Detecting and Analyzing the Effects of Time-Varying Parameters in DSGE Models

Fabio Canova, Norwegian Business School, and CEPR
Filippo Ferroni, Federal Reserve Bank of Chicago
Christian Matthes, Federal Reserve Bank of Richmond *

April 16, 2019

Abstract

We study how structural parameter variations affect the decision rules and economic inference. We provide diagnostics to detect parameter variations and to ascertain whether they are exogenous or endogenous. A constant parameter model poorly approximates a time-varying DGP, except in a handful of relevant cases. Linear approximations do not produce time-varying decision rules; higher order approximations can do this only if parameter disturbances are treated as decision rule coefficients. Structural responses are time invariant regardless of order of approximation. Adding endogenous variations to the parameter controlling leverage in Gertler and Karadi's (2010) model substantially improves the fit of the model.

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

JEL Classification: C10, E27, E32.

*We thank Jesus Fernandez Villaverde (the editor), two anonymous referees, Stephane Bonhomme, Francesco Bianchi, Ferre de Graeve, Marco del Negro, James Hamilton, Lars Hansen, Michele Lenza, Frank Schorfheide, Harald Uhlig, and Tao Zha as well as participants of many seminars and conferences for their comments and suggestions. Canova acknowledges the financial support from the Spanish Ministerio de Economia y Competitividad through the grants ECO2012-33247; ECO2015-68136-P; and FEDER, UE. The views presented in this paper are not necessarily those of the Federal Reserve Bank of Richmond, the Federal Reserve Bank of Chicago, or the Federal Reserve System. Earlier versions of this paper circulated under the title "Approximating Time Varying Structural Models with Time Invariant Structures."

1 Introduction

In macroeconomics it is standard to work with 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 crucial, for example, to distinguish structural from reduced-form models and to conduct correctly designed policy counterfactuals.

Dueker et al. (2007), Fernandez-Villaverde and Rubio-Ramirez (2007), Canova (2009), Rios-Rull and Santaeularia-Llopis (2010), Liu et al. (2011), Vavra (2014), Dew-Backer (2014), Meier and Sprengler (2015), Seoane (2016), Castelnuovo and Pellegrino (2018) among others, have shown that the parameters of dynamic stochastic general equilibrium (DSGE) models are not time invariant and that variations are small but persistent. Parameter variations do not necessarily imply that DSGE models are not structural (see, e.g., Cogley and Yagihashi, 2010, Chang et al., 2013, Schmitt-Grohe and Uribe, 2003, Hansen and Sargent, 2010, and Cogley et al., 2015), but they create concerns about the economic interpretation of the results.

Recently, DSGE models with time-varying parameters have begun to appear. In modeling time variations, investigators have followed the vector autoregression (VAR) literature: parameter variations are assumed to be exogenous, drifting smoothly as independent random walks (as in Cogley and Sargent, 2005; and Primiceri, 2005) or switching between a finite number of states (as in Sims and Zha, 2006). Many economic questions, however, hint at the possibility that parameter variations may instead be endogenous. For example, does a central bank react to inflation in the same way in an expansion as in a contraction (see Davig and Leeper, 2006)? Do fiscal multipliers depend on the level of inequality (see, e.g., Brinca et al., 2016)? Are households as risk averse when they are wealthy as when they are poor? Clearly, analyses conducted under the assumption of time invariant models or exogenous rather than endogenous forms of time variations may lead to misleading conclusions regarding, e.g., the welfare costs of business cycles, and to invalid policy prescriptions.

This paper is concerned with the consequences of time-varying misspecification, that is, the misspecification induced by neglected parameter variations or incorrect assumptions about time variation in structural parameters. We focus on DSGE models with smoothly evolving parameters and work with first and higher order perturbed solutions. We characterize the approximate decision rules when parameter variations are present; discuss the conditions under which constant parameter models provide a good approximation to the data generating process (DGP) and the potential distortions that emerge when the DGP features parameter variations. We also examine whether the evidence produced by time-varying parameter VAR and DSGE models can be matched and provide two diagnostics to detect time-varying misspecification.

The literature is generally silent on these issues. Earlier work by Parkin (1988) studied whether the parameters of a Real Business Cycle (RBC) model are a function of one particular omitted variable. Seoane (2016) finds endogenous parameter variations, interprets it as evidence of misspecification, and proposes a more complex constant parameter model that fits the data better. Kulish and Pagan (2017) characterize the decision rules and the likelihood function of a DSGE model when predictable structural breaks occur. Magnusson and Mavroedis (2014) examine how time variations in certain parameters may affect the identification of other structural parameters. Justiniano and Primiceri (2008) and Fernandez-Villaverde et al. (2011) investigate to what extent variations in the shock volatility matter for real variables.

Given lack of work in the area, understanding how models with smoothly varying parameters work, documenting the distortions induced by estimating constant parameters models, and designing diagnostics to detect parameter variations are prerequisites to answer broader questions - such as the distinction between shocks and parameters when the latter are allowed to vary or the form of structural parameter variations (smoothly changing, Markov switching, etc.) that best captures patterns in the data.

The next section presents a motivating example to set ideas. We consider a simple RBC model where certain structural parameters are constant or time varying and, in the latter case, we allow variations to be either exogenous or endogenous. We examine what time variations imply for the optimality conditions of the problem and discuss their implications for responses to technology disturbances. Constant and varying parameter models generally produce different dynamic responses because income and substitution effects are altered. Endogenous parameter variations add to the uncertainty of the environment, making agents prefer to consume more today relative to the future for the same transitory fluctuations in income.

Section 3 formally shows that when parameter variations are present, it is generally impossible for a constant parameter model to "reasonably" approximate the DGP. When linearized solutions are considered, there are two special cases when the dynamics in response to structural shocks will be isomorphic, but even in these cases, historical and variance decomposition exercises are distorted. When second order solutions are considered, structural responses in time-varying and constant-parameter models are proportional only when parameter variations are exogenous . For higher order solutions, the structural responses will be highly distorted.

Linear approximations of time-varying parameter models do not produce time-varying decision rules and higher order approximations can do this only if parameter disturbances are interpreted as (reduced form) decision rule coefficients. Still, regardless of order of approximation employed, structural responses will be time invariant. Thus, smoothly varying parameter VAR are not the natural reduced form counterpart of the smoothly varying DSGE models we consider.

In Section 4, we design diagnostics to detect time-varying misspecification. In the context of a Monte Carlo exercise, we show that the diagnostics are able to signal potential problems and to detect the true DGP with high probability.

Section 5 focuses on linear approximations and briefly discusses the identification and inferential repercussions that neglected time variations may have for the estimation time-invariant parameters. Supporting evidence for this section is in the online appendix. In the context of a Real Business Cycles (RBC) example, we show that pathologies occur: the likelihood is flattened, twisted, and moved away from the true parameter values. Moreover, there are important biases in the estimates of the parameters controlling shock persistence and income and substitution effects, which do not die away as sample size increases. Decision rule misspecification and shock misaggregation account for the distortions.

In Section 6, we estimate a few structural parameters of Gertler and Karadi's (GK) (2010) model of unconventional monetary policy, apply the diagnostics to detect parameter variations, and estimate versions of the model where the bank's moral hazard parameter is allowed to vary over time. We find that a fixed parameter model is misspecified; that making parameter variations an endogenous function of net worth is preferable; and that the dynamic effects of capital quality shocks on the spread and on bank net worth may be different than previously thought. Section 7 concludes.

2 A motivating example

To motivate our interest in time-varying parameter models, we use a closed-economy RBC model. For ease of presentation, we restrict the discussion to first order solutions. The representative agent maximizes:

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

subject to the sequence of constraints:

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

$$Y_t = \zeta_t K_{t-1}^{\alpha} N_t^{1-\alpha}, \tag{3}$$

where Y_t is output, C_t consumption, K_t the stock of capital, N_t is hours worked, and G_t is government expenditure. The system is perturbed by two exogenous structural disturbances: one to technology ζ_t and one to government spending G_t, both assumed to follow time-invariant AR(1) processes:

$$\begin{aligned} \ln \zeta_t &= (1-\rho_\zeta)\ln\zeta + \rho_\zeta \ln \zeta_{t-1} + e_t^{\zeta}, \\ \ln G_t &= (1-\rho_G)\ln G + \rho_G \ln G_{t-1} + e_t^{G}, \end{aligned} \tag{4}$$

with variables without time subscript denoting steady state quantities. There are six structural parameters (α, the capital share; η, the risk-aversion coefficient; γ, the inverse of the Frisch elasticity of labor supply; A, the constant in utility; β_t, the time discount factor; and δ_t, the depreciation rate), and six auxiliary parameters (the steady-state values of government expenditure and of TFP, (ζ, G); their autoregressive parameters, (ρ_ζ, ρ_g); and their standard deviations (σ_ζ, σ_g)). Because we care about time-varying structural parameters, we let the auxiliary parameters be time invariant.

For illustration, we assume that α, γ, A, η are constant and allow β_t and δ_t to be time-varying. The law of motion of (β_t, δ_t) is described next. Dueker et al. (2007), Liu et al (2011), Gourio (2012), and Meier and Sprenger (2015) provide evidence that these parameters are indeed varying over time. None of the arguments here depend on which parameter is allowed to be time-varying. The optimality conditions are:

$$
\begin{aligned}
AC_t^{\eta}N_t^{\gamma} &= (1-\alpha)Y_t/N_t, && (5)\\
\beta_t C_t^{-\eta} &= E_t\left(\beta_{t+1}C_{t+1}^{-\eta}(\frac{\alpha Y_{t+1}}{K_{t+1}}+1-\delta_{t+1})\right) &&\\
&+ E_t\left(\frac{\partial\beta_{t+1}}{\partial K_t}u(C_{t+1},N_{t+1})-\frac{\partial\delta_{t+1}}{\partial K_t}K_t\right), && (6)\\
Y_t &= C_t+K_t-(1-\delta_t)K_{t-1}+G_t, && (7)\\
Y_t &= \zeta_t K_{t-1}^{\alpha}N_t^{1-\alpha}. && (8)
\end{aligned}
$$

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 an additional effect due to variations in the derivatives of β_{t+1} and δ_{t+1} with respect to the endogenous states (see equation (6)). As the optimality conditions clearly show, time-varying parameters cannot be treated as "wedges" in the sense of Chari et al. (2007), because they imply cross-equation restrictions. To understand what time variations in β_t, δ_t imply, we consider a number of cases.

2.1 Model A: Constant parameters.

Let $\beta_t = \beta$ and $\delta_t = \delta$. The optimality conditions are

$$
\begin{aligned}
E_t\left[f(X_{t+1},X_t,X_{t-1},Z_{t+1},Z_t,\mu)\right] &= \\
E_t\begin{pmatrix} AC_t^{\eta}N_t^{\gamma+1}-(1-\alpha)Y_t \\ C_t^{-\eta}-E_t\beta C_{t+1}^{-\eta}\left(\alpha Y_{t+1}/K_t+1-\delta\right) \\ Y_t-C_t+K_t-(1-\delta)K_{t-1}-G_t \\ Y_t-\zeta_t K_{t-1}^{\alpha}N_t^{1-\alpha}\end{pmatrix} &= 0, \qquad (9)
\end{aligned}
$$

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

$$
\frac{K}{Y}=\frac{\alpha}{\delta-1+1/\beta};\ \frac{C}{Y}=1-\delta\frac{K}{Y}-\frac{G}{Y};\ \frac{N}{Y}=\zeta^{\frac{1}{1-\alpha}}\left(\frac{K}{Y}\right)^{\frac{\alpha}{\alpha-1}};\ Y=\left[\frac{A}{(1-\alpha)}\left(\frac{C}{Y}\right)^{\eta}\left(\frac{N}{Y}\right)^{1+\gamma}\right]^{-\frac{1}{\eta+\gamma}}. \qquad (10)
$$

The (first order) decision rules are $(X_t - X_A) = P(\mu)(X_{t-1} - X_A) + Q(\mu)Z_t$, where X_A are the steady states and μ the vector of model's parameters.

2.2 Model B: Exogenous parameter variations

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

$$
\begin{aligned}
u_{d,t+1} &= \rho_d u_{d,t} + \epsilon_{d,t+1}, && (11)\\
u_{\delta,t+1} &= \rho_\delta u_{\delta,t} + \epsilon_{\delta,t+1}. && (12)
\end{aligned}
$$

We restrict the support of ϵ so that variations in d_t generate finite expected utility [1]. Here, $\frac{\partial(\beta_{t+1}/\beta_t)}{\partial K_t} = \frac{\partial \delta_{t+1}}{\partial K_t} = 0$ and the optimality conditions are

$$\begin{aligned} E_t\left[f(X_{t+1}, X_t, X_{t-1}, Z_{t+1}, Z_t, \Theta_{t+1}, \Theta_t)\right] &= \\ E_t \begin{pmatrix} AC_t^{\eta} N_t^{\gamma+1} - (1-\alpha)Y_t \\ 1 - d_t C_{t+1}^{-\eta}/C_t^{-\eta}\left(\alpha Y_{t+1}/K_t + 1 - \delta_{t+1}\right) \\ Y_t - C_t - K_t + (1-\delta_t)K_{t-1} - G_t \\ Y_t - \zeta_t K_{t-1}^{\alpha} N_t^{1-\alpha} \end{pmatrix} &= 0, \end{aligned} \tag{13}$$

With the selected parameterization, the steady-state values of β_t and δ_t are β and δ so that $(\frac{K}{Y}, \frac{C}{Y}, \frac{N}{Y}, Y)$ coincide with those of the constant parameter model and $X_A = X_B$. In addition, since time variations in (d_{t+1}, δ_{t+1}) are exogenous, the (first order) decision rules for model B are $(X_t - X_B) = P(\mu)(X_{t-1} - X_B) + Q(\mu)z_t + R(\mu, \mu_B)\epsilon_t$, where μ_B are the parameters specific to model B. Thus, shocks to the parameters play the role of additional disturbances, but they do not affect the transmission of structural shocks, which are regulated by the matrices P(μ) and Q(μ).

2.3 Model C: Endogenous parameter variations, no internalization

We assume that d_t and δ_t depend on K_t. Making the growth rate of the discount factor a function of the capital stock captures the idea that agents may have different saving rates depending on their level of wealth. Similarly, making the depreciation rate a function of the capital stock may capture "congestions" effects. We specify:

$$\Theta_{t+1} = [\Theta_u - (\Theta_u - \Theta_l) \circ e^{-\phi_a(K_t - K)}] + [\Theta_u - (\Theta_u - \Theta_l) \circ e^{\phi_b(K_t - K)}] + U_{\theta,t+1} \tag{14}$$

where $\circ$ is the Hadamart (elementwise) product, $\phi_a, \phi_b, \Theta_u, \Theta_l$ are (2×1) vectors of parameters, and $U_{\theta,t+1}$ is a zero mean, i.i.d. vector of shocks with bounder domain. We restrict $|| - \phi_a + \phi_b|| < M$ for some small M, so that expected utility exists. In (14), we specify parameter variations in the same way the literature has modeled, e.g., the relationship between the capital depreciation rate and capacity utilization (see, e.g., Justiniano and Primiceri, 2008). Our setup is more flexible and allows for endogenous and exogenous variations to simultaneously occur. Furthermore, depending on the choice of $\phi' s$, we can accommodate linear or quadratic relationships, which are symmetric or asymmetric. If we set $\Theta_l = (\beta/2,\, \delta/2)$, $X_C = X_A$.

We assume that agents treat the capital stock appearing in (14) as an aggregate variable. This assumption is similar to the 'small k - big k' situation or the 'internal- external' habit formation distinction encountered in fixed parameter rational expectations models. Since agents' first order conditions do not take into account the fact that their optimal capital choice changes d_t and δ_t, $\partial\beta_{t+1}/\partial K_t = \partial\delta_{t+1}/\partial K_t = 0$ and the equilibrium conditions are as in (13). Still, since in the

[1] A sufficient condition is that $\lim_{T\to\infty} P(\prod_i^t d_i < \exp(-at), \forall t > \bar{T}$, for any $a > 0)$=1.

aggregate d_t and δ_t depend on K_t, the (first order) decision rules for model C are $(X_t - X_C) = P(\mu, \mu_C)(X_{t-1} - X_C) + Q(\mu, \mu_C)z_t + R(\mu, \mu_C)\epsilon_t$, where μ_C are the parameters specific to model C. Here, shocks to the parameters still play the role of additional disturbances, but they may affect the transmission of structural shocks. It turns out that, with the parameterization used, variations in d_t, δ_t will affect P and Q if their law of motion is asymmetric, i.e. $\phi_{a,\beta} \neq \phi_{b,\delta}$ and/or $\phi_{a,\delta} \neq \phi_{b,\delta}$.

2.4 Model D: Endogenous parameter variations, internalization.

We still assume that time variations in (β_t, δ_t) are as in equation (14). Contrary to model C, agents internalize the effects their capital decisions have on parameter variations. The relevant derivatives are

$$\begin{aligned} d'_{t+1} &\equiv \partial d_{t+1}/\partial K_t = -(\beta_u - \beta/2)[-\phi_{a,\beta}e^{-\phi_{a,\beta}(K_t-K)} + \phi_{b,\beta}e^{\phi_{b,\beta}(K_t-K)}] \quad (15)\\ \delta'_{t+1} &\equiv \partial \delta_{t+1}/\partial K_t = -(\delta_u - \delta/2)[-\phi_{a,\delta}e^{-\phi_{a,\delta}(K_t-K)} + \phi_{b,\delta}e^{\phi_{b,\delta}(K_t-K)}]. \quad (16)\end{aligned}$$

In order for $X_D = X_A$ we restrict $\phi_{a,\beta} = \phi_{b,\beta} = \phi_\beta$, $\phi_{a,\delta} = \phi_{b,\delta} = \phi_\delta$. The optimality conditions now are:

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

$$E_t \begin{pmatrix} AC_t^\eta N_t^{\gamma+1} - (1-\alpha)Y_t \\ 1 - d'_t\ u(C_{t+1}, N_{t+1})/C_t^{-\eta} - d_t\ C_{t+1}^{-\eta}/C_t^{-\eta}(\alpha Y_{t+1}/K_{t+1} + 1 - \delta_{t+1}) + \delta'_{t+1}K_t) \\ Y_t - C_t - K_t + (1-\delta_t)K_{t-1} - G_t \\ Y_t - \zeta_t K_{t-1}^\alpha N_t^{1-\alpha} \end{pmatrix}, \quad (17)$$

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

$$\begin{pmatrix} d_{t+1} \\ \delta_{t+1} \\ d'_{t+1} \\ \delta'_{t+1} \end{pmatrix} = \Phi(\Theta, K_t, U_{t+1}) = \begin{pmatrix} 2d_u - (d_u - \beta/2)[e^{-\phi_\beta(K_t-K)} + e^{\phi_\beta(K_t-K)}] + U_{\beta,t+1} \\ 2\delta_u - (\delta_u - \delta/2)[e^{-\phi_\delta(K_t-K)} + e^{\phi_\delta(K_t-K)}] + U_{\delta,t+1} \\ -(d_u - \beta/2)\phi_1[-e^{-\phi_\beta(K_t-K)} + e^{\phi_\beta(K_t-K)}] \\ -(\delta_u - \delta/2)\phi_3[-e^{-\phi_\delta(K_t-K)} + e^{\phi_\delta(K_t-K)}] \end{pmatrix}. \quad (18)$$

The (first order) decision rules for model D are $(X_t - X_D) = P(\mu, \mu_D)(X_{t-1} - X_D) + Q(\mu, \mu_D)z_t + R(\mu, \mu_D)\epsilon_t$, where μ_D are the parameters specific to model D. As in models B and C, shocks to the parameters play the role of additional disturbances but they will affect the transmission of structural shocks regardless of whether the relationship between parameters and endogenous variables is symmetric or not.

2.5 Impulse responses

Why is the transmission of structural shocks z_t in models C and D potentially different from the transmission in models A and B? To give some intuition, we compute responses to technology shocks. For the parameters common to all models, we choose $\alpha = 0.30$,

$\beta = 0.99$, $\delta = 0.025$, $\gamma = 2$, $\eta = 2$, $A = 4.50, \zeta = 1, \rho_\zeta = 0.90$, $\sigma_\zeta = 0.00712$, $G = 0.28$, $\rho_g = 0.50$, and $\sigma_g = 0.052$. For the other parameters, we choose the following:

- For μ_B we select $\rho_\beta = 0.90, \rho_\delta = 0.80$, $\sigma_\beta = 0.0008$, and $\sigma_\delta = 0.01$.
- For μ_C, we select $\phi_{a,\beta} = 0.01, \phi_{b,\beta} = 0.03, \phi_{a,\delta} = 0.2, \phi_{b,\delta} = 0.1$, $\sigma_d = 0.008$, $\sigma_\delta = 0.005$, $\beta_u = 0.999$, and $\delta_u = 0.025$.
- For μ_D, we select $\phi_{a,\beta} = 0.001, \phi_{b,\beta} = 0.016, \phi_{a,\delta} = 0.2, \phi_{b,\delta} = 0.1$, $\sigma_d = 0.009$, $\sigma_\delta = 0.001$, $\beta_u = 0.999$, and $\delta_u = 0.025$.

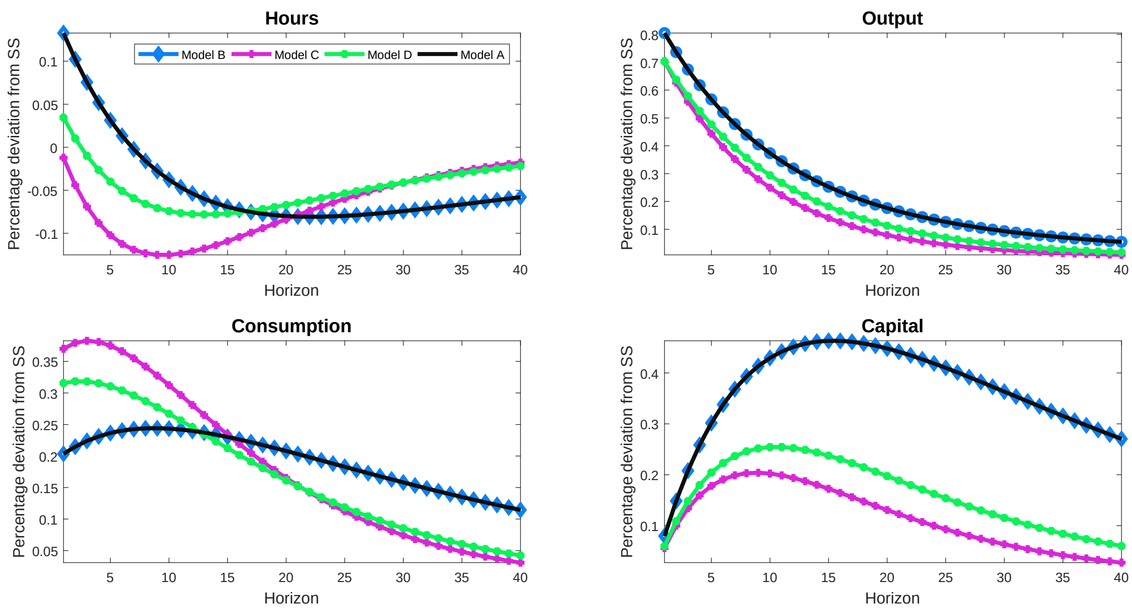

Figure 1: Responses to technology shocks, first order approximation.

Figure 1 reports the responses of hours, capital, consumption, and output to technology shocks in percentage deviation from the steady states in the four models. The sign of the responses is unchanged by the presence of parameter variations and, as expected from the above discussion, the dynamics in models A and B are the same. The shape and persistence of hours, consumption, and capital responses in models C and D instead differ because the income and substitution effects are altered. In particular, in response to technology shocks, agents work and save less and consume more in models with endogenously varying parameters. Thus, endogenous parameter variations are similar to uncertainty shocks. Note that the responses of model D are generally between those of models C and A because agents internalize the second order effects that lower capital accumulation has on the endogenous variables.

3 Our framework of analysis

The class of DSGE models we are interested in studying is:

$$\begin{aligned} 0 &= E_t\left[f(X_{t+1}, X_t, X_{t-1}, Z_{t+1}, Z_t, \Theta_{t+1}, \Theta_t)\right], && (19)\\ Z_{t+1} &= \Psi(Z_t, \sigma\Sigma_z\epsilon^z_{t+1}), && (20)\\ \Theta_{t+1} &= \Phi(\Theta, X_t, U_{t+1}), && (21)\\ U_{t+1} &= \Omega(U_t, \sigma\Sigma_u\epsilon^u_{t+1}), && (22)\end{aligned}$$

(19) are the equilibrium conditions, where X_t is an $n_x \times 1$ vector of endogenous variables; Z_t an $n_z \times 1$ vector of strictly exogenous variables; and Θ_t a vector of possibly time-varying structural parameters. Since the distinction between variables and parameters is blurred when we allow for parameter variations, we use the convention that parameters are the variables typically assumed to be constant by economists (discount factor, Frisch elasticity of substitution, etc.). (20) is the law of motion of the exogenous variables; ϵ^z_{t+1} is a $n_\epsilon \times 1$ vector of independent identically distributed (i.i.d.) structural disturbances with mean zero and identity covariance matrix, $n_z \geq n_\epsilon$; $\sigma \geq 0$ is an auxiliary scalar; and Σ_ϵ is a known $n_\epsilon \times n_\epsilon$ matrix. (21) is the law of motion of the structural parameters; U_{t+1} is a $n_u \times 1$ vector of exogenous disturbances, and Θ is a vector of constants. (22) describes the evolution of the exogenous component of parameter variations U_{t+1}; 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. (20)-(22) are known to the agents when they optimize. We assume that f, Ψ, Φ and Ω are continuous and differentiable up to some order q, and that (20)-(22) induce stationary fluctuations in Z_{t+1} and U_{t+1}.

We posit that the decision rules are of the form:

$$X_t = h(X_{t-1}, W_t, \sigma\Sigma\epsilon_t, \Theta), \tag{23}$$

where h is continuous and differentiable, $\epsilon_t = [\epsilon^{z\prime}_t, \epsilon^{u\prime}_t]'$, $\Sigma = diag[\Sigma_z, \Sigma_u]$, $W_t = [Z'_t, U'_t]'$.

It is useful to highlight three features of our setup. First, (21) permits parameters be a constant, exogenously or endogenously drifting, or both, depending on whether the derivatives of Φ with respect to X_t and U_{t+1} are zero or not, and Θ_t will be serially correlated if X_t, U_{t+1}, or both are serially correlated. Second, the setup allows for time variations in the parameters regulating preferences, technologies, and constraints but does not consider variations in the auxiliary parameters regulating the law of motion of Z_t and U_{t+1} or the mapping Φ. Thus, we do not study time variations due to stochastic volatility, GARCH, or rare event phenomena (as in, e.g., Andreasen, 2012), nor those driven by evolving persistence of the exogenous processes. Finally, while we examine stationary environments, non stationarities can be dealt with, as usual, scaling the endogenous variables by the common growth process. [2]

[2] An interesting question not addressed in the paper is whether stationary solutions that look like (23) can be obtained when non stationarities are present in all parameters or only in a selected group of them.

For the rest of this section we focus on two issues. Under what conditions would an econometrician using a constant parameter model approximate well a time varying parameter model? Would a time varying parameter VAR model the be correct reduced form counterpart a time varying parameter DSGE?

3.1 First order approximations

Linearly expanding (19) around the steady states leads to

$$0 = E_t \left[F x_{t+1} + G x_t + H x_{t-1} + L z_{t+1} + M z_t + N \theta_{t+1} + O \theta_t \right], \tag{24}$$

where $F = \frac{\partial f}{\partial X_{t+1}}$, $G = \frac{\partial f}{\partial X_t}$, $H = \frac{\partial f}{\partial X_{t-1}}$, $L = \frac{\partial f}{\partial Z_{t+1}}$, $M = \frac{\partial f}{\partial Z_t}$, $N = \frac{\partial f}{\partial \Theta_{t+1}}$, and $O = \frac{\partial f}{\partial \Theta_t}$, all evaluated at the steady-state values of (X_t, Z_t, Θ_t), with lowercase letters indicating deviations from the steady states. Linearly expanding (23) leads to:

$$x_t = P x_{t-1} + Q z_t + R u_t, \tag{25}$$

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

Proposition 1 describes how to compute the matrices P, Q, R and Corollary 2 highlights the differences with the constant parameters case.

Proposition 1. *The matrices P, Q, R satisfy:*

- *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 $V = \psi_z' \otimes F + I_{n_z} \otimes (FP + G + N\phi_x)$.*
- *Given P, R solves $WR = -vec(N\phi_u\omega_u + O\phi_u)$, and $W = \omega_u' \otimes F + I_{n_\theta} \otimes (FP + G + N\phi_x)$,*

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 vec denotes the columnwise vectorization.

Corollary 2. *If $\phi_x = \phi_u = 0$, R=0, P solves $FP^2 + GP + H = 0$ and, given P, Q solves $VQ = -vec(L\psi_z + M)$, where $V = \psi_z' \otimes F + I_{n_z} \otimes (FP + G)$.*

(The proof of propositions and corollaries are in the online appendix A.)

The linear decision rules of a time varying parameter model differ from those of a constant parameter model because there will be an additional set of disturbances driving the endogenous variables, and because the responses of the endogenous variables to structural shocks may be altered. Thus, a constant parameter model, in general, incorrectly measures the structural dynamics and the relative importance of different sources of fluctuations in endogenous variables.

Corollary 3 gives conditions under which the structural dynamics of a time varying parameter model are reproduced with a constant parameter model.

Corollary 3. *If $\phi_x = 0$, or if $\phi_u = 0$ and $N\phi_x$ and $O\phi_x$ are zero, the dynamics in response to z_t shocks are identical to those obtained when parameters are constant.*

When parameter variations are purely exogenous, $\phi_x = 0$, the variability of x_t will be altered but the dynamics produced by structural disturbances will not. Thus, for example, the dynamics induced by technology shocks do not depend on whether the discount factor is constant or exogenously varying, provided technological and discount factor innovations are uncorrelated. Nevertheless, variance or historical decomposition exercises will be distorted, since the constant parameter model omits sources of variations (the ϵ_t^u disturbances). When parameter variations are purely endogenous, $\phi_u = 0$, there will be no extra source of variability in x_t. However, to know if a constant parameter model correctly characterizes the responses to structural shocks we need to check the entries of $N\phi_x$ and $O\phi_x$ matrices.

Although in (21) Θ_{t+1} depends on the endogenous variables X_t, endogenous time variations can also be obtained by making parameters a function of the exogenous variables, $\Theta_{t+1} = \Phi(\Theta, Z_t, U_{t+1})$ as, e.g., in Ireland (2007). The equations for P, Q, R are now different (P now solves $FP^2+GP+H = 0$; given P, Q solves $VQ = -vec(L\psi_z+ M + N\phi_z\psi_z + O\phi_z)$ and $V = I_{n_z} \otimes (FP + G + F\phi_z)$; and given P, R solves $WR = -vec(N\phi_u\omega_u + O\phi_u)$, where $W = I_{n_z} \otimes (FP + G + F\omega_u)$). However, the conclusions we derived hold with ϕ_z replacing ϕ_x in proposition 1 and corollaries 2 and 3.

Proposition 1 is derived under continuous and smooth parameter variations. An alternative would be to assume a Markov switching specification for the time variations (see e.g., Bianchi and Melosi, 2016). In our setup, parameters are treated as variables because it is computationally infeasible to solve the model taking into account all (infinite) future parameter paths. In Markov switching models, the number of future parameter paths is finite so that the model can be explicitly solved for these paths. This difference has pros and cons. In switching models, non linearities due to parameter variations are retained, while in smoothly varying parameters they appear only with higher order solutions. However, for computational reasons, the number of states must be kept small; when it is large, a smoothly varying specification provides a good approximation to a Markov switching specification. Also, while in smoothly varying parameter models the likelihood function is typically taken to be normal, in Markov switching specifications it is a mixture of normals.

Our preference for smoothly varying specifications comes from the empirical evidence, e.g., in Stock and Watson (1996), and with the practice employed in numerous time-varying parameter VARs. Note that the framework is flexible and can accommodate once-and-for-all breaks (at a known date), as long as the transition between states is smooth. For example, a smooth threshold exogenously switching specification can be approximated with $\theta_{t+1} = (1-\rho)\theta + \rho\theta_t + a\exp(t-T_0)/(b+exp(t-T_0))$, $t = 1, \ldots, T_0-1, T_0, T_0+1, \ldots T$, where a and b are vectors; and $\theta_{t+1} = (1-\rho)\theta + \rho\theta_t + a\exp(-(X_t - X))/(b + exp(-(X_t - X))$, where X is the steady-state value of X_t, can approximate smooth threshold endogenously switching specifications. Since models with occasionally binding constraints are special cases of switching models with endogenous probability of smooth transitions (see Binning and Maih, 2017), our analysis is applicable also to these situations. However, variations occurring at unknown dates, as in, e.g., Liu et al. (2011), or abrupt changes, as in, e.g., Davig and Leeper

(2006), are not covered by Proposition 1, since the smoothness condition is violated.

The (linear) solution (25) is a VAR(1) with fixed coefficients. Thus, linearized DSGE models with time-varying parameter do not generate new issues when it comes to time aggregation, invertibility, or non fundamentalness relative to a linearized fixed parameter DSGE models. In a linear framework, the P and Q matrices will be time-varying only if Φ is time-varying. Thus, it is incorrect to consider time-varying parameter VARs as the reduced form counterpart of continuously varying parameter linearized DSGE models: variations in DSGE parameters cannot produce the time-varying correlation structure VAR models generate unless the auxiliary parameters vary (see, e.g., Ascari et al., 2018) or there is learning (see, e.g., Cogley et al., 2015).

Kulish and Pagan (2017) developed solution and estimation procedures for models with abrupt breaks and learning between the states. Their solution for the pre-break and post-break period is a constant coefficients VAR, while for the learning period it is a time-varying coefficients VAR. Thus, a few words distinguishing Kulish and Pagan's approach from ours are needed. First, they are interested in characterizing the solution during the learning period when the structure is unchanged, while we are interested in the decision rule when parameters are continuously varying. Second, their modeling of time variations is abrupt, and the solution is designed to deal with that situation. Third, in our setup, expectations are varying with the structure; in Kulish and Pagan, they vary only in anticipation of a (foreseeable) break.

Finally, as (25) indicates, it is hard to distinguish linearized time-varying models from linearized time-invariant models featuring an additional set of shocks. In fact, models with n_1 structural shocks and n_2 time-varying parameters, models with $n = n_1 + n_2$ structural shocks and models with n_1 structural shocks and n_2 measurement errors are observationally equivalent:

$$x_t = Px_{t-1} + Qz_t + Ru_t, \tag{26}$$

$$x_t = Px_{t-1} + Q^* z_t^*, \tag{27}$$

$$x_t = Px_{t-1} + Qz_t + v_t, \tag{28}$$

where $Q^* = [Q, R]$; $z_t^* = [z_t', u_t']'$; and $v_t = Ru_t$. In applications, procedures like the one by Seoane (2016), can be used to select the interpretation of the additional shocks.

3.2 Higher order approximations

The online Appendix A shows that, in a time-varying parameter model, the second order approximate decision rule is

$$x_t = Px_{t-1} + Qz_t + Ru_t + \mathbf{C}\widetilde{\Lambda}_t + \mathbf{D}, \tag{29}$$

where, by construction, P, Q, and R are the same as in the first order solution,

$$\widetilde{\Lambda}_t = vec\left(\begin{bmatrix} x_t \\ x_{t-1} \\ z_t \\ u_t \end{bmatrix} \begin{bmatrix} x_t' & x_{t-1}' & z_t' & u_t' \end{bmatrix}\right),$$

and $(\mathbf{C}, \mathbf{D})$ are matrices. The second order approximate decision rule in a constant parameter model is

$$x_t = P^{cc} x_{t-1} + Q^{cc} z_t + \mathbf{C}^{cc} \widetilde{\Lambda}_t + \mathbf{D}^{cc}, \tag{30}$$

where (P^{cc}, Q^{cc}) are the same as in the first order solution. Note that $(\mathbf{C}^{cc})$ will have zero entries corresponding for all cross terms involving u_t (see appendix for details).

(29) and (30) will differ for four reasons: $P^{cc} \neq P, Q^{cc} \neq Q, R \neq 0$ as in the first order case; there will additional cross terms not present in the constant parameter solution and the quadratic terms in x_t will have different coefficients ($\mathbf{C}^{cc} \neq \mathbf{C}$); the adjustment due to risk will be different ($\mathbf{D}^{cc} \neq \mathbf{D}$). Thus, even when $P = P^{cc}$ and $Q = Q^{cc}$, the second order responses to structural shocks in the two models will differ. To illustrate this situation, consider:

$$\begin{aligned} E_t y_{t+1} &\equiv f(x_t, \theta_t) = \theta_t x_t^{0.95}, && (31)\\ x_t - \bar{x} &= 0.8(x_{t-1} - \bar{x}) + \epsilon_t^z, && (32)\\ \theta_t \equiv \Phi(x_{t-1}, \epsilon_t^u) &= 2 - 0.5[\exp(-0.9(x_{t-1} - \bar{x})) + \exp(0.9(x_{t-1} - \bar{x}))] + \epsilon_t^u, && (33) \end{aligned}$$

where both ε_t^z and ε_t^u are i.i.d. and $\bar{x} \equiv Ex_t = 1$. The second order solution is

$$\begin{aligned} y_t - \bar{y} &= 1.37(x_{t-1} - \overline{x}) + 1.71\epsilon_t^z + \epsilon_t^u - 0.11(x_{t-1} - \overline{x})^2 - 0.04\epsilon_t^{z2} + 0.07(x_{t-1} - \overline{x})\epsilon_t^z \\ &+ 0.76(x_t - \overline{x})\epsilon_t^u + 0.95\epsilon_t^z\epsilon_t^u, \end{aligned} \tag{34}$$

while the second order solution of the constant parameter version of the model is

$$y_t - \bar{y} = 1.37(x_{t-1} - \overline{x}) + 1.71\epsilon_t^z - 0.03(x_{t-1} - \overline{x})^2 - 0.04\epsilon_t^{z2} + 0.07(x_{t-1} - \overline{x})\epsilon_t^z \tag{35}$$

The linear responses to ϵ_t^z computed with (34) and (35) are the same, since $N\phi_x$ and $O\phi_x$ are zero. However, second order responses will differ since there is second order effect from x_{t-1} to θ ($\phi_{xx} = -0.08$) [3].

There is one case of interest when the constant solution will only produce mild distortions: when parameter variations are exogenous, the responses to structural shocks will be proportional and the proportionality factor depends on differences in the steady states due to risk terms.

Proposition 4. *If $\phi_x = 0$, then $\mathbf{C}^{cc} = \mathbf{C}$ and the dynamics of $(x_t - \mathbf{D})$ and $(x_t - \mathbf{D}^{cc})$ in response to z_t shocks in time-varying and constant parameter models are the same.*

For higher order solutions, the decision rules in constant and time-varying parameter models differ and the responses to structural shocks will be incorrectly characterized. This is because, for example, in a third order approximation, the optimality conditions feature terms requiring a correction of the linear terms to account for uncertainty. Since shocks are omitted in constant parameter models, the correction terms differ.

[3] In this example, the risk terms do not appear because (32)-(33) are linear

3.3 Time-varying decision rules?

There are typically two reasons for using a time-varying parameter structural model. The first is to improve the fit of a constant parameter model: by allowing additional sources of variations, not necessarily structurally interpretable, one hopes to absorb both unexplained variability and serial correlation. As we have seen, even linearized solutions can serve this purpose. The second reason is to allow changes over time in the data correlation and in the dynamics induced by the structural shocks. Linearized solutions are incapable of producing these time variations. Does the conclusion change when higher order solutions are considered? Equation (29) can be rewritten as

$$\begin{aligned} x_t \;&\simeq\; Px_{t-1} + Qz_t + Ru_t + C_{22}vec(x_{t-1}x_{t-1}^{'}) + C_{33}vec(z_t z_t^{'}) + C_{44}vec(u_t u_t^{'}) \\ &+C_{23}vec(x_{t-1}z_t^{'}) + C_{24}vec(x_{t-1}u_t^{'}) + C_{34}vec(z_t u_t^{'}). \end{aligned} \tag{36}$$

If u_t is treated as an (exogenous) variable, (36) is again a fixed coefficient representation. Thus, higher order solutions do not necessarily produce time-varying decision rules. However, if we interpret u_t as a "parameter," letting $\gamma_{1t} = P + C_{24}u_t$, and $\gamma_{2t} = Q + C_{34}u_t$ and neglecting a number of square terms, we have

$$x_t \approx A + \gamma_{1t}x_{t-1} + \gamma_{2t}z_t + Ru_t, \tag{37}$$

a time-varying decision rule. Thus, to hope to match the evidence produced by a time-varying VAR, one must consider at least second order solutions, neglect a number of terms, and treat u_t as a reduced form parameter vector for the purpose of inference. However, even under the "parameter" interpretation, structural responses will be time invariant. Expression (37) can generate time-varying responses to z_t shocks if and only if these shocks have effects on u_t, which is excluded a-priori since z_t and u_t innovations are uncorrelated. To be clear, with endogenous parameter variations, shocks to z_t may affect θ_t, and this will be reflected in the P and C_{24} matrices. However, since u_t is zero at all t when computing responses to elements of z_t, structural responses will be time invariant, regardless of the form of parameter variations.

4 Characterizing time-varying misspecification

Because an econometrician using the decision rules of a constant parameter model generally misspecifies the structural relationships when the DGP features parameter variations, it is important to detect time-varying misspecification problems. This section considers two diagnostics: one based on the optimality "wedges" and one based on the forecast errors of the constant parameter model. To see what the wedge diagnostic involves, consider the optimality conditions of a constant parameter model:

$$E_t\left[\mathcal{F}(X_{t-1}^{cc}, W_t, \sigma\Sigma\epsilon_t^z, \Theta)\right] = 0 \tag{38}$$

where for X_t we use the constant parameter decision rule: $X_t^{cc} = h^{cc}(X_{t-1}^{cc}, W_t, \sigma\Sigma\epsilon_t^z, \Theta)$. Here $\mathcal{F}$ is a martingale difference process. If instead X_t has been generated by a time-varying parameter model $X_t = h(X_{t-1}, W_t, \sigma\Sigma\epsilon_t, \Theta)$, $\mathcal{F}(X_{t-1}, W_t, \sigma\Sigma\epsilon_t^z, \Theta)$ will not be a martingale difference process since $\sigma\Sigma\epsilon_t^z \neq \sigma\Sigma\epsilon_t$ and $h \neq h^{cc}$. Moreover, it will be predictable using lags of X_{t-1}.

While these implications hold in general, we illustrate the argument using a first order approximate solution. The optimality wedge in this case is

$$\begin{aligned} (F(P-P^{cc})^2 + G(P-P^{cc}))x_{t-1} &+ \\ (F(Q-Q^{cc})\psi_z + G(Q-Q^{cc}) + F(P-P^{cc})(G-G^{cc}))z_t &+ \\ (F(P-P^{cc})R + GR + FR\omega_u)u_t. & \end{aligned} \tag{39}$$

When $P^{cc} = P, Q^{cc} = Q$, as in an exogenously varying model, the wedge reduces to $(GR+FR\omega_u)u_t$, which is non zero if $R \neq 0$ and predictable using $x_{t-j}, j \geq 1$ if $\omega_u \neq 0$. When $P^{cc} \neq P, Q^{cc} \neq Q$, as in an endogenously varying model, the wedge will be non zero, even when $R = 0$, and predictable using x_{t-j}, even when $\omega_u = 0$.

The wedge diagnostic shares with standard generalized method of moments (GMM) overidentification tests the idea of using a Lagrange-type test to detect deviations from the null, but it differs in two important respects: parameter estimates will be obtained with likelihood-based methods; and wedges can be non zero even without overidentification restrictions.

To detect time-varying misspecification, one can estimate a constant parameter model using approximate decision rules of different orders, compute optimality wedges in each case, and regress them on lags of the observables. If the regression coefficients obtained with different approximations are jointly significant, the martingale difference condition is violated, and there is evidence of time-varying parameters. Note that the diagnostic assumes that the model is correctly specified up to parameter variations. If this is not the case, lags of the observables may be significant, even without time-varying parameters (see, e.g., Inoue et al., 2015). Monte Carlo evidence on the properties of the wedge diagnostic when other forms of misspecifications are present is in table 1.

The logic of the forecast error diagnostic is similar. Because the argument is independent of the order of the approximation used, we present it for second order decision rules. The constant parameter solution is $x_t \simeq P^{cc}x_{t-1} + Q^{cc}z_t + C_{22}^{cc}vec(x_{t-1}x_{t-1}^{'}) + C_{33}^{cc}vec(z_t z_t^{'}) + C_{23}^{cc}vec(x_{t-1}z_t^{'})$, and the time-varying solution is $x_t \simeq Px_{t-1} + Qz_t + Ru_t + C_{22}vec(x_{t-1}x_{t-1}^{'}) + C_{33}vec(z_t z_t^{'}) + C_{44}vec(u_t u_t^{'}) + C_{23}vec(x_{t-1}z_t^{'}) + C_{24}vec(x_{t-1}u_t^{'}) + C_{34}vec(z_t u_t^{'})$. Let v_t^{cc} be the forecast error in predicting x_t using the constant parameter decision rule and the data generated from the time-varying parameter model. Then

$$\begin{aligned} v_t^{cc} &\simeq x_t - P^{cc}x_{t-1} - C_{22}^{cc}(x_{t-1}x_{t-1}^{'}) \\ &\simeq Qz_t + Ru_t - (P^{cc} - P)x_{t-1} - (C_{22}^{cc} - C_{22})vec(x_{t-1}x_{t-1}^{'}) + C_{33}vec(z_t z_t^{'}) \\ &+ C_{44}vec(u_t u_t^{'}) + C_{23}vec(x_{t-1}z_t^{'}) + C_{24}vec(x_{t-1}u_t^{'}) + C_{34}vec(z_t u_t^{'}). \end{aligned} \tag{40}$$

Thus, when $P^{cc} \neq P$ and $C_{22}^{cc} \neq C_{22}$, forecast errors are functions of lags of the observables x_{t-1} . When $P^{cc} = P$ and $C_{22}^{cc} = C_{22}$, forecast errors may still depend on the lags of the observables if u_t is serially correlated. Hence, an alternative way to check for parameter variations involves estimating the constant parameter model using approximate decision rules of different orders, in-sample predicting the endogenous variables, regressing the resulting forecast errors v_t^{cc} on lags of the observables, and checking the joint significance of the regression coefficients.

DGP	Estimated model	Optimality wedge		Forecast errors	
		T=500	T=150	T=500	T=150
Fixed parameter	Fixed parameter	0.00	0.00	0.00	0.00
Exogenously varying	Fixed parameter	0.58	0.28	1.00	0.78
Endogenously varying	Fixed parameter	0.99	0.25	1.00	0.99
Endogenously varying (internalization)	Fixed parameter	0.60	0.05	1.00	0.99
Fixed parameter second order solution	Fixed parameter first order solution	0.00	0.00	0.00	0.00
Fixed parameter time to build	Fixed parameter no time to build	0.12	0.08	0.49	0.11
Fixed parameter capacity utilization	Fixed parameter no capacity utilization	0.00	0.00	1.00	0.80
Fixed parameter stochastic volatility	Fixed parameter constant volatility	0.00	0.00	0.00	0.00
Fixed parameter second order, adjustment costs	Fixed parameter first order, no adjustment costs	0.00	0.00	0.00	0.00
Two states, one switch	Fixed parameter	1.00	0.20	0.99	0.97
Occasionally binding	Fixed parameter	0.59	0.41	1.00	0.99

Table 1: Percentage of rejections at the 0.05 confidence level of the null of no time variations in 200 experiments. The dependent variable is either the Euler wedge or the forecast error in the output equation. The regressors are lagged consumption and lagged real rate for the Euler wedge; lagged output, lagged consumption, and lagged hours for the forecast error.

We apply the two diagnostics to samples constructed using the RBC model of section 2. The parametrization is as in section 2.5. Table 1 reports the rejection rate of an F-statistic for the null hypothesis that all regression coefficients are zero at the 5 percent confidence level for two sample sizes (T=150, T=500) and a number of designs. The first four rows consider the models discussed in section 2: constant parameter, exogenously time-varying β and δ, endogenously time-varying β and δ as a function of the capital stock, with or without internalization. The next five rows consider situations where the estimating model neglects nonlinearities, high order terms, or structural features present in the data-generating process. The last two rows consider the case of a switching model with once-and-for-all switch in a number of parameters and an occasionally binding constraint on investment decisions (details on the models used are in the online appendix B).

The wedge diagnostic has good size properties (does not reject the null of constant parameters) when the estimating model is correctly specified and when there are a number of important forms of misspecification. However, it displays some distortion with the one-period time to built DGP. It has instead conservative power properties (does not reject the null very often) when the DGP features parameter variations.

This occurs primarily in the smaller sample, and it is due to the fact that parameter variations are small - with the chosen calibration they explain 3-6 percent of the output variance. If parameter variations are larger - they explain, say, 15-20 percent of output variance - the rejection rate is close to its nominal value (see table B.1 in the online appendix). The statistic has also good power in detecting once-and-for-all breaks or time variation due to occasionally binding constraints.

The forecast error diagnostic has similar size properties, except when the DGP features capacity utilization, but better power properties, even when time variations explain a small portion of output variance. Differences in power are due to the fact that the wedge diagnostic uses the nonlinear optimality conditions and thus needs either larger samples or larger parameter variations to detect time-varying misspecification.

4.1 Exogenous versus endogenous parameter variations

If the diagnostics indicate the presence of parameter variations, one may want to know whether they are of the exogenous or endogenous type. To distinguish the two options, we use the logic of the DGSE-VAR methodology of Del Negro and Schorfheide (2004).

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 a 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 for the VAR and $h_1(y|\gamma_1, M_1) > h_2(y|\gamma_2, M_2)$, there is better support in the data for g_1.

	Data Generating Process					
	Prior, T_1=150			Prior, T_1=450		
	Model B	Model C	Model D	Model B	Model C	Model D
Data added from B	1.00	0.00	0.00	1.00	0.00	0.00
Data added from C	0.01	0.97	0.01	0.00	0.99	0.00
Data added from D	0.00	0.00	1.00	0.00	0.00	1.00

Table 2: Fraction of 100 experiments when the Bayes factor exceeds 3.0. Marginal likelihoods are obtained using $T = 150$ data points produced by the models listed in the first column and T_1 data from the model listed in the rows. When rows do not sum up to one, the Bayes factor is inconclusive (below 3.0).

We use a similar idea to examine whether a model with exogenous or endogenous variations is better suited to explain the data. Given a sample of data, one uses a model with either exogenous or endogenous variations as a prior and generates data form it. A statistically larger marginal likelihood, say, when the exogenously varying parameter model is used as prior, would indicate that the available sample is more likely to be generated by a model with this feature. We prefer to use the DSGE-VAR device, rather than comparing the marginal likelihood of different models directly to avoid small sample distortions. Since we can add as many data points as we like from the prior, small sample distortions are less of an issue in the DSGE-VAR setup.

Table 2 reports Monte Carlo results for the RBC example. The parameterization is as in section 2, and a first order approximate solution is used to generate the data. We use proper priors on all parameters which are estimated using consumption, output, capital, and hours as observables and three measurement errors. Bayes factors are computed when the sample has $T = 150$ observations and $T_1 = 150$ or $T_1 = 450$ observations from the models listed in the rows are added to the sample. The statistic is powerful: marginal likelihood differences are quite large, even when $T_1 = 150$.

4.2 Some practical suggestions

To diagnose and analyze time variations in structural parameters of a given model we suggest researchers to use the following steps

i) Estimate the constant parameter version of the model, possibly allowing for time variations in the variance of the shocks.

ii) Run the diagnostics using the estimated model and, if time variations are pervasive, check whether endogenous versus exogenous variations are more likely. Because the two diagnostics have different small sample properties, it seems wise to use both of them in empirical applications.

When the model is of large scale, running regressions on lags of all endogenous variables is likely to lead to overparameterization and multicollinearity. To make the test powerful, users should employ the states of the model in the regressions. Similarly, when performing exogenous versus endogenous checks, having the proper state variables for the endogenous specification is important to make the comparison fair. To avoid overparameterization, it is a good idea to a-priori shrink the coefficients of the auxiliary regressions toward zero. Rejection of the no time variations null will give researchers stronger confidence that parameter variations are indeed present.

iii) When time variations are detected, one needs to choose which parameters are time-varying. One option is to specify time variations in all structural parameters and design a prior that allows the variance of some elements of U_t to be zero (see, e.g., Ferroni et al., 2017). Alternatively, one could introduce time variations only in parameters known to be unstable or suspected to be time-varying.

iv) Estimate the time-varying parameter model and compare its structural dynamics with those of the constant parameter model. One can proceed in two ways: enlarge the number of observable variables whenever a new shock to the parameters is introduced; or maintain the same number of observables, even if a new shock is introduced. We follow these steps in the empirical application of section 6.

5 Inference

Given the results of section 3, important inferential distortions may occur using a constant parameter specification when the DGP features time varying parameters. In this paper, as in the majority of the literature, estimation and inference are based on the likelihood function of the linearized constant parameter model.

The likelihood function is typically constructed via the Kalman filter and, thus, it is a function of the forecast errors computed with the constant parameter model. Thus, one should expect the forecast error misspecification described in section 4 to spread to the likelihood function. Two problems are relevant: the matrix P entering the prediction formula is generally incorrect; a smaller number of shocks is assumed to drive the endogenous variables. If structural and parameter disturbances are AR(1), shock aggregation produces a lower dimensional ARMA(2,1) process. Because the disturbances of a constant parameter model are assumed to be AR(1), distortions are likely to occur in the serial correlation properties of the estimated shocks and the parameters regulating the internal propagation of the model.

We have investigated how these two problems affect the identification of the constant parameters and the conclusions concerning the propagation and the relative importance of structural shocks in the context of the RBC model of section 2. The online appendix C presents the results of a Monte carlo exercise. Two main conclusions emerge. First, the identification issues highlighted by Canova and Sala (2009) are generally important when time varying misspecification is present. When a constant parameter model is incorrectly assumed, the maximum of the likelihood function changes location, its shape is twisted, and its curvature flattened. It turns out that both shock misaggregation and decision rule misspecification induce identification pathologies and that the distortions created by the latter are considerably worse.

Likelihood estimation of a constant parameter model shows an interesting pattern of biases. Because the decision rules are misspecified and shocks misaggregated, the parameters most distorted are those regulating the estimated persistence of the structural shocks and those controlling income and substitution effects. Thus, one is more likely to find that very persistent processes are required for the time-invariant model to fit the data when the DGP features time-varying parameters and that parameters such as the labor share or the intertemporal elasticity of substitution tend to be biased.

6 Time-varying financial frictions?

We apply the technology we developed to study time variations in the parameters of Gertler and Karadi's (2010) - GK for short - model. Our contribution is threefold. First, we provide likelihood estimates of the model-specific parameters (the fraction of capital diverted by banks λ, the proportional transfer to entering bankers ω, and the survival probability of bankers θ), which the authors have calibrated to match the steady-state spread, the steady-state leverage, and a notional length of bank activity. Second, we use the diagnostics of section 3 to gauge the extent and the sources of parameter variations. Third, we estimate models where λ is time-varying and examine what it implies for the responses of capital quality shocks. We use U.S. data on the growth rates of output, of consumption, of leverage, of intermediary demand for assets (credit) and a spread measure - the difference between the yield of a BAA 10-year corporate bond and the 10-year Treasury bond of constant maturity - in estimation. Yield, real personal consumption expenditures, and GDP are from the Federal Reserve Economic Data (FRED). Leverage and credit are from Haver Analytics. The former measures Tier 1 (core) capital as a percent of average total assets; the latter, total loans scaled by the size of the U.S. population. We consider two samples: 1985:2- 2014:3 and 1985:3-2007:4. The linearized equations and the prior used are in appendix D.

The posterior modal estimates for the full sample are $\lambda = 0.170$, $\theta = 0.452$, and $\omega = 0.012$; the posterior standard deviations are small (0.007, 0.008, and 0.0005). For the shorter sample, posterior modal estimates are $\lambda = 0.138$, $\theta = 0.399$, and $\omega = 0.01$, and the posterior standard deviations are 0.007, 0.011, and 0.0006. For comparison, GK used $\lambda = 0.318$, $\theta = 0.972$, $\omega = 0.002$. Note that λ regulates private leverage: the value used by GK implies a steady-state leverage of 1.38. Our full sample estimate implies a steady-state leverage of 3.39, closer to the value in U.S. corporate and non corporate business sectors.

We run our diagnostics to check for parameter variations (see table 3). With the full sample, the forecast errors of all equations but output are predictable, and typically, lagged output and the lagged spread are significant. In addition, lagged consumption and lagged investment to output ratios significantly explain movements of the Euler wedge (coefficients are, respectively, -0.10 and 0.72, with t-statistics of -7.96 and 5.37). Because the conclusions for the shorter sample are similar, the time variations we detect are not due to the financial crisis.

	C_{t-1}	Y_{t-1}	Credit$_{t-1}$	Leverage$_{t-1}$	Spread$_{t-1}$	
	t-statistic					F-statistic
Forecast error	Sample 1985:3-2014:3					
C	-0.10	3.11	-0.01	0.58	6.13	3.64
Y	-1.33	1.37	0.38	-0.60	0.68	1.24
Credit	-0.22	3.38	0.87	-0.36	6.02	5.74
Leverage	0.14	-3.27	-0.91	0.40	-5.81	5.51
Spread	-0.10	-4.07	-0.15	0.48	-5.76	7.25
	Sample 1985:3-2007:4					
C	-0.19	3.11	-0.01	-0.58	6.13	3.64
Y	-1.33	1.37	0.38	0.60	0.68	1.24
Credit	-0.22	3.38	0.98	-0.36	6.02	5.74
Leverage	0.14	-3.27	-0.91	0.40	-5.81	5.51
Spread	-0.10	-4.07	0.15	0.48	-5.76	7.25

Table 3: Forecast error diagnostic. The left-hand side regression variable is the forecast error in the equation listed in the first column; the right-hand side variables are listed in columns 2 through 5. Critical values for the null of zero coefficients are F(5,112)=2.56 (full sample) and F(5,85)=2.90 (shorter sample).

While all parameters could a-priori be varying, we choose to study time variation in only λ, because it regulates leverage and drives movements in the credit and spread equations, whose forecast errors are highly predictable. We specify:

$$\lambda_t = (1-\rho_\lambda)\lambda + \rho_\lambda \lambda_{t-1} + \sigma_\lambda e_{t,\lambda} \quad \text{(Exogenous variations)} \tag{41}$$

$$\lambda_t = (2\lambda_u - (\lambda_u - \frac{\lambda}{2}) * (\exp(-\phi_1(X_{t-1} - X^s)) + \exp(\phi_2(X_{t-1} - X^s))) + \sigma_\lambda e_{t,\lambda} \quad \text{(Endogenous variations),} \tag{42}$$

where X is net bank wealth and X^s its steady-state value. We select bank net wealth as the relevant state variable in (42) because of its importance in determining the magnitude of the spread and the dynamics of credit. Depending on the values of $\lambda_u, \phi_1, \phi_2$, (42) may generate variations that affect the steady states, the dynamic responses to shocks, or both, and the variations in dynamic responses could be symmetric (if $\phi_1 = \phi_2$) or asymmetric (if $\phi_1 \neq \phi_2$). We add to the original model either equation (41) or (42) and assume that agents know them when optimizing. When considering endogenous time variation, we focus on the case where agents do not internalize the effects of their choices on the parameters, which is arguably more relevant for this application. We calculate the marginal likelihood for each DSGE-VAR style specification, averaging the marginal likelihood obtained on a grid when $\rho_\lambda = [0.7, 0.9]$ and $\sigma_\lambda = [0.01, 0.05]$ for the exogenously varying model and when $\sigma_\lambda = [0.01, 0.05], \phi_1 = [0.01, 0.1]$, and $\phi_2 = [0.05, 0.2]$ for the endogenous specification (we keep $\lambda_u = 0.8$ fixed), setting $T_1 = T = 115$ (our full sample size). The endogenously varying specification is strongly preferred: the average difference in the log marginal likelihood of the two specifications exceeds 300 and nowhere in the chosen intervals is the difference smaller than 100.

Parameter	Constant Parameters		Exogenous time variations		Endogenous time variations	
h	0.431	(0.006)	0.235	(0.029)	0.213	(0.003)
λ	0.170	(0.007)	0.628	(0.028)	0.331	(0.045)
ω	0.012	(0.0005)	0.014	(0.0009)	0.050	(0.0008)
θ	0.452	(0.008)	0.504	(0.010)	0.445	(0.036)
ρ_ξ	0.672	(0.020)	0.990	(0.003)	0.726	(0.066)
σ_{xi}	0.159	(0.012)	0.155	(0.010)	0.180	(0.028)
ρ_λ			0.998	(0.002)	0.599	(0.016)
σ_λ			0.028	(0.002)	0.027	(0.004)
λ_u					0.859	(0.036)
ϕ_1					0.022	(0.008)
ϕ_2					0.182	(0.029)
Log ML	-608.36		1550.52		1574.94	

Table 4: Parameter estimates, Gertler and Karadi model, sample 1985:3-2014:3. Posterior standard deviations in parenthesis. Log ML lists the log marginal likelihood of each specification.

We estimate the parameters allowing λ to be time-varying. The priors for the new parameters are in appendix D. Table 4 reports estimates of selected parameters.

In the model with exogenously varying parameters, variations in λ_t are very persistent. Compared with the constant parameter model, estimates of (λ, θ) are larger, the lifetime of bankers slightly increase, and the persistence of the capital quality shock increases. With the endogenous specification, the estimate of λ is intermediate between the other two estimates but bankers' survival probability is roughly unchanged. The data requires a strong asymmetric specification for time variations ($\phi_1 < \phi_2$), implying a negative relationship between the fraction of funds that bankers can divert and their net worth. Finally, confirming our DSGE-VAR analysis, the endogenous specification has larger marginal likelihood than the alternative specifications.

Figure 2 plots estimates of the responses of output, inflation, investment, net worth, leverage, and the spread to a 1 percent capital quality shock. The constant parameter specification closely replicates the dynamics in GK's figure 3. There is a decline in output, investment temporarily falls and then increases because capital is below its steady state. Bankers' net worth falls, and the spread sharply increases.

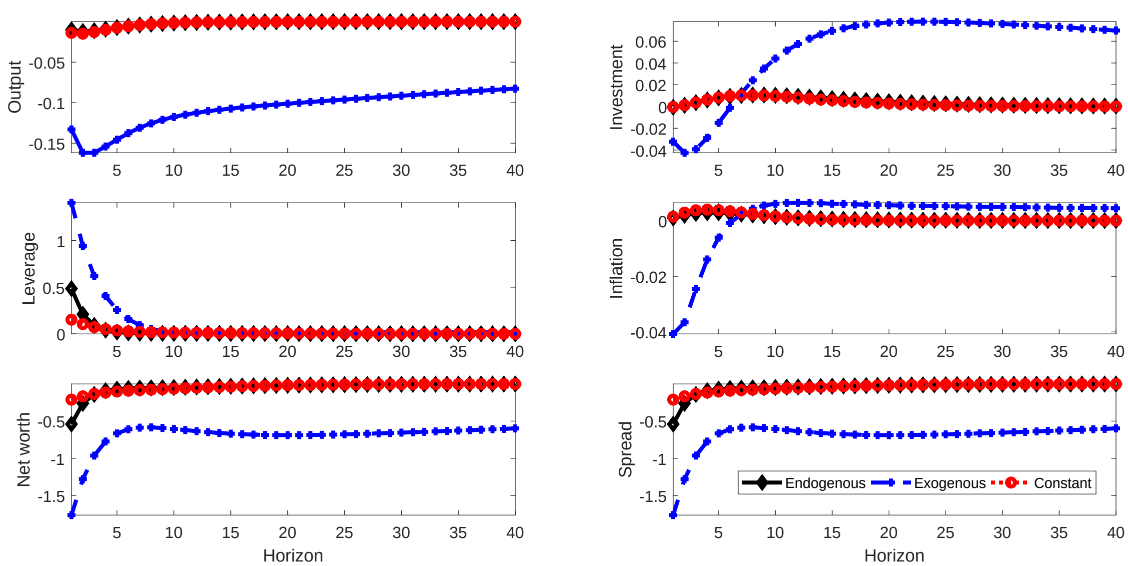

Figure 2: Dynamics in response to capital quality shock.

With an exogenously varying λ, the responses are more persistent, primarily because the persistence of the capital quality shock is abnormally high. Quantitatively, output, net worth, and the spread fall more in the short run, primarily because the steady-state value of λ increases. Thus, making λ exogenously time-varying enhances the model's ability to capture the impact recessionary effects of capital quality shocks.

With endogenously varying λ, the medium-term dynamics reproduce those of the constant parameter model, but in the short run, the effect on leverage, net worth and the spread is stronger. This is because shocks that lower net worth make the share of funds diverted by banks higher, and this produces stronger dynamics in the financial block. Since the dynamic responses of net worth are larger, the spread increases more.

Nevertheless, differences in the responses of investments and output relative to the time invariant model are small. Interestingly, estimates of the habit parameter h dramatically fall in the time-varying specifications. Thus, in the estimated time invariant model, h captures, in part, the missing dynamics due to the time variations of λ.

We take this evidence as suggestive of the potential problems one encounters estimating time invariant structural models and indicative that the relationship between λ and net worth needs a proper micro fundation for policy counterfactuals to be interpreted. Providing this micro fundation is beyond the scope of the paper, but attempts to endogenize crucial parameters in models like Gertler and Karadi's exist (see, e.g., Bigio, 2012; Ferrante, 2018). In these models, an endogenous deterioration of the quality of loans generally leads to higher aggregate leverage, higher aggregate risks, and a larger probability of bank runs.

7 Conclusions

This paper is concerned with the misspecification induced by neglected parameter variation and with the consequences of assuming incorrect forms of time variation and provides researchers with a new set of tools to assess the quality of their models. We characterize the approximate decision rules of a DSGE model when parameter variation is present; discuss whether constant parameter models provide a good approximation to the DGP; and examine whether time-varying parameter DSGE models generate decision rules comparable to those of time-varying parameter VARs. We provide diagnostics to detect time-varying misspecification and study the consequences of using time-invariant models when the DGP features parameter variations.

When parameter variations are present, a constant parameter model does not "reasonably" approximate the DGP. When linearized solutions are considered, there are two special cases when the dynamics in response to structural shocks will be the same. When second order solutions are considered, structural responses in time-varying and constant parameter models are proportional only when parameter variations are exogenous. For higher order solutions, the structural responses will be highly distorted. Constant and time-varying parameter models produce dynamics that are different because income and substitution effects are altered. Disturbances to the parameters add to the uncertainty of the environment, making agents prefer to consume more today relative to the future for the same transitory fluctuations in income.

We show that linear approximations do not produce time-varying decision rules and higher order approximations can do this only if parameter disturbances are interpreted as (reduced form) decision rule coefficients. Still, regardless of order of approximation employed, structural responses will be time invariant.

The diagnostics we design are able to detect neglected parameter variations and distinguish exogenous and endogenous forms of time variations. We highlight that certain identification problems noted in the literature may be the result of neglected time variations. Our Monte Carlo study indicates that parameter and impulse response distortions may be large, even for modest parameter variations.

We show that the parameter regulating moral hazard in the Gertler and Karadi (2010) model is likely to be time-varying. When we allow variations to be a function of net worth, the fit of the model dramatically improves because there is an additional propagation channel that makes spread and output responses to capital quality shocks stronger and more persistent.

8 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., Bonomolo P. and H. F. Lopes (2018). Walk on the wild side: temporarily unstable paths and multiplicative sunspots, forthcoming, American Economic Review.

Bianchi, F., and L. Melosi (2016). Modeling the evolution of expectations and uncertainty in general equilibrium. International Economic Review, 57, 717-756.

Bigio, S. (2012). Financial risk capacity. Columbia University manuscript.

Binning, A. and J. Maih (2017). Modeling occasionally binding constraints using regime switching. Norges bank working paper, 23-1017.

Brinca, P., H. Holter, P. Krusell, and L. Malafry (2016). Fiscal multipliers in the 21st century. Journal of Monetary Economics, 77, 53-69.

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.

Castelnuovo, E. and G. Pellegrino (2018). Uncertainty-dependent effects of monetary policy shocks: A new-Keynesian interpretation. Journal of Economic Dynamics and Control, 93, 277-293.

Chang, Y., S. Kim 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., P. Kehoe and E. Mc Grattan (2007). Business cycle accounting. Econometrica, 75, 781-836.

Cogley, T., C. Matthes and A. Sbordone (2015). Optimized Taylor rules for disinflation when agents are learning. Journal of Monetary Economics, 72, 131-147

Cogley, T., and T. Sargent (2005). Drifts and volatilities. Monetary policy and outcomes in post WWII US. Review of Economic Dynamics, 8, 262-302.

Cogley, T., and T. Yagihashi, (2010). Are DSGE approximating models invariant to policy shifts? The B.E. Journal: Macroeconomics Contributions, vol 10, 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.

Dew-Becker, I. (2014) Bond pricing with a time-varying price of risk in an estimated medium-scale DSGE model. Journal of Money Credit and Banking, 46, 837-888.

Dueker, M. A. Fischer, and R. Dittman, (2007). Stochastic Capital depreciation and the co-movements of hours and productivity. Berkeley Journals: Topics in Macroeconomics, vol 6, article 6.

Ferroni, F., S. Grassi, and M. Leon Ledesma (2017). Selecting Structural Innovations in DSGE models. Forthcoming, Journal of Applied Econometrics.

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

Fernandez-Villaverde, J. , P. Quintana, and J. Rubio-Ramirez (2011). Risk matters: the real effects of volatility shocks. American Economic Review, 101, 2530-2563.

Ferrante, F. (2018). A model of endogenous loan quality and the collapse of the shadow banking system. American Economic Journal: Macroeconomics, 10, 152-202.

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

Gourio, F., (2012). Disaster risk and business cycles. American Economic Review, 102, 2734-2766.

Hansen, L., and T. Sargent (2010). Wanting robustness in macroeconomics, in B. Friedman and M. Woodford (eds.) Handbook of Monetary Economics, 3, 1097-1157.

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.

Inoue, A., C.H. Kuo, and B. Rossi, (2015) Identifying sources of model misspecification, UPF manuscript.

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

Justiniano, A., and G. Primiceri (2008). The time-varying volatility of macroeconomic fluctuations. American Economic Review, 98, 604-641.

Kulish, M., and A. Pagan (2017). Estimation and solution of models with expectation and structural changes. Journal of Applied Econometrics, 32, 255-274.

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

Meier S. , and C. Sprengler (2015). Temporal stability of time preferences. Review of Economics and Statistics, 97(2), 273-286.

Parkin, M. (1988) A method for determining whether the parameters in aggregative models are structural. Carnegie Rochester Conference Series in Public Policy, 29, 215-252.

Primiceri, G. (2005) Time-varying structural vector autoregression and monetary policy. Review of Economic Studies, 72, 821-852.

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

Seoane, H. (2016). Parameter drifts, misspecification and the real exchange rate in emerging countries. Journal of International Economics, 98, 204-215.

Sims, C., T. Zha (2006). Were there regime switches in the US monetary policy? American Economic Review, 96(1), 54-81.

Stock, J., and M. Watson (1996). Evidence of structural instability in macroeconomic relationships. Journal of Business and Economic Statistics,14, 11-30.

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