variability present in the endogenous variables is by increasing the persistence of both shocks. In models C and D the distortions become considerably larger and, for example, the mean posterior estimate of inverse of the Frisch elasticity is now negative. In addition, the distribution of estimates is typically skewed and multimodal. Thus, neglected parameter variations are more detrimental when they account for a significant portion of the variability of the endogenous variables.

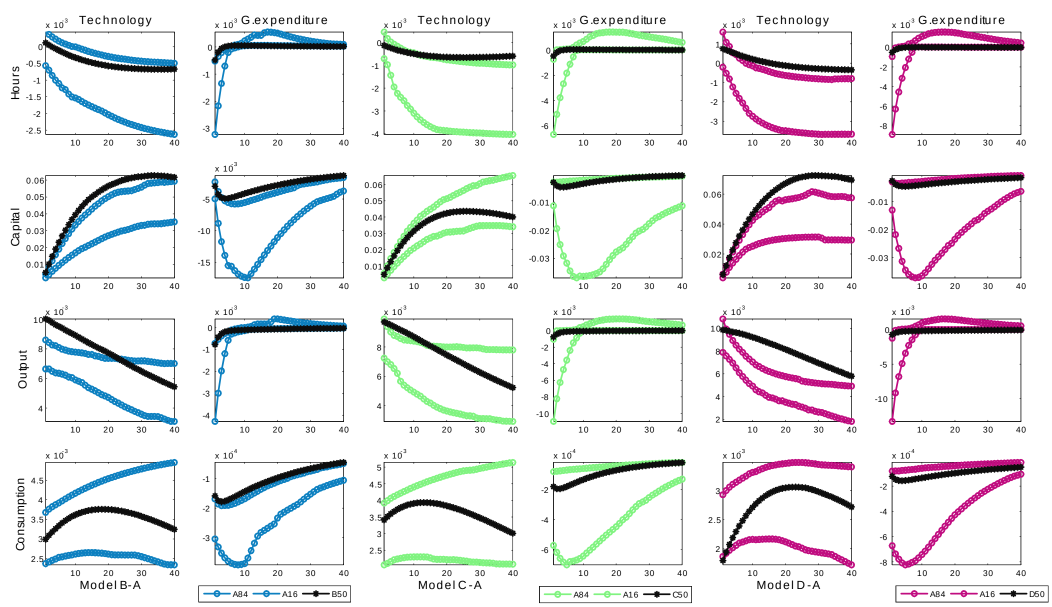

Figure 4:Impulse responses, DGP1

The impulse responses are in line with these conclusions. When parameter variations explain a small fraction of the variability of output, the responses to technology shocks are off in terms of impact magnitude, in particular for output; and the response produced with estimates obtained with the true model tend to be on the upper limit of the estimated 68 percent band produced with estimates of the incorrect model. Interestingly, output responses are those more poorly characterized with the time invariant model and, consistent with previous findings, the misspecification obtained when the true model features exogenous time variations is larger. The responses to government expenditure shocks obtained with a time invariant model are different from those obtained estimating the correct model in terms of magnitude, shape and persistence. Since the signal that government expenditure produces in the model is weak, it is not surprising that it is obscured by the presence of time variations.

The distortions obtained when parameter variations are important for the variance of output are generally larger. Here, for example, the persistence of the responses to technology shocks is poorly estimated: while responses obtained estimating the true model tend to zero, the bands obtained estimating a time invariant model do not include zero even after 10 years.

Variable	Technology	Government	Technology	Government
	Model B		Time invariant	
Y	94.100	0.300	0.997	0.004
C	89.500	0.200	0.999	0.001
N	60.200	0.500	0.986	0.014
K	70.200	0.400	0.995	0.006
	Model C		Time invariant	
Y	97.200	0.300	0.988	0.016
C	88.100	0.300	0.999	0.001
N	44.600	0.600	0.990	0.012
K	84.400	0.200	0.990	0.014
	Model D		Time invariant	
Y	98.000	0.100	0.993	0.015
C	92.200	0.200	0.998	0.003
N	35.900	0.500	0.973	0.034
K	96.600	0.300	0.992	0.012

Table 5: Long run variance decomposition, DGP1

What is the contribution of structural shocks to the variability of the endogenous variables when the forecast errors of the time invariant model are used to construct the likelihood function? One should expect the structural shocks of the time invariant model to be a contaminated version of the structural shocks of the time varying DGP for two reasons. First, the wrong P matrix is used to compute forecast errors. Second, we are aggregating m (structural and parameter) shocks into $n < m$ (structural) shocks, thus generating VARMA decision rules where the n structural shocks are functions of the leads and lags of the original disturbances (see e.g. Canova and Paustian, 2011).

Thus, even if the P matrix were correctly specified, distortions should occur, unless the shocks to the parameters are unimportant and feature low persistence.

When parameter variations do not explain a large portion of the variance of output, technology shocks in the time invariant model absorb the missing variability, regardless of whether parameter variations are exogenous or endogenous and the effect seems particularly strong for hours worked. When parameter variations explain a larger portion of the variance of output, technology shock still absorb a large amount of the missing variability. However, in some cases, both shocks capture the missing variability. For example, while government spending shocks explain only 0.1 percent of the variability of output in the long run in the true model when parameter variations are endogenous, they explain 7-8 percent when a time invariant model is used.

In sum, for the DGP we consider and the parameterization employed, estimating a constant parameter model when the DGP features time varying parameters leads to distortions, regardless of the sample size, of whether variations are exogenous or endogenous, and of whether parameter variations matter for output variability or not. The parameters mostly affected are those regulating the estimated persistence of the shocks and those controlling income and substitution effects. [3]

6 Structural dynamics and SVAR methods

Our results suggest that if the DGP features parameter variations, and these variations are neglected in the likelihood function, parameter estimates are biased and structural responses distorted. Because of these problems, one may wonder whether less structural and computationally less demanding methods can be used if structural dynamics is all that matters to the investigator. Canova and Paustian (2011) have shown that when the model misses features of the DGP, VAR methods which employ robust sign restrictions can be effective in capturing qualitative features of structural dynamics induced. Here we ask if VARs are good also when the model neglects parameter variations.

The exercise we conduct in this section is as follows. Using the illustrative RBC model, we simulate data from the decision rules of models B, C, and D when parameters variations generate small output volatility (DGP1). We then compute VAR residuals using the population P matrix of the correct model and of the constant coefficient model, rotate the resulting residuals using an orthonormal matrix, and keep responses if technology shocks generate a positive response of hours, capital, output and consumption on impact and if government expenditure

[3] We have also performed a Monte Carlo exercise where the labor share is also time varying. Variations in the labor share have been documented in the literature (see e.g. Rios Rull and Santaeularia Llopis, 2010) and there is evidence that variations in this parameter are strongly countercyclical. This is relevant for our exercise because, in this case, all four optimatility conditions are affected by time variations. Thus, the strength of the income and substitution effect distortions are likely to be larger. Indeed, we do find that distortions in this case become quite large and it many cases it becomes difficult to estimate the time invariant model regardless of the DGP (results for this setup are available on request).

shocks generate a negative response of hours, output, consumption and capital - these signs are those present in figure 1 and hold for variations of the (constant) structural parameters within a reasonable range. We repeat the exercise 150 times and collect the distribution of structural responses for the correct and the time invariant specifications. Figure 5 plots the median response obtained with the correct model (red line in each box) and the 16 and 84 percentile of the distribution of responses obtained with the incorrect model.

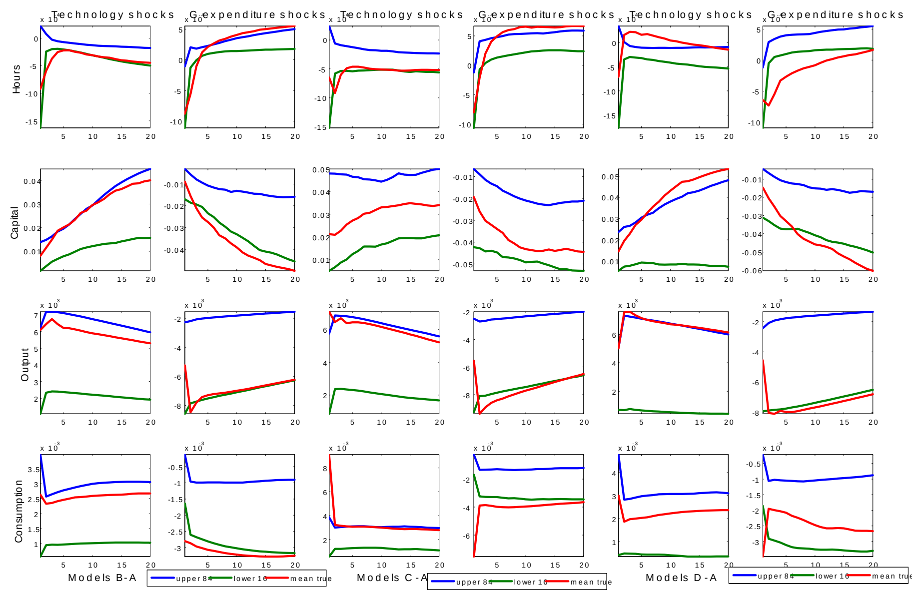

Figure 5: Impulse responses,SVAR models

Overall, SVAR methods are competitive with structural methods when parameter variations are neglected. When the DGP is model B, the sign and the shape of the responses is correctly captured. Although the responses to technology shocks obtained with the true model tend to be on the upper bound of the band obtained with the incorrect model and the responses to government spending shocks obtained with the true model tend to be on the lower bound of the bands obtained with the incorrect model, no major distortions occur.

The performance with the other two DGPs is similar. However, with model C it is the magnitude of the dynamic response of consumption which is misrepresented, while with model D it is primarily the persistence of certain responses that is somewhat underestimated.

Recall that there are two sources of misspecification: the P matrix is incorrect; aggregation problems are present. Our analysis indicates, that with the DGP we use, i) distortion in the P matrix are small; ii) the Q matrix is not very strongly affected by misspecification; iii) shock misaggregation is minor. Because shocks to the parameters are i.i.d., timing distortions are also small.

7 Time varying financial frictions?

We apply the technology we developed to the unconventional monetary policy model of Gertler and Karadi (GK) (2010). Our contribution is three fold. We provide likelihood based estimates of the parameters specific to the model (the fraction of capital that can be diverted by banks λ, the proportional transfer to entering bankers ω, and the survival probability of bankers θ), that the authors have informally calibrated to match a steady state spread, a steady state leverage and a notional length of bank activity; we use the diagnostics developed in the paper to gauge the extent of parameter variations; we estimate time variations in λ and compare responses to capital quality shocks in the fixed coefficient and the time varying coefficient models.

The equations of the GK model are summarized in appendix B. We use US data from 1985Q2 to 2014Q3 on the growth rate of output, growth rate of consumption, growth rate of leverage, and growth intermediary demand for assets (credit) and the spread. The spread is measured by the difference between BAA 10 years corporate bond yields and a 10 year treasury constant maturity and it is from the FRED, as are real personal consumption expenditures and GDP data. Leverage is from Haver and measures Tier 1 (core) capital as a percent of average total assets. Credit is measured as total loans (from Haver), scaled by size of US population.

Using a flat prior, the posterior mode estimates are $\lambda = 0.245$, $\theta = 0.464$, $\omega = 0.012$; the standard error are tight (0.0182, 0.0008, 0.0098) making the estimates highly significant. For comparison, GK calibrated these three parameters to $\lambda = 0.318$, $\theta = 0.972$, $\omega = 0.002$. In the GK model λ regulates private leverage: our estimate implies a higher steady state leverage than the one implied by the authors (our estimate is 3.32, GK is 1.38), which closer to the leverage found in the US in corporate and non-corporate business sectors over the sample. Our estimates also suggest that the survival probability of bankers is much lower than the one assumed by GK (about 10 years).

With these parameter estimates, we perform forecast error and wedge mispecification diagnostics. Table 6 indicates that the forecast errors of all equations but consumption are predictable and typically lagged consumption and lagged spread matter. The mean value of the Euler wedge is 0.02 with a standard error of 0.03; but both lag consumption and lag investment to output ratios significantly explain its movements (coefficients are respectively -0.10 and 0.72, with standard errors of 0.01 and 0.13). Thus, misspecification exists and parameter variation could be the reason for it.

Armed with this preliminary evidence, we estimate models allowing λ to be time varying. Time variations are specified as

$$\lambda_t = (1-\rho_\lambda)\lambda + \rho_\lambda \lambda_{t-1} + e_{t,\lambda} \qquad \text{Exogenous variations} \tag{39}$$

$$\lambda_t = (2 * \lambda_u - (\lambda_u - \frac{\lambda}{2}) * (\exp(-\phi_1 * (X_{t-1} - X^s)) + \exp(\phi_2 * (X_{t-1} - X^s))) + e_{t,\lambda} \qquad \text{Endogenous variations} \tag{40}$$

where X is net bank wealth N. Table 7 reports estimates of selected parameters

Equation	T-stat					F-stat
	Y_{t-1}	C_{t-1}	Credit$_{t-1}$	Leverage$_{t-1}$	Spread$_{t-1}$	
Y	0.84	2.61	0.24	0.52	10.00	4.39
C	-0.85	1.11	0.85	-0.65	0.33	1.26
Credit	1.06	2.61	1.65	-0.58	8.49	7.11
Leverage	-1.11	-2.50	-1.63	0.63	-8.25	7.04
Spread	-1.26	-3.06	-1.10	0.81	-8.46	8.16

Table 6: Regression diganostic for time variation. The left hand side of the regression is the forecast in the equation listed in the first column; the right hand side the variables listed in the second to the fifth column. Sample size T=116. Critical level of the F-stat(5,112)=2.56.

In the model with exogenously varying parameters, variations in λ_t are very persistent. Furthermore, estimates of λ, ω, θ are now larger making steady state leverage drop to about 2.9 and the lifetime of bankers to increase. When the endogenous specification is used, estimates of λ and ω further increase, making steady state leverage fall to 1.9, but bankers survival probability is roughly unchanged. The data seem to require a very strong asymmetric specification ($\phi_1 < \phi_2$) implying a strong negative relationship between the fraction of funds that bankers can divert and their net worth. Finally, note that the endogenous specification is superior to the specifications with exogenous time variations and fixed coefficients in a marginal likelihood sense.

Parameter	Time Invariant		Exogenous TVC		Endogenous TVC	
					Function of net worth	
h	0.43	(0.006)	0.19	(0.03)	0.09	(0.02)
λ	0.24	(0.01)	0.37	(0.03)	0.55	(0.03)
ω	0.01	(0.008)	0.02	(0.002)	0.11	(0.008)
θ	0.46	(0.009)	0.54	(0.01)	0.52	(0.02)
ρ_λ			0.99	(0.004)		
σ_λ			0.02	(0.002)	0.03	(0.003)
λ_u					0.98	(0.008)
ϕ_1					0.02	(0.007)
ϕ_2					0.15	(0.009)
Log ML	-167.97		1546.18		1628.69	

Table 7: Parameter estimates, Gertler and Karadi model.

To investigate how inference differs in the three estimated model we plot in figure 7 the responses of output, inflation, investment, net worth, leverage and the spread to a one percent capital quality shock. The constant coefficient specification closely replicate the dynamics presented by GK in their figure 3. There is a persistent decline in output and a temporary but strong decline in inflation. Investment temporarily falls but it then increase because capital is below its steady state. Bankers net worth falls and there is a sharp increase in the spread.

When we allow λ to be exogenously varying, the qualitative features of the responses are similar. Quantitatively, output falls more on impact but less in the short run; net worth falls less and the spread increases less in the short run. Thus, making λ exogenously time varying, reduces the ability of the model to capture recessionary effects on impact.

When variations are endogenous, the model possesses an additional mechanism of propagation of shocks since lower net worth implies higher share of funds diverted by banks and generally stronger accelerator dynamics. Since the dynamic responses of net worth are highly persistent, the spread persistently increases making investment increase less and output to fall more and more persistently relatively to the other two cases. Thus, neglecting endogenous variations in λ could impair our ability to correctly measure the effects of capital quality shocks.

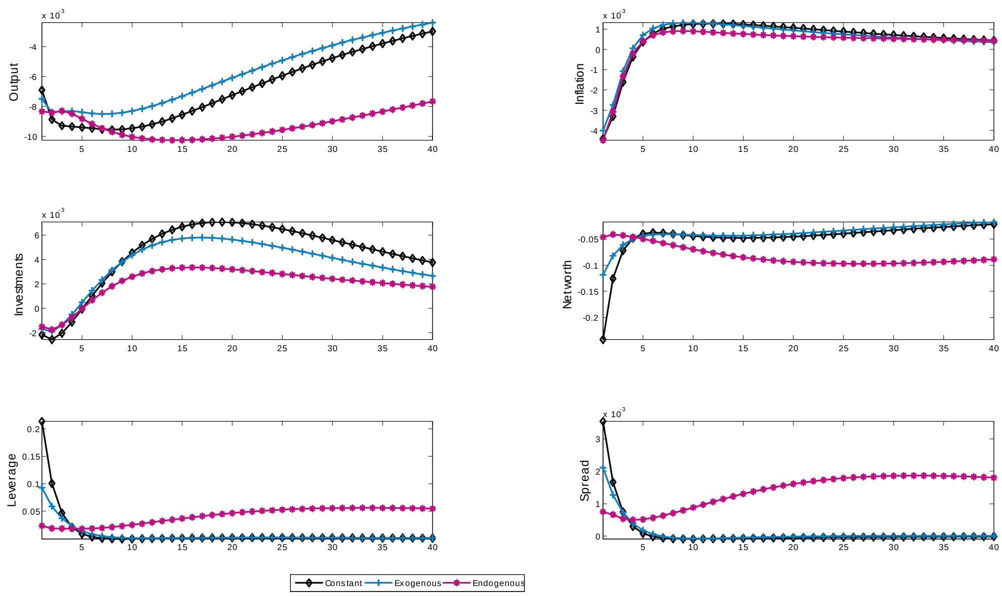

Figure 6: Dynamics in response to a capital quality shock

8 Conclusions

This paper has three main goals. It is interested in i) characterizing the decision rules of a DSGE when parameter variations are exogenous or endogenous, and in the latter case, when agents internalize or not the effects that their optimal decisions may have on parameter variations; ii) providing diagnostics to detect the misspecification driven by parameter variations; and iii) studying the consequences of using time invariant models when the parameters are time varying in terms of identification, estimation, inference, and policy analyses and comparing likelihood-based and SVAR-based estimates of the structural responses.

We show that if parameter variations are purely exogenous, the contemporaneous impact and the dynamics induced by structural shocks are the same as in a model with no parameter variations. However, if parameter variations are endogenous, the dynamics in responses to structural shocks may be different from the one of a constant coefficient model and the extent of the differences depends on the detail of the model. We provide simple and powerful diagnostics to detect the misspecification due to neglected parameter variations based on the optimality wedges of Chari et al. (2007) and the forecast errors. We also describe a marginal likelihood diagnostics which can help to recognize whether the detected time variations are of exogenous or endogenous nature.

We highlight certain parameter identification problem noted in the literature may be the results of misspecification due to neglected time variations. Our Monte Carlo study indicates that parameter and impulse response distortions may be large even for modest time variations in the parameters and that they tend to be stronger when variations are endogenous. It also shows that, when parameter variations are neglected, SVAR methods are competitive with more structural likelihood based methods as far are the responses to structural disturbances are concerned. Thus, the edge that likelihood based methods have when the model is correctly specified vanishes when misspecification is present.

In the context of the Gertler and Karadi (2010) model, we show that there is evidence that the parameter regulating the amount of moral hazard is indeed time varying, and that variations are possibly linked to the amount of net worth bankers have. When we allow for this link, the fit dramatically improves, primarily because the model acquires an additional propagation channel which makes spread and thus output responses stronger and much more persistent.

Our analysis provides researcher with a new set of tools that can help them to assess the quality of their models and respecify certain problematic features. There are a few interesting questions we do not study in this paper. How do we distinguish a model with time varying parameters from a model measurement errors are present? Is a model with m structural shocks observationally equivalent to a model with m_1 structural shocks and m_2 time varying parameters where $m_1 + m_2 = m$? To what extent parameter variations capture variations in the variances (or in higher moments) of the structural shocks? We leave these questions for future research.

9 References

Andreasen, M. (2012). On the effects of rare disasters and uncertainty shocks for risk premia in non-linear DSGE models. Review of Economic Dynamics, 15, 293-316.

Ascari, G. and A. Sbordone (2014). The macroeconomics of Trend Inflation. Journal of Economic Literature, 52, 679-739.

Brinca, P., Holter, H., P. Krusell, and L. Malafry (2014). Fiscal multipliers in the 21st century. EUI Schumann center working paper, 2014-119.

Canova, F. (2009). What Explains the Great Moderation in the U.S?: A Structural Analysis. Journal of the European Economic Association, 7, 697-721.

Canova, F. and L. Sala (2009). Back to square one: Identification issues in DSGE models. Journal of Monetary Economics, 56, 431-449.

Canova, F. and M. Paustian (2011). Business cycle measurement with some theory. Journal of Monetary Economics, 48, 345-361.

Chang, Y., Kim, S. and F. Schorfheide (2013). Labor market heterogeneity and the policy-(in)variance of DSGE model parameters'. Journal of the European Economic Association, 11,193-220.

Chari, V.V., Kehoe, P. and Mc Grattan, E. (2007). Business Cycle Accounting. Econometrica, 75, 781-836.

Cogley, T. and T. Yagihashi, (2010). Are DSGE approximating models invariant to policy shifts? The B.E. Journal of Macroeconomics (Contributions), Article 27.

Davig, T. and E. Leeper (2006). Endogenous monetary policy regime changes. NBER International Seminar in Macroeconomics, 345-391. National Bureau of Economic Research.

Del Negro, M. and F. Schorfheide (2004). Prior from general equilibrium models for VARs. International Economic Review, 45, 643-673.

Dueker, M. Fisher, A. and Dittman, R. (2006). Stochastic Capital Depreciation and the Co-movements of Hours and Productivity. Berkeley Journals: Topics in Macroeconomics, vol 6,3, article 6.

Fernandez Villaverde, J. and J. Rubio Ramirez (2004). Comparing dynamic equilibrium economies to the data. Journal of Econometrics, 123, 153-187.

Fernandez Villaverde, J. and J. Rubio Ramirez (2007). How structural are structural parameters. NBER Macroeconomics Annual, 22, 83-132.

Galvao, A. , Giraitis, L., Kapetanios, G., and K. Petrova (2014) Local Bayesian estimation and forecasting with a TVP model, manuscript

Gertler, M. and Karadi, P. (2010) A Model of unconventional monetary policy. Journal of Monetary Economics, 58, 17-34.

Huang, N. (2014). Weak Inference for DSGE models with time varying parameters. Boston College, manuscript.

Hurwicz, L. (1962). On the structural form of interdependent systems, in E. Nagel, P. Suppes, A. Tarski (eds.) Logic, Methodology and Philosophy of science: proceedings of a 1960 international congress. Stanford University Press.

Koop, G., Pesaran, H., and R. Smith (2013). On the identification of Bayesian DSGE models. Journal of Business and Economic Statistics, 31, 300-314

Kulish, M. and A. Pagan (2014) Estimation and Solution of models with expectation and structural changes, Dynare working paper 34.

Ireland, P. (2004). Technology shocks in a New Keynesian model. Review of Economics and Statistics, 83, 926-933.

Ireland, P. (2007). Changes in the Federal Reserve Inflation target: Causes and Consequences. Journal of Money Credit and Banking, 39, 1851-1882.

Liu, Z., Waggonner, D. and T. Zha (2011). Sources of Macroeconomic fluctuations: A regime switching DSGE approach. Quantitative Economics, 2, 251-301.

Magnusson, L. and S. Mavroeidis (2014). Identification using stability restrictions. Econometrica 82, 1799-1851.

Meier S. and C.Sprengler (forthcoming). Temporal Stability of Time preferences. Review of Economics and Statistics.

Mutschler, W. (2014) Identification of DGSE Models - the effect of higher order Approximation and Pruning, University of Munster, manuscript.

Rios Rull J. V. and R. SantaEularia-Llopis, (2010). Redistributive shocks and productivity shocks. Journal of Monetary Economics,37, 931-948.

Schmitt Grohe, S. and M. Uribe (2003). Closing small open economy models. Journal of International Economics, 62, 161-185.

Schmitt Grohe, S. and M. Uribe (2004). Solving Dynamic General Equilibrium Models using second order approximation to the policy function. Journal of Economic Dynamics and Control, 28, 755-775.

Seoane, H. (2014) Parameter Drifts, Misspecification and the Real Exchange Rate in Emerging Countries, manuscript, Universidad Carlos III Madrid.

Stock, J. and Watson, M. (1996). Evidence of Structural Instability in Macroeconomic Relationships. Journal of Business and Economic Statistics,14, 11-30.

Vavra, J. (2014). Time varying Phillips curve, NBER working paper 19790.

Appendix A: Additional Monte Carlo figures and tables

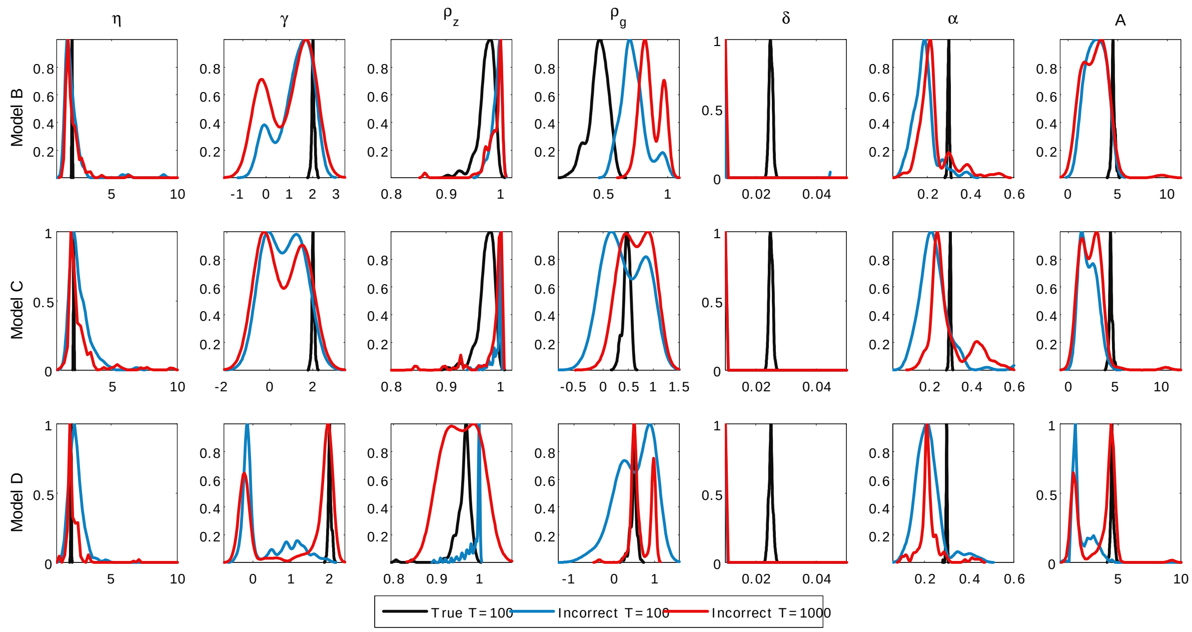

Figure A1:Density of estimates; DGP1 (time variations explain little of output variance).

True Parameter	Correct Mean	Time invariant Mean	5th percentile	95th percentile	Time invariant Mean	5th percentile	95th percentile
	T=150	T=150			T=1000		
DGP Model B							
$\eta = 2.0$	2.00	2.29	1.53	3.87	2.45	1.61	3.09
$\gamma = 2.0$	2.01	1.11	-0.33	2.06	0.25	-0.27	1.95
$\rho_z = 0.9$	0.94	0.99	0.96	1.00	0.99	0.97	1.00
$\rho_g = 0.5$	0.47	0.76	0.62	0.96	0.91	0.79	0.98
$\delta = 0.025$	0.03	0.01	0.01	0.03	0.01	0.01	0.01
$\alpha = 0.3$	0.30	0.19	0.11	0.41	0.21	0.10	0.34
$A = 4.5$	4.53	2.73	1.33	4.14	1.80	1.14	4.16
DGP Model C							
$\eta = 2.0$	2.00	3.40	1.56	7.51	5.19	1.77	22.90
$\gamma = 2.0$	2.00	-0.08	-0.32	0.73	-0.19	-0.35	0.35
$\rho_z = 0.9$	0.88	0.99	0.93	1.00	0.99	0.90	1.00
$\rho_g = 0.5$	0.48	0.56	0.08	0.97	0.91	0.59	0.98
$\delta = 0.025$	0.02	0.02	0.01	0.07	0.02	0.01	0.07
$\alpha = 0.3$	0.30	0.26	0.15	0.34	0.26	0.19	0.35
$A = 4.5$	4.50	1.71	1.25	2.77	2.27	1.24	8.17
DGP Model D							
$\eta = 2.0$	2.00	3.05	1.68	4.59	2.40	1.98	4.81
$\gamma = 2.0$	2.00	-0.06	-0.28	0.54	1.63	-0.27	1.98
$\rho_z = 0.9$	0.88	0.98	0.90	1.00	0.92	0.91	1.00
$\rho_g = 0.5$	0.47	0.42	-0.46	0.96	0.50	0.32	0.97
$\delta = 0.025$	0.02	0.01	0.01	0.03	0.01	0.01	0.01
$\alpha = 0.3$	0.30	0.23	0.15	0.32	0.21	0.13	0.27
$A = 4.5$	4.49	1.91	1.45	3.57	4.10	1.65	4.51

Table A1:Distributions of estimates, Parameter variations explain 20 percent of output variance.

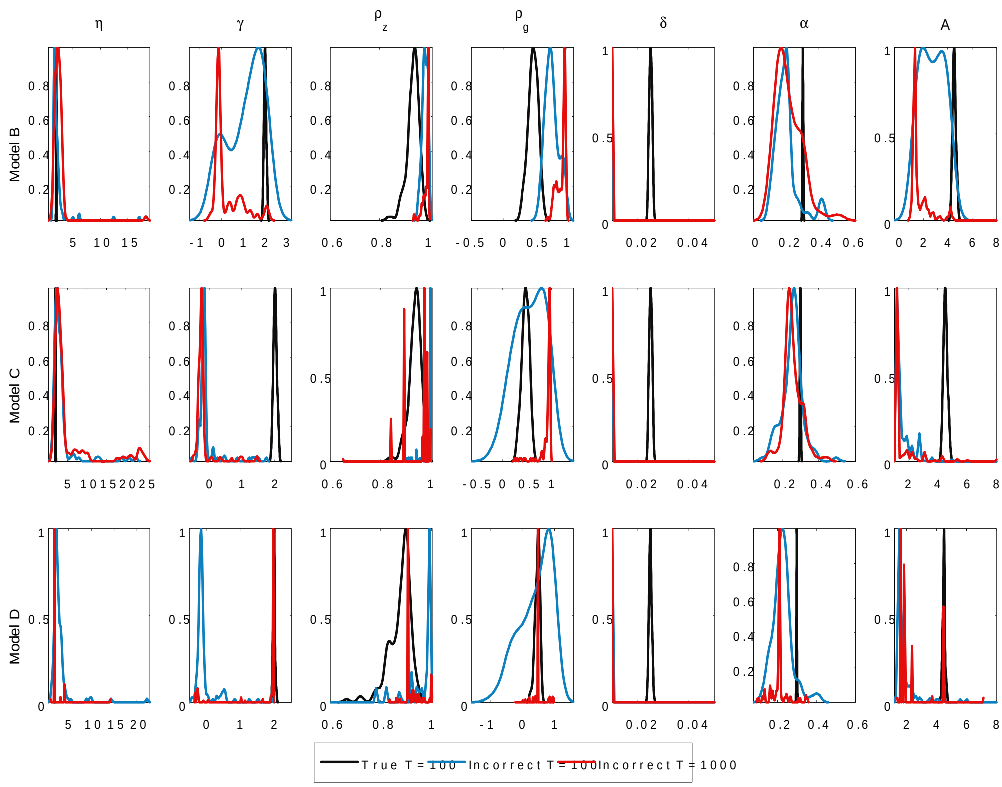

Figure A2:Density of estimates (time variations explain 20 percent of output variance)

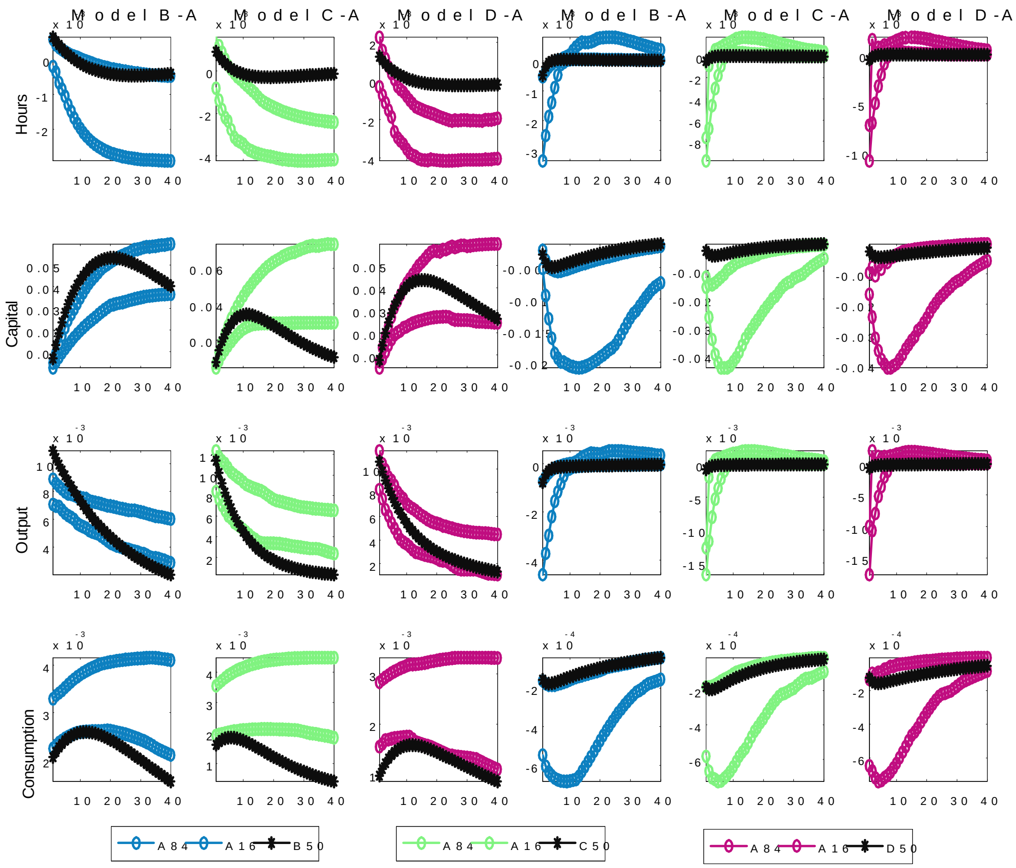

Figure A3:Impulse responses; parameter variations explain 20 percent of output variance

Variable	Technology	Government	Technology	Government
	Model B		Time invariant	
Y	81.300	0.100	0.998	0.006
C	55.300	0.100	0.998	0.002
N	15.600	0.400	0.978	0.025
K	40.600	0.100	0.994	0.008
	Model C		Time invariant	
Y	81.900	0.100	0.927	0.082
C	26.500	0.100	0.999	0.001
N	5.400	0.400	0.966	0.039
K	37.400	0.100	0.974	0.030
	Model D		Time invariant	
Y	82.200	0.100	0.936	0.072
C	32.800	0.100	0.996	0.008
N	10.200	0.500	0.928	0.079
K	60.000	0.400	0.979	0.028

Table A2: Variance decomposition, Parameter variations explain 20 percent of output variance.

Appendix B : The equations of Gertler and Karadi model

$$\exp(\varrho_t) = (\exp(C_t) - h\exp(C_{t-1}))^{-\sigma} - \beta h(\exp(C_{t+1}) - h\exp(C_t))^{-\sigma} \quad (41)$$

$$1 = \beta\exp(R_t)\exp(\Lambda_{t+1}) \quad (42)$$

$$\exp(\Lambda_t) = \frac{\exp(\varrho_t)}{\exp(\varrho_{t-1})} \quad (43)$$

$$\chi * \exp(L_t)^{\varphi} = \exp(\varrho_t)\exp(P_{m,t})(1-\alpha)\frac{\exp(Y_t)}{\exp(L_t)} \quad (44)$$

$$\exp(\nu_t) = (1-\theta)\beta\exp(\Lambda_{t+1})(\exp(R_{k,t+1}) - \exp(R_t)) + \beta\exp(\Lambda_{t+1})\theta\exp(x_{t+1})\exp(\nu_{t+1}) \quad (45)$$

$$\exp(\eta_t) = (1-\theta) + \beta\exp(\Lambda_{t+1})\theta\exp(z_{t+1})\exp(\eta_{t+1}) \quad (46)$$

$$\exp(\phi_t) = \frac{1}{(1-\psi_t)}\frac{\exp(\eta_t)}{\lambda - \exp(\nu_t)} \quad (47)$$

$$\exp(z_t) = (\exp(R_{k,t}) - \exp(R_{t-1}))(1-\psi_{t-1})\exp(\phi_{t-1}) + \exp(R_{t-1}) \quad (48)$$

$$\exp(x_t) = \frac{\exp(\phi_t)(1-\psi_t)}{(\exp(\phi_{t-1})(1-\psi_{t-1}))}\exp(z_t) \quad (49)$$

$$\exp(K_t) = \frac{\exp(\phi_t)\exp(N_t)}{\exp(Q_t)} \quad (50)$$

$$\exp(N_t) = \exp(Ne_t) + \exp(Nn_t) \quad (51)$$

$$\exp(Ne_t) = \theta\exp(z_t)\exp(N_{t-1})\exp(-e_{Ne,t}) \quad (52)$$

$$\exp(Nn_t) = \omega(1-\psi_{t-1})\exp(Q_t)\exp(\xi_t)\exp(K_{t-1}) \quad (53)$$

$$\exp(R_{k,t}) = (\exp(P_{m,t})\alpha\frac{\exp(Y_{m,t})}{\exp(K_{t-1})} + \exp(\xi_t) * (\exp(Q_t) - \frac{\exp(\delta))}{\exp(Q_{t-1})} \quad (54)$$

$$\exp(Y_{m,t}) = \exp(a_t) * (\exp(\xi_t) * \exp(U_t) * \exp(K_{t-1}))^{\alpha} * \exp(L_t)^{1-\alpha} \quad (55)$$

$$\begin{aligned}\exp(Q_t) &= 1 + 0.5\eta_i(\frac{(In_t + I^s)}{(In_{t-1} + I^s} - 1)^2 + \eta_i(\frac{(In_t + I^s)}{(In_{t-1} + I^s} - 1)\frac{(In_t + I^s)}{(In_{t-1} + I^s s)} \\ &- \beta\exp(\Lambda_{t+1})\eta_i(\frac{(In_{t+1} + I^s)}{(In_t + I^s} - 1)(\frac{(In_{t+1} + I^s)}{(In_t + I^s)})^2 \end{aligned} \quad (56)$$

$$\exp(\delta) = \delta_c + b/(1+\zeta) * \exp(U_t)^{1+\zeta} \quad (57)$$

$$\alpha\frac{\exp(Y_m)}{\exp(U_t)} = \frac{b\exp(U_t)^{\zeta}\exp(\xi_t) * \exp(K_{t-1})}{\exp(P_{m,t})} \quad (58)$$

$$In_t = \exp(I_t) - \exp(\delta) * \exp(\xi_t) * \exp(K_{t-1}) \quad (59)$$

$$\exp(K_t) = \exp(\xi_t) * \exp(K_{t-1}) + In_t \quad (60)$$

$$\exp(G_t) = G^s * \exp(g_t) \quad (61)$$

$$\exp(Y_t) = \exp(C_t) + \exp(G_t) + \exp(I_t) + 0.5\eta_i(\frac{(In_t + I^s)}{(In_{t-1} + I^s} - 1)^2(In_t + I^s) + \tau\psi\exp(K_t) \quad (62)$$

$$\exp(Y_{m,t}) = \exp(Y_t) * \exp(D_t) \quad (63)$$

$$\begin{aligned}
\exp(D_t) &= \gamma * \exp(D_{t-1}) * \exp(infl_{t-1})^{-\gamma_P * \epsilon} \exp(infl_t)^{\epsilon} \\
&+ (1-\gamma)((1-\gamma \exp(infl_{t-1})^{\gamma_P(1-\gamma)} \exp(infl_t)^{\gamma-1})/(1-\gamma))^{-\epsilon/(1-\gamma)} && (64) \\
\exp(X_t) &= 1/\exp(P_{m,t}) && (65) \\
\exp(F_t) &= \exp(Y_t) * \exp(P_{m,t}) + \beta\gamma \exp(\Lambda_{t+1}) \exp(infl_{t+1})^{\epsilon} (\exp(infl_t))^{-\epsilon\gamma_P} \exp(F_{t+1}) && (66) \\
\exp(Z_t) &= \exp(Y_t) + \beta\gamma \exp(\Lambda_{t+1}) \exp(infl_{t+1})^{\epsilon-1} \exp(infl_t)^{\gamma_P*(1-\epsilon)} \exp(Z_{t+1}) && (67) \\
\exp(infl_t^*) &= \frac{\epsilon}{\epsilon-1} \frac{\exp(F_t)}{\exp(Z_t)} \exp(infl_t) && (68) \\
(\exp(infl_t))^{1-\epsilon} &= \gamma \exp(infl_{t-1})^{\gamma_P(1-\epsilon)} + (1-\gamma)(\exp(infl_t^*))^{1-\epsilon} && (69) \\
\exp(i_t) &= \exp(R_t) * \exp(infl_{t+1}) && (70) \\
\exp(i_t) &= \exp(i_{t-1})^{\rho_i} (\beta^{-1} \exp(infl_t)^{\kappa_\pi} * (\exp(X_t)/(\epsilon/(\epsilon-1)))^{\kappa_y})^{1-\rho_i} \exp(e_{i,t}) && (71) \\
\psi_t &= \kappa * (R_{k,t+1} - R_t - R_k^s + R^s) + e_{\psi,t} && (72) \\
a_t &= \rho_a * a_{t-1} - \sigma_a * e_{a,t} && (73) \\
\xi_t &= \rho_\xi * \xi_{t-1} - \sigma_\xi * e_{\xi,t} && (74) \\
g_t &= \rho_g * g_{t-1} - e_{g,t} && (75) \\
e_{\psi,t} &= \rho_\psi * e_{\psi,t-1} + e_{\psi,t}; && (76)
\end{aligned}$$