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 2015

Introduction

- Typical to assume that DSGE are structural, i.e. preference, technology parameters are invariant to policy interventions (Hurwicz, 1962).

- Mounting evidence (Dueker et al., 2007, Fernandez and Rubio, 2007, Canova, 2009, Rios and Santaeularia, 2010, Liu et al., 2011, Galvao, et al., 2014, Vavra, 2014, Seoane, 2014, Meier and Sprengler, forthcoming) that DSGE parameters are not time invariant.

- Parameter variations may be due to misspecification of a time invariant structure (Cogley and Yagihashi, 2010; Chang, et al., 2013, Basile and Carvalho, 2015).

- Parameter variations may be needed to insure the existence of stationary equilibrium (see e.g. Schmitt Grohe and Uribe, 2003).

• Standard approach to modelling parameter variations follows VAR literature: parameters are exogenously drifting as a RW with small variance.

• Justification comes from Stock and Watson (1996): economic relationship display small but persistent variations.

• Many questions suggest that time variations may be endogenous:

- Is it reasonable to assume that the Federal Reserve reacts with the same coefficients to inflation in an expansion or in a contraction? (Davig and Leeper (2006): state dependent rule).

- Are households as risk averse when they are wealthy or poor? Or as impatient when the capital stock is high or low?

- Does the propagation of shocks depends on the state of private and government finances? Or on inequality? (e.g. Brinca et al, 2014).

Goals of the paper

- Characterize differences in decisions rules, impulse responses, etc. when parameters variations are exogenous or state dependent. In the latter case, examine externalized or internalized variations.

- Provide diagnostics to detect misspecification driven by parameter variations and to distinguish exogenous from state dependent variations.

- Study the consequences of considering time invariant models when parameters are time varying in terms of identification, estimation, inference. Compare Likelihood and SVAR based estimates of structural dynamics.

- Apply the technology to Gertler-Karadi (2010) model.

Related literature

- Seoane (2014): Use patterns of time variations as model respecification tool.

- Kulish and Pagan (2012): Likelihood estimation in models with structural breaks; Galvao et al. (2014): time varying structural estimation.

- Magnusson and Mavroedis (2014, Econometrica): time variations helps identification of time invariant parameters in GMM. Does it apply to likelihood based estimation?

- Huang (2014): Moderate (exogenous) variations in weakly identified DSGE parameters do not make ML asymptotics and standard break test wrong. Moderate variations in strongly identified parameters do.

- Ireland (2007, JMCB): trend inflation reflects exogenous structural shocks. Ascari and Sbordone (2014, JEL): trend inflation reflects monetary choices.

Results

- If parameter variations are exogenous, structural dynamics are the same as in a time invariant model. Little inferential loss if structural disturbances are correctly identified (variance/ historical decompositions distorted).

- If parameter variations are endogenous, structural dynamics may be altered. Extent of the differences depends on the specification.

- Identification and inferential problems may arise when parameter variations are neglected.

- SVAR methods competitive with likelihood based methods.

- Financial (moral hazard) friction in Gertler-Karadi model is time varying and endogenous.

Plan of the talk

1. Some analytical results and an example.

2. Diagnostics for time varying misspecification; and for exogenous vs. endogenous time variations.

3. Identification and misspecification.

4. Some likelihood based and semi-structural (SVAR) based MC evidence.

5. An application.

The setup

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

X_t: $n_x \times 1$ vector of the endogenous variables; Z_t: $n_z \times 1$ vector of the exogenous variables; Θ_{1t}: $n_{\theta 1} \times 1$ vector of possibly time varying (TV) structural parameters; f is continuous and differentiable up to order q.

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

Ψ is continuous and differentiable up to order q; $\epsilon^z_{t+1} \sim iid(0, I)$ a $n_e \times 1$ vector, $n_z \geq n_e$; $\sigma \geq 0$ an auxiliary scalar; Σ_ϵ a known $n_e \times n_e$ matrix.

Evolution of $\Theta_t = [\Theta_{1t}, \Theta_{2t}]$, where Θ_{2t} : $n_{\theta 1} n_x \times 1$ vector:

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

Φ is continuous and differentiable up to order q, Θ is a vector of constants, U_t : $n_u \times 1$ vector of disturbances, $n_\theta = n_{\theta 1}(1 + n_x) \geq n_u$.

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

Ω is continuous and differentiable up to order q; $\epsilon^u_t \sim iid(0, I)$ is $n_u \times 1$ vector, uncorrelated with ϵ^z_{t+1}, Σ_u is a known $n_u \times n_u$ matrix.

Let $\epsilon_{t+1} = [\epsilon^z_{t+1}, \epsilon^u_{t+1}]', \Sigma = diag[\Sigma_z, \Sigma_u]$. Decision rule:

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

• Time variations only affect structural parameters (see e.g. Andreasan, 2012 for TV in auxiliary parameters Σ_ϵ).

First order approximations

The linear approximation of (1) is

$$0 = E_t \left[F x_{t+1} + G x_t + H x_{t-1} + L z_{t+1} + M z_t + N \theta_{1t+1} + O \theta_{1t} \right] \quad (6)$$

where $F = \partial f / \partial X_{t+1}$, $G = \partial f / \partial X_t$, $H = \partial f / \partial X_{t-1}$, $L = \partial f / \partial Z_{t+1}$, $M = \partial f / \partial Z_t$, $N = \partial f / \partial \Theta_{1t+1}$, $O = \partial f / \partial \Theta_{1t}$.

The linear approximation to the decision rule:

$$x_t = P x_{t-1} + Q z_t + R u_t \quad (7)$$

where $P = \partial h / \partial X_{t-1}$, $Q = \partial h / \partial Z_t, R = \partial h / \partial U_t$.

Proposition 1 *The matrices P, Q, R are obtained as follows:*

- P *solves* $FP^2 + (G + N\phi_x)P + (H + O\phi_x) = 0$

- *Given* P*,* $VQ = -vec(L\psi_z + M)$ *and* $V = \psi_z' \otimes F + I_{n_z} \otimes (FP + G + N\phi_x)$ *where* vec *denotes the columnwise vectorization*

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

where $\phi_u = \partial\Phi/\partial U_{t+1}$, $\phi_x = \partial\Phi/\partial X_t$, $\psi_z = \partial\Psi/\partial Z_t, \omega_u = \partial\Omega/\partial U_t$,where ω_u is a $n_u \times n_u$ matrix and ψ_z is a $n_z \times n_z$ matrix, both with all eigenvalues strictly less than one in absolute value.

Corollary 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 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 1: have time varying parameters equivalent to having additional shocks to the model.

- Corollary 2: with exogenous parameters variations, contemporaneous and lagged dynamics to structural shocks are the same as in a time invariant model. Can we recover structural shocks with a time invariant model?

- Corollary 3: with endogenous parameter variations, contemporaneous and lagged dynamics to structural shocks may be affected. Need to know whether $N\phi_x$ or $O\phi_x$ are zero or not.

Higher order approximations

- Do the conclusions change?

Let $W_t = [Z_t, U_t]'$, $Y_{t+1} = [X_{t+1}, X_t]'$, Using the decision rule is $X_t = h(X_{t-1}, W_t, \sigma\boldsymbol{\Sigma}\epsilon_{t+1})$, (1) is

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

- The second order approximation of (8) is

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

The second order expansion of the decision rule 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 4 *Time variations in the parameters affect* h_{xx}, h_{xw}, h_{ww} *if only if* h_x *and* h_w *are affected*

- In higher order approximations there are terms of the form $F_{x\sigma\sigma} \neq 0, F_{w\sigma\sigma} \neq 0$. These require a correction of the linear terms in the decision rule to account for uncertainty.

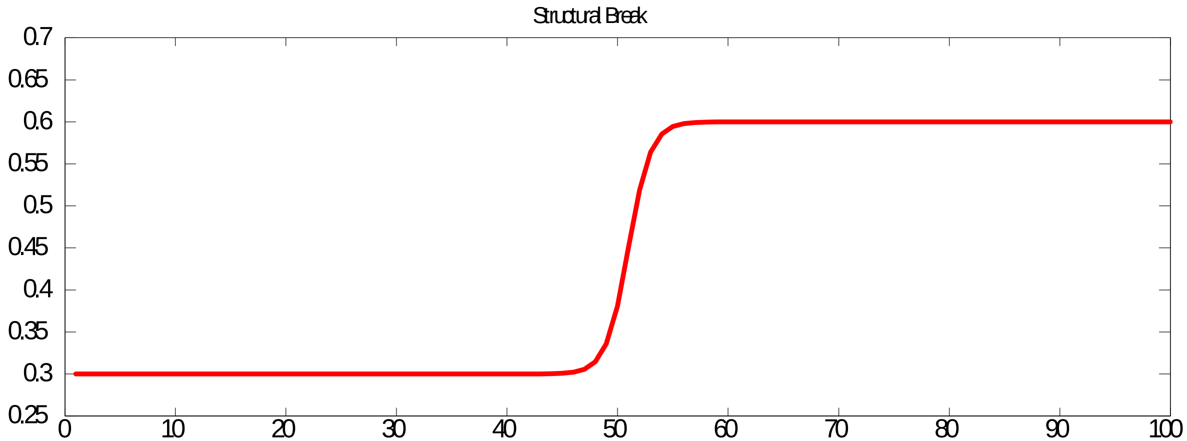
- The dynamics induced by structural shocks in fixed coefficient and TV coefficient models will be different (some shocks are omitted in fixed coefficient models).

Discussion

- Time variations are assumed to be continuous.

- Can accommodate once-and-for-all breaks (at known dates) with smooth transition (e.g. $\theta_{t+1} = (1-\rho)\theta + \rho\theta_t + a * \exp(t)/(b + exp(t))$, $t = -T_1, \ldots, -1, 0, 1, \ldots T_2$, exogenous; $\theta_{t+1} = (1-\rho)\theta + \rho\theta_t + a \exp(-(K_t - K + U_{\theta,t+1}))/(b + exp(-(K_t - K + U_{\theta,t+1}))$, endogenous).

- Can not allow for Markov switching variations in the parameters or non-smooth transition e.g. $\theta_{t+1} = (1 - I(K_{t-1} > K))\theta_0 + I((K_{t-1} > K))\theta_1$ (see Davig and Leeper (2006)).



- Kulish and Pagan (2014) have a solution method with abrupt structural breaks and learning - valid for first order approximations only. We use final form to derive results (no distinction between states and controls).

- Are time varying parameters CKMcG wedges? No: across equation restrictions; less time varying parameters than optimality conditions.

A RBC example

$$\max E_0 \sum_{t=1}^{\infty} \beta_t \left(\frac{C_t^{1-\eta}}{1-\eta} - A\frac{N_t^{1+\gamma}}{1+\gamma}\right) \tag{11}$$

$$Y_t(1-g_t) = C_t + K_t - (1-\delta_t)K_{t-1} \tag{12}$$

$$Y_t = \zeta_t K_{t-1}^{\alpha} N_t^{1-\alpha} \tag{13}$$

Y_t is output, C_t consumption, K_t the stock of capital, N_t hours worked and $g_t = \frac{G_t}{Y_t}$ the share of government expenditure.

Exogenous disturbances:

$$\ln \zeta_t = (1-\rho_{\zeta)}) \ln \bar{\zeta} + \rho_\zeta \ln \zeta_{t-1} + e_t^z \tag{14}$$

$$\ln g_t = (1-\rho_g) \ln \bar{g} + \rho_g \ln g_{t-1} + e_t^g \tag{15}$$

- 12 parameters: structural (α, η, γ, A, β_t δ_t), auxiliary ($\bar{\zeta}, \bar{g}, \rho_\zeta, \rho_g, \sigma_\zeta, \sigma_g$).

- β_t and δ_t allowed to be time varying (Meier and Sprengler, forthcoming; Dueker, et al., 2007, Liu et al., 2011). Note: Ireland (2004): β and δ weakly identified. Canova and Sala (2009): partially identified.

The optimality conditions:

$$
\begin{aligned}
AC_t^{\eta} N_t^{\gamma} &= (1-\alpha)(1-g_t)Y_t/N_t && (16)\\
C_t^{-\eta} &= E_t\left[\frac{\beta_{t+1}}{\beta_t} C_{t+1}^{-\eta}\left(\frac{\alpha(1-g_{t+1})Y_{t+1}}{K_{t+1}} + 1 - \delta_{t+1}\right)\right] \\
&+ E_t\left[\frac{\partial \beta_{t+1}}{\partial K_t} u(C_{t+1}, N_{t+1}) - \frac{\partial \delta_{t+1}}{\partial K_t} K_t\right] && (17)\\
(1-g_t)Y_t &= C_t + K_t - (1-\delta_t)K_{t-1} && (18)\\
Y_t &= \zeta_t K_{t-1}^{\alpha} N_t^{1-\alpha} && (19)
\end{aligned}
$$

Two effects of parameter variations on optimality conditions:

- direct effect in the Euler equation and in the resource constraint when β_t and δ_t are time varying.

- If agents take into account that their decisions affects parameter variations, there is an additional (indirect) effect due to the derivatives of β_{t+1} and δ_{t+1} with respect to the endogenous states (the capital stock).

Model A: Constant coefficients.

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

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

In the steady state:

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

Model B: Exogenous parameter variation

Let $d_t = \beta_{t+1}/\beta_t$;
$\Theta_{1t+1} - \Theta \equiv (d_{t+1} - (1-\rho_\beta)\beta, \delta_{t+1} - (1-\rho_\delta)\delta)' = U_{t+1}$

$$u_{\beta,t+1} = \rho_d u_{\beta,t} + e_{\beta,t+1} \tag{21}$$

$$u_{\delta,t+1} = \rho_\delta u_{\delta,t} + e_{\delta,t+1} \tag{22}$$

Since $\partial\beta_{t+1}/\partial k_t = \partial\delta_{t+1}/\partial k_t = 0$, the optimality conditions are

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

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

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

- Since $E(d_t) = \beta, E(\delta_t) = \delta$, the steady states are as in model A.

- Since $\phi_x = 0$, P and Q are as in model A.

- Is R $\neq 0$? Check whether the columns of $N\phi_u\omega_u + O\phi_u$ are zero.

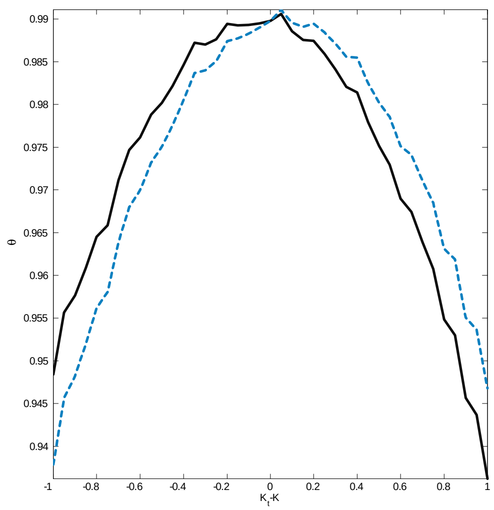
$$N\phi_u\omega_u + O\phi_u = \begin{pmatrix} 0 & 0 \\ -1/\beta & -\rho_\delta/\beta \\ 0 & -k \\ 0 & 0 \end{pmatrix} \neq 0 \qquad (24)$$

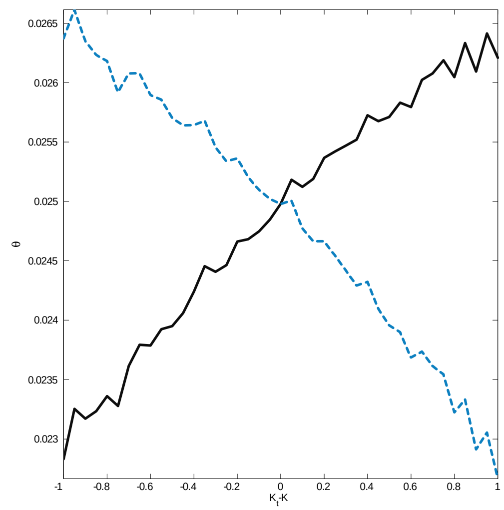
• Note: if $d_{t+1} = \beta_{t+1}/\beta_t$ (d_t is a fast moving variable), time variations in d_t matter only if $\rho_d \neq 0$

Model C: State dependent parameter variations, no internalization

$$\Theta_{1t+1} = [\Theta_u - (\Theta_u - \Theta_l)e^{-\phi_1(K_t-K)}] + [\Theta_u - (\Theta_u - \Theta_l)e^{\phi_2(K_t-K)}] + u_{\theta,t+1} \tag{25}$$

$\phi_1, \phi_2, \Theta_u, \Theta_l$ are vectors. Let $u_{\theta,t+1}$ be zero mean, iid shocks.





- If $d_l = \beta/2$ and $\delta_l = \delta/2$, the steady states are as model A.

- Assume that agents treat K_t appearing in (25) as an aggregate variable: $\partial\beta_{t+1}/\partial K_t = \partial\delta_{t+1}/\partial K_t = 0$. The optimality conditions are as in (23).

- Are the P and Q matrices affected?

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

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

- P and Q differ from those of model A if $\phi_1 \neq \phi_2$ or $\phi_3 \neq \phi_4$

- Do time variation in Θ_{1t} affect X_t?

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

- $R \neq 0$ if $\phi_1 \neq \phi_2$ or $\phi_3 \neq \phi_4$.

Model D: State dependent parameter variations, internalization

The relevant derivatives are

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

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

If $\phi_1 = \phi_2 = \phi_1, \phi_3 = \phi_4 = \phi_3$, the steady states are as in model A. Here

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

$$E_t \begin{pmatrix} AC_t^{\eta} N_t^{\gamma+1} - (1-\alpha)(1-g_t)Y_t \\ 1 - d'_t\ u(C_{t+1}, N_{t+1})/C_t^{-\eta} - d_t\ C_{t+1}^{-\eta}/C_t^{-\eta}(\alpha(1-g_{t+1})Y_{t+1}/K_{t+1} + 1 - \delta_{t+1} + \delta'_{t+1}K_t) \\ (1-g_t)Y_t - C_t - K_t + (1-\delta_t)K_{t-1} \\ Y_t - \zeta_t K_{t-1}^{\alpha} N_t^{1-\alpha} \end{pmatrix} = 0$$

where $X_t = (K_t, Y_t, C_t, N_t)'$, $Z_t = (\zeta_t, g_t)'$ but now

$$\Theta_t = \begin{pmatrix} d_{t+1} \\ \delta_{t+1} \\ d'_{t+1} \\ \delta'_{t+1} \end{pmatrix} = \Phi(\Theta, K_t, u_{t+1}) \quad =$$

$$\begin{pmatrix} 2d_u - (d_u - \beta/2)[e^{-\phi_1(K_t - K + u_{\beta,t+1})} + e^{\phi_1(K_t - K + u_{\beta,t+1})}] \\ 2\delta_u - (\delta_u - \delta/2)[e^{-\phi_3(K_t - K + u_{\delta,t+1})} + e^{\phi_3(K_t - K + u_{\delta,t+1})}] \\ -(d_u - \beta/2)\phi_1[-e^{-\phi_1(K_t - K)} + e^{\phi_1(K_t - K)}] \\ -(\delta_u - \delta/2)\phi_3[-e^{-\phi_3(K_t - K)} + e^{\phi_3(K_t - K)}] \end{pmatrix} \tag{31}$$

The relevant matrices are:

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

$$O = \frac{\partial f}{\partial \Theta_t} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ -1/\beta & 0 & 0 & 0 \\ 0 & -K & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

$$\phi_x = \frac{\partial \Phi}{\partial X_t} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -2(d_u - \beta/2)\phi_1^2 & 0 & 0 & 0 \\ -2(\delta_u - \delta/2)\phi_3^2 & 0 & 0 & 0 \end{pmatrix}$$

$$\phi_u = \frac{\partial \Phi}{\partial u_{t+1}} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \\ -2(d_u - \beta/2)\phi_1^2 & 0 \\ 0 & -2(\delta_u - \delta/2)\phi_3^2 \end{pmatrix}$$

Then $N\phi_x \neq 0$, $O\phi_x = 0$ and $N\phi_u\omega_u + O\phi_u = 0$ since $\omega_u = 0_{2\times 2}$. Thus, P is affected even when $\phi_1 = \phi_2$ and $\phi_3 = \phi_4$.

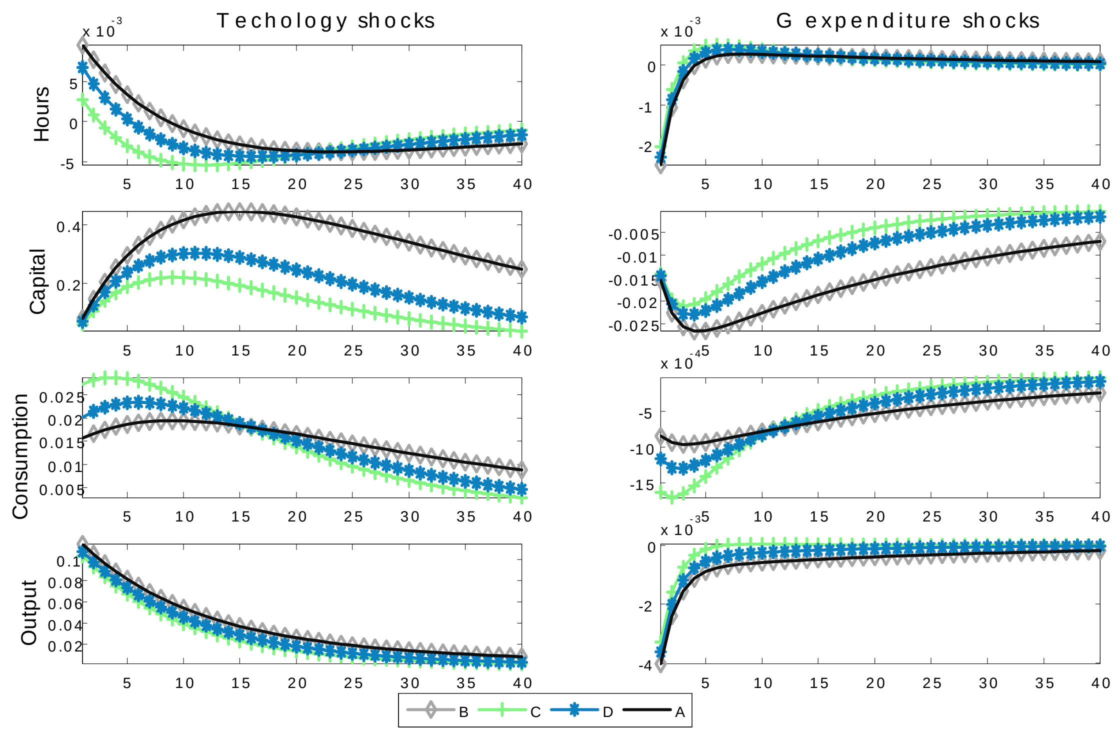
Why are the structural dynamics in models C and D different?

- Common parameters: $\alpha = 0.30$, $\beta = 0.99$, $\delta = 0.025$, $\gamma = 2$, $\eta = 2$, $A = 4.50$, $\bar{\zeta} = 1$; $\rho_z = 0.90$, $\sigma_z = 0.0712$, $\bar{g} = 0.18$, $\rho_g = 0.50$ and $\sigma_g = 0.0052$.

- Model B : $\rho_\beta = 0.985$; $\rho_\delta = 0.95$; $\sigma_\beta = 0.002$;$\sigma_\delta = 0.07$.

- Model C : $\phi_{1\beta} = 0.01$; $\phi_{2\beta} = 0.03$; $\phi_{1\delta} = 0.2$; $\phi_{2\delta} = 0.1$; $\sigma_\beta = \sigma_\delta = 0.5$; $d_u = 0.999$; $\delta_u = 0.025$.

- Model D : $\phi_{1\beta} = 0.0001$; $\phi_{2\beta} = 0.016$; $\phi_{1\delta} = 0.2$; $\phi_{2\delta} = 0.1$; $\sigma_\beta = 0.0001$,$\sigma_\delta = 0.1$; $d_u = 0.999$; $\delta_u = 0.025$.



- Income and substitution effects are different!

Punchline

- Having exogenous parameter variations, is like adding shocks to the original model without altering the existing dynamics.

- Approximating an exogenous TVC model with a constant coefficient model may be less costly than approximating a endogenous TVC model.

- Dynamics in models with endogenous time variations may differ from those of exogenous and time invariant models because income and substitution effects are altered.

Characterizing time invariant misspecification

Two ways:

- "Wedges" (see Chari et al., 2008).

- Forecast errors.

Wedges

CC Model:

$$0 = [F(X_{t-1}, W_t, \sigma\Sigma^z\epsilon^z_{t+1}, \Theta)] = 0 \quad (32)$$

$$X_t = h(X_{t-1}, W_t, \sigma\Sigma^z\epsilon^z_{t+1}, \Theta) \quad (33)$$

TVC Model:

$$0 = [F^*(X^*_{t-1}, W_t, \sigma\Sigma\epsilon_{t+1}, \Theta)] = 0 \quad (34)$$

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

Wedge:

1) $[F(X^*_{t-1}, W_t, \sigma\Sigma^z\epsilon^z_{t+1}, \Theta)] \neq 0$ (since $\Sigma^z\epsilon^z_{t+1} \neq \Sigma\epsilon_{t+1}; h \neq h^*$).

2) $[F(X^*_{t-1}, W_t, \sigma\Sigma^z\epsilon^z_{t+1}, \Theta)]$ predictable using past X^*_{t-1}

In first order system the wedge is

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

• If $P^* = P, Q^* = Q$ wedge is

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

Different from zero if $R^* \neq 0$ and predictable if $\omega_u \neq 0$

• If $P^* \neq P, Q^* \neq Q$ wedge is non zero (even if $R = 0$) and predictable using past x_t even if $\omega_u = 0$.

Forecast errors

- Linearized decision rule in time invariant model: $x_t = P^x_{t-1} + Qz_t$

- Linearized decision rule in TVC model: $x^*_t = P^* x^*_{t-1} + Q^* z_t + R^* u_t$.

Let v^*_t be the forecast error in predicting x^*_t using the decision rules of the constant coefficient model and TVC data

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

• Forecast error is function of the lags of the observables x^*_{t-1}.

• True when $P^* \neq P$ but also if P^* if u_t are serially correlated (they affect x^*_{t-1}).

Misspecification diagnostic: RBC example

Euler wedge.

DGP	c_{t-1}	r_{t-1}
B	-0.008	1.06
	(0.04)	(0.14)
C	0.06	1.81
	(0.02)	(0.21)
D	-0.05	-0.59
	(0.09)	(0.17)

Forecast error: hour equation.

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

- Counterfactual: Constant coefficients but one period time to build model. First lag real rate -0.03, first lag of consumption growth -0.06, both insignificant.

Puchline

• Can detect time invariant misspecification by estimating a time invariant model and a bunch of regressions.

• Much less costly than, e.g. rolling estimation or estimating different versions of the model (with and without TV coefficients).

• Exploit the structure of optimality conditions and of forecast errors.

How to detect exogenous vs. endogenous time variations?

- Use a DSGE-VAR (Del Negro and Schorfheide, 2004) setup.

Idea:

- Simulate T_1 data point for each model. Add it to actual T data points.

- If simulated data come from the DGP, precision of estimates improve, ML increase. If simulated data does not come to DGP, noise is added, precision decrease (bias may be generated), ML decrease.

- Compare ML of adding T_1 data from the exogenous model to the ML of adding T_1 data from the endogenous model.

RBC example: T=150

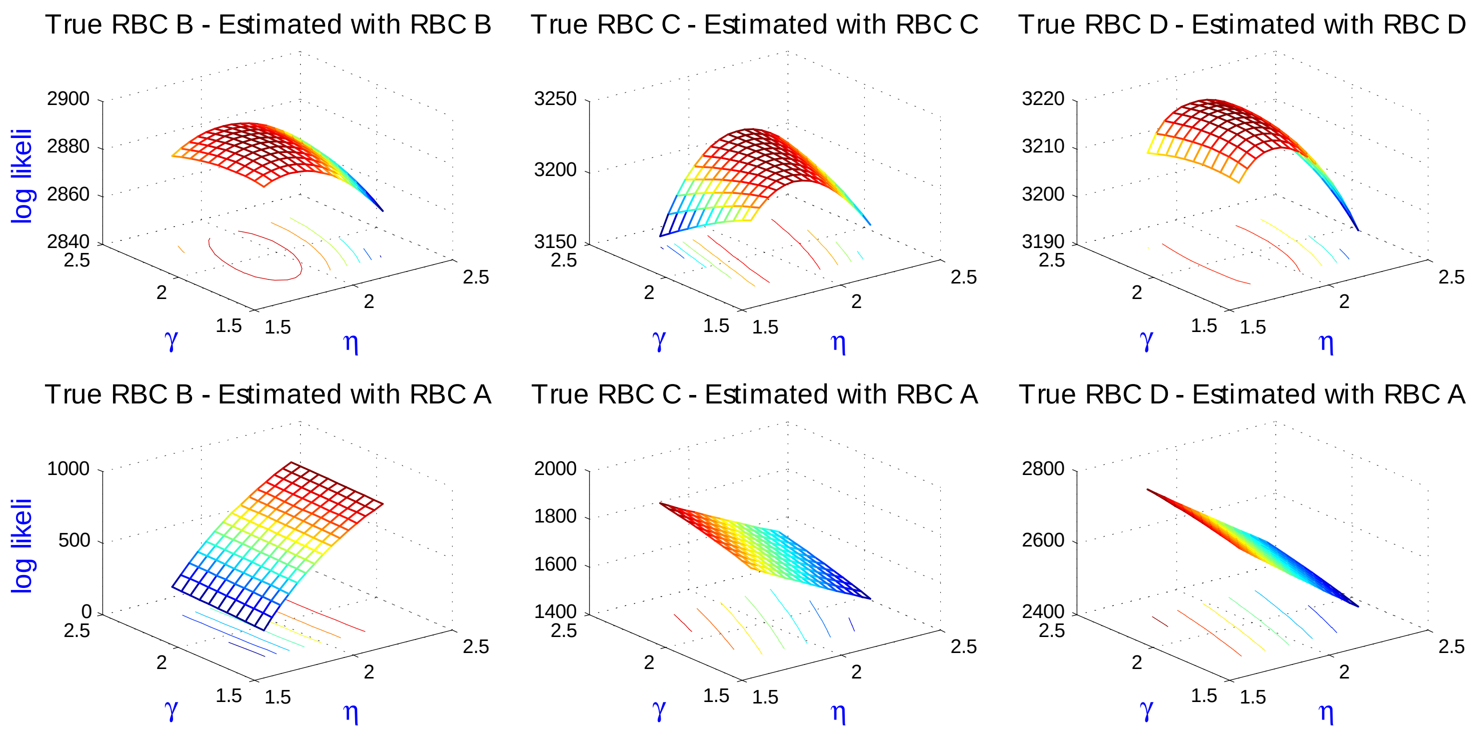
	Log Marginal likelihood					
	T_1=150			T_1=750		
DGP	Model B	Model C	Model D	Model B	Model C	Model D
Simulated from B	1586	-6709	-5108	9714	-3478	-12597
Simulated from C	1421	2005	-855	7480	4828	-409
Simulated from D	697	-2649	1864	6083	622	11397

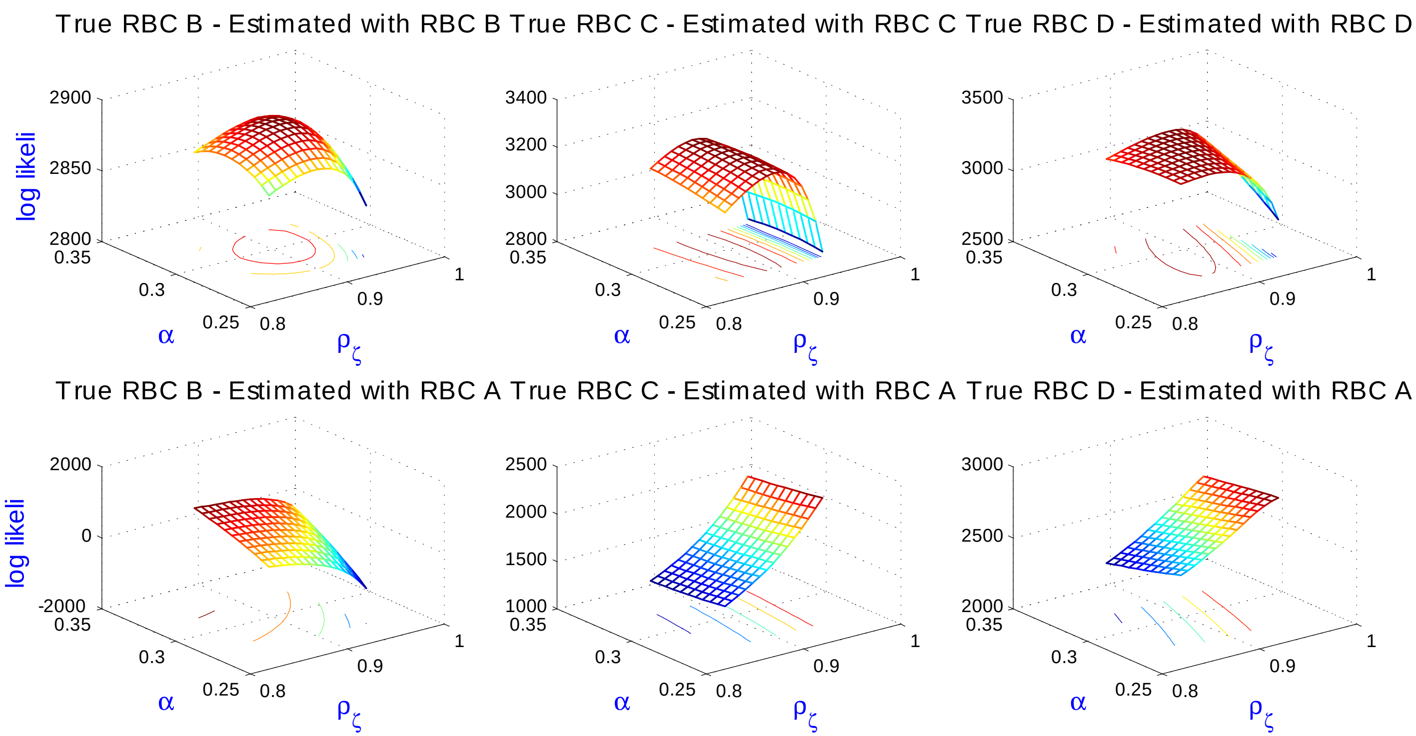
Inferential distortions 1: parameter identification

• Can time invariant parameters be identified from a potentially misspecified likelihood function?

- Canova and Sala (2009): DSGE models have POPULATION identification problems.

- Can identification problems arise because parameter variations are neglected? What should we expect to happen to the likelihood function when time invariant parameters are assumed?


True RBC B - Estimated with RBC B
True RBC C - Estimated with RBC C
True RBC D - Estimated with RBC D
True RBC B - Estimated with RBC A
True RBC C - Estimated with RBC A
True RBC D - Estimated with RBC A
log likeli
log likeli
γ
η
1.5
2
2.5
2840
2860
2880
2900
3150
3200
3250
3190
3200
3210
3220
0
500
1000
1400
1600
1800
2000
2400
2600
2800


True RBC B - Estimated with RBC B
True RBC C - Estimated with RBC C
True RBC D - Estimated with RBC D
True RBC B - Estimated with RBC A
True RBC C - Estimated with RBC A
True RBC D - Estimated with RBC A
log likeli
α
ρ ζ
2900
2850
2800
3400
3200
3000
2800
3500
3000
2500
2000
0
-2000
2500
2000
1500
1000
3000
2500
2000
0.35
0.3
0.25
0.8
0.9
1

Koop, Pesaran, Smith identification 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

- Curvature of the correct likelihood OK. Maximum is at $\gamma = 2, \eta = 2$, $\alpha = 0.3, \rho_z = 0.9$ for all models.

- When the decision rules of model A are used and the truth are models B-C-D, distortions are large and some partial identification problem exist (notice ridges in contours).

- Misspecification of P and Q distorts the likelihood. Neglecting existence of other shocks tilts the likelihood without changing location.

- KPS statistic agrees: weak identification only for ρ_g with some DGPs.

- Results consistent with Huang (2014).

Inferential distortions 2: ML based parameter and response estimates

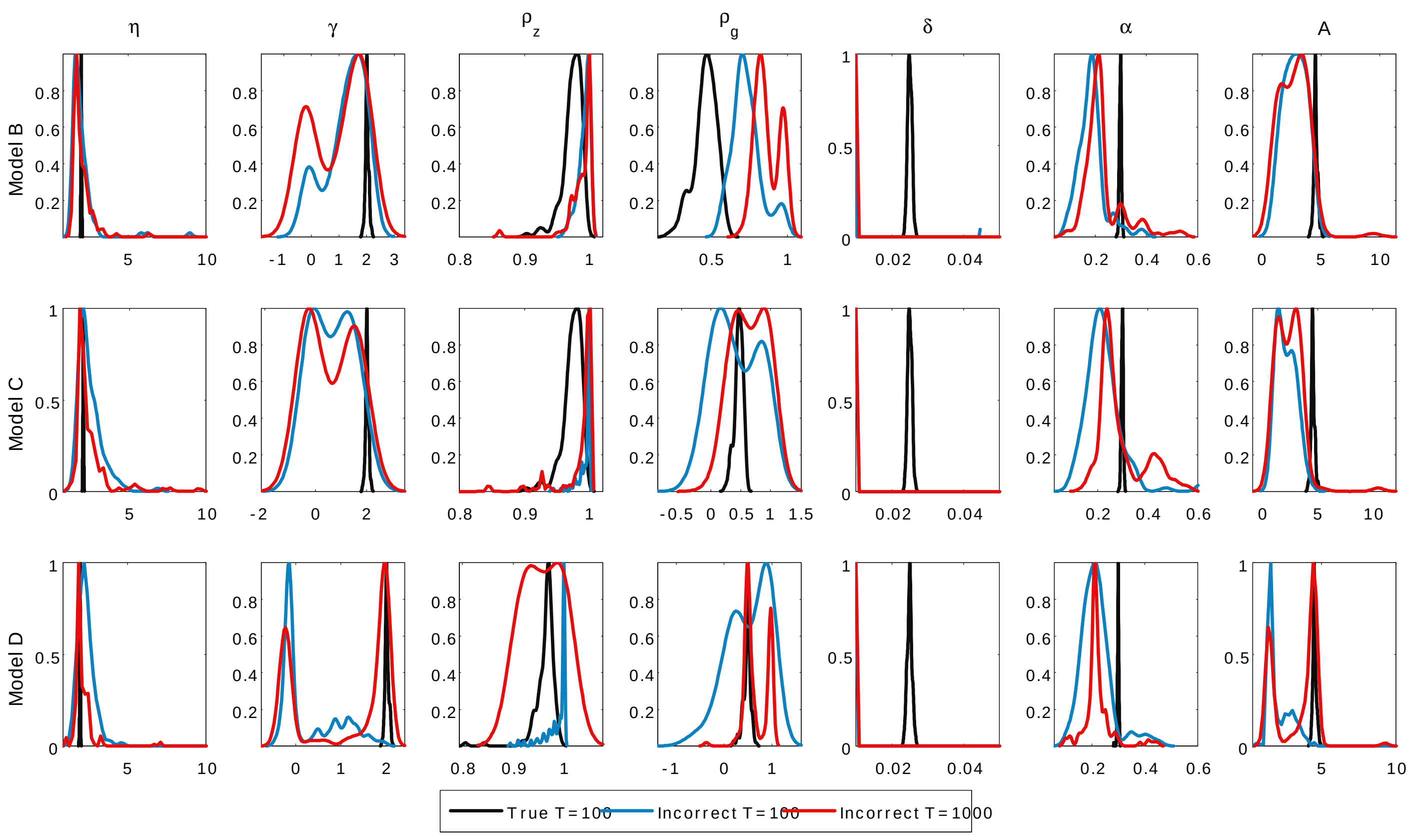
- Use RBC models B, C, D; N=150; T=100 or T=1000.

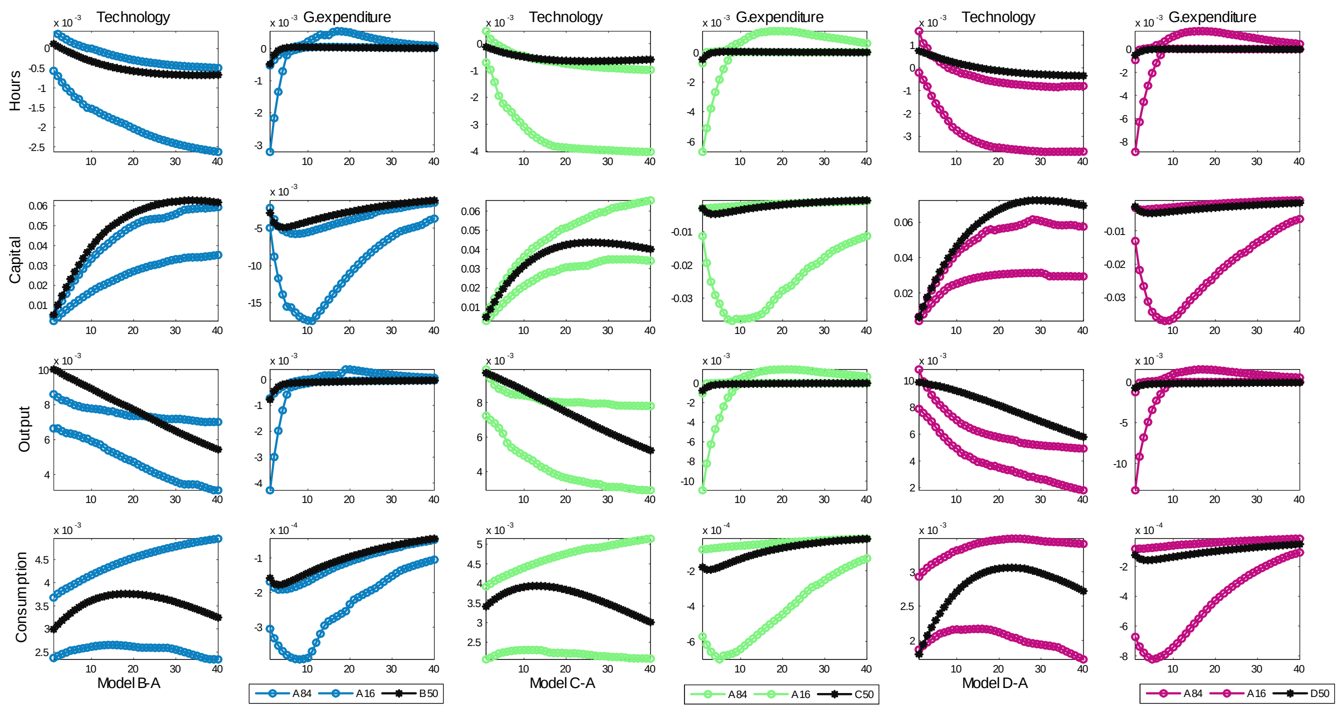
- Two DGPs: one where parameter variations explain little of output variability (less than 5 percent); one where parameter variations explain a sizable portion of output variability (20 percent).

- Estimate a time invariant model and the correct model (to control for possible identification and numerical problems).

- Compare parameter estimates, impulse responses and variance decompositions.

Small time variations							
True Parameter	Correct Mean T=150	Time invariant Mean	5th percentile	95th percentile	Time invariant Mean	5th percentile	95th percentile
		T=150			T=1000		
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

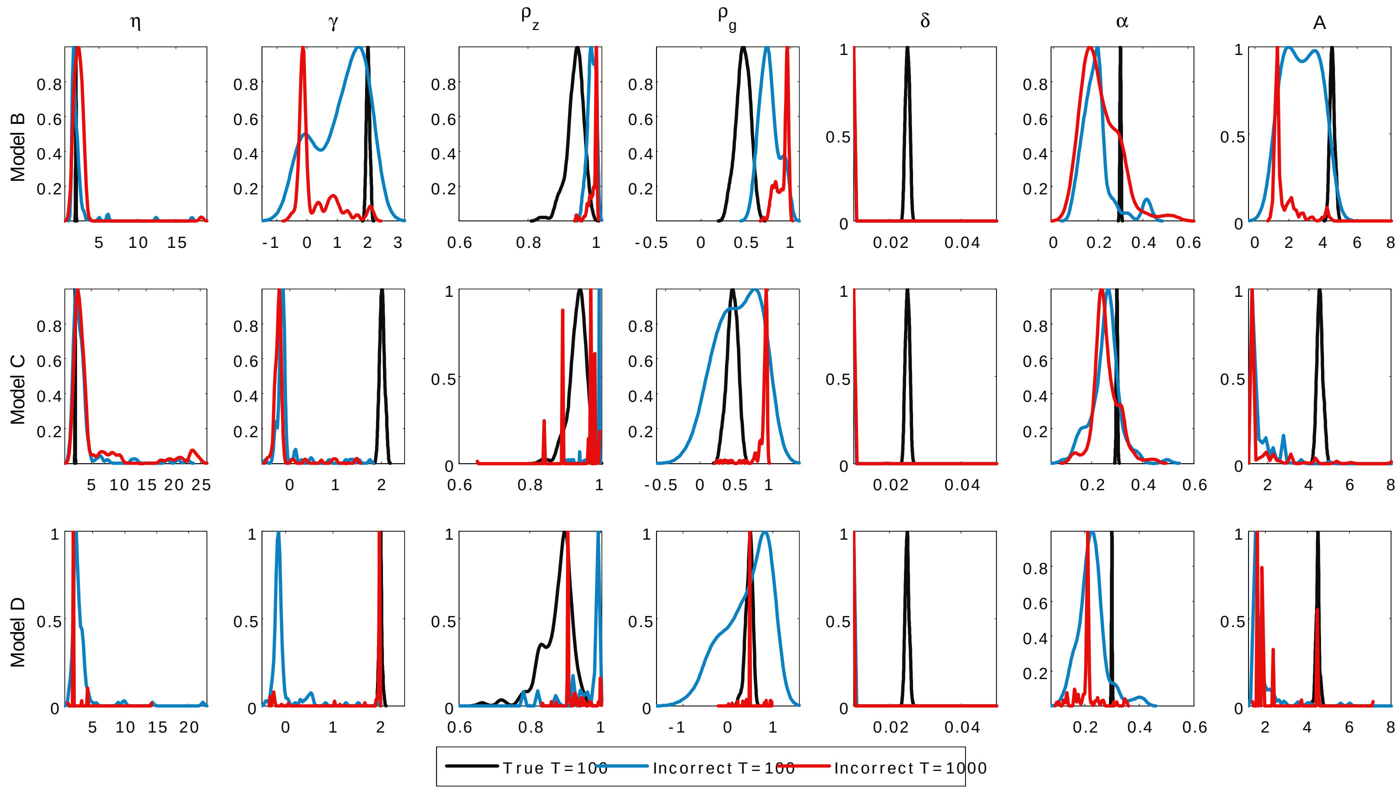

η
γ
ρ z
ρ g
δ
α
A
Model B
Model C
Model D
True T=100
Incorrect T=100
Incorrect T=1000



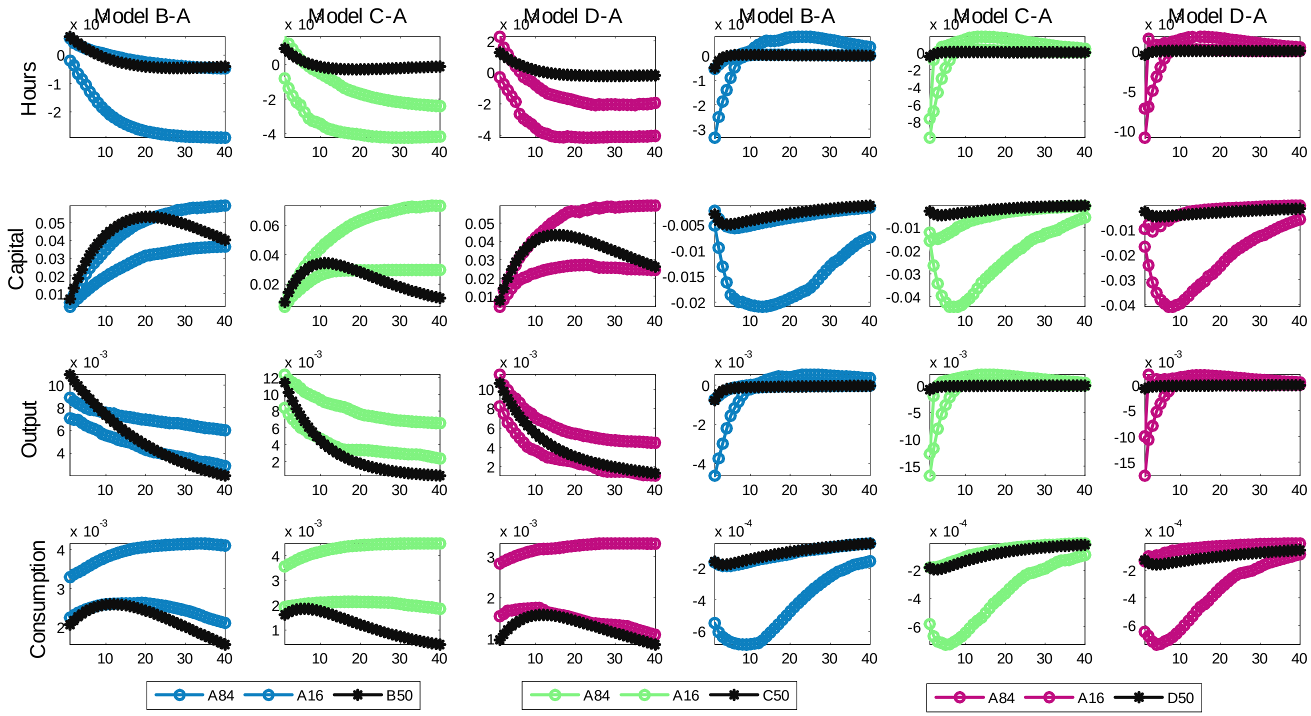
Impulse responses, DGP1

	Variance decomposition			
	DGP: Small time variations			
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

Large time variations							
True Parameter	Correct Mean	Time invariant			Time invariant		
		Mean	5th percentile	95th percentile	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



η
γ
ρ z
ρ g
δ
α
A
Model B
Model C
Model D
True T=100
Incorrect T=100
Incorrect T=1000



Impulse responses, DGP2

	Variance decomposition			
	DGP:Large time variations			
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

Punchline

- Large distortions in γ and ρ's , i.e. the parameter controlling income and substitution effects

- Results with T=1000 do not improve

- If the DGP has more important time variations, results worsen.

SVAR based inference

- Can less structural methods be useful to recover structural dynamics?

- Canova and Paustian (2011): robust SVAR methods can be used if theoretical model is misspecified in certain way.

- Is the conclusion still true if we neglect time variations in the structural parameters?

- Use the DGP where parameter variations explain less than 5 percent of the variance of output.

- Simulate (long) data using the decisions rules of a TVC model.

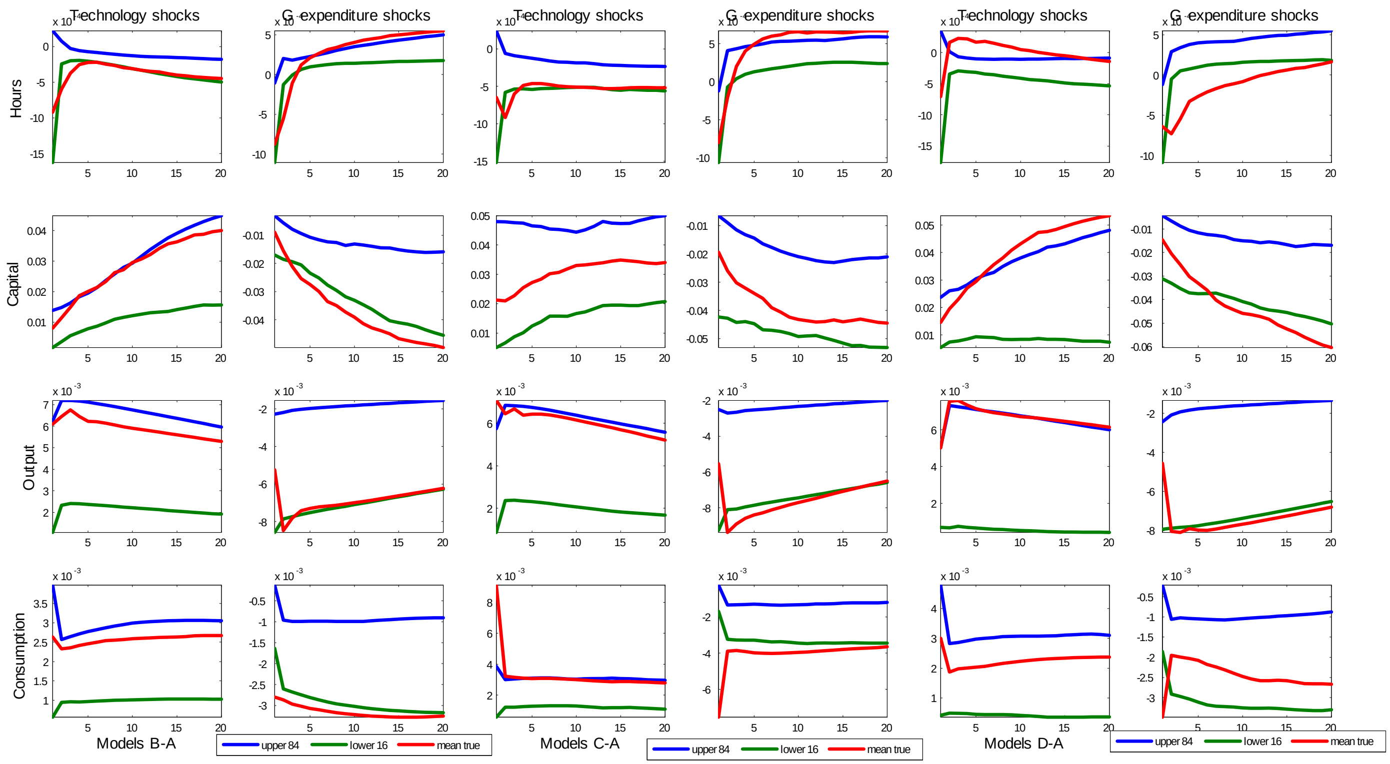
- Construct VAR errors using the dynamics (matrix P) of a constant coefficient model.

- Rotate VAR errors until they satisfy robust sign restrictions (common to all models):

Technology shocks: Y, H, K up;

Government expenditure shocks: Y, H, C and K down.

- For each replication $l = 1, \ldots, L$ store the median of the true and incorrectly specified model. Report the 68 percent interval for the median of the incorrectly specified model and the median value of median for the correctly specified model.



Technology shocks
G expenditure shocks
Technology shocks
G expenditure shocks
Technology shocks
G expenditure shocks
Hours
Capital
Output
Consumption
Models B-A
Models C-A
Models D-A
upper 84
lower 16
mean true

- Performance good also with models C and D. Only consumption responses not well captured - entries of P matrix off primarily in one row.

- Performance is good since shock misaggregation is minor.

Estimation of Gertler and Karadi (2010) model

Three goals:

i) Estimate the parameters specific to the model (λ : the share of projects that bankers can steal, ω : the fraction of wealth given to new bankers, θ : the lifetime horizon of a banker).

ii) Apply diagnostics for TV misspecification

iii) Estimate TVC models. Compare responses to a capital quality shock to those in a fixed coefficient model.

Parameter estimates

Parameter	Mode	Gertler- Karadi
λ	0.2457	0.381
	(0.0182)	
ω	0.0115	0.002
	(0.0008)	
θ	0.4646	0.972
	(0.0098)	
SS Leverage	3.32	1.38
Lifetime bank	$<$1y	$\approx$10y

Forecast error regressions diagnostic

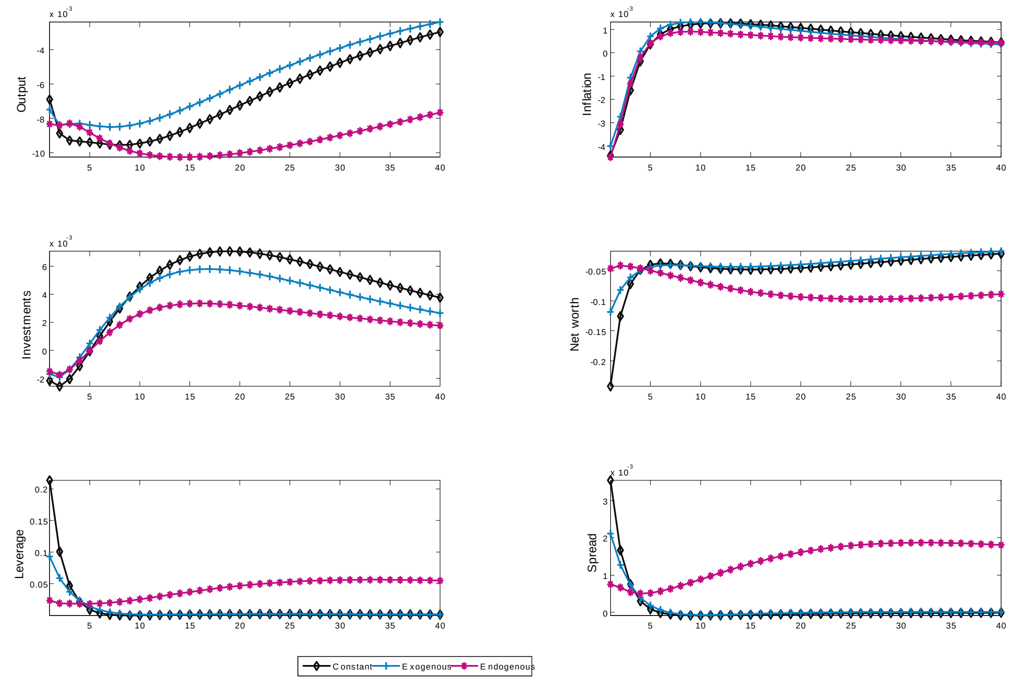
Equation	T-stat					F-stat
	Y_{t-1}	C_{t-1}	Credit$_{t-1}$	Leverage$_{t-1}$	Spread$_{t-1}$	
Y	0.84	2.61	0.24	0.52	10.00	4.39
C	-0.85	1.11	0.85	-0.65	0.33	1.26
Credit	1.06	2.61	1.65	-0.58	8.49	7.11
Leverage	-1.11	-2.50	-1.63	0.63	-8.25	7.04
Spread	-1.26	-3.06	-1.10	0.81	-8.46	8.16

Wedge diagnostic

	Mean	C_{t-1}	IY_{t-1}
Euler wedge	0.02	-0.10	0.72
	(0.03)	(0.01)	(0.13)

Estimates with TVC models

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	



Capital quality shock

A remaining question

- How do you distinguish a model with m shocks from a model with m_1 shocks and m_2 time varying parameters,$m_1 + m_2 = m$? Or with m_1 shocks and m_2 measurement errors?

Solution with m_1 shocks and m_2 time varying parameters:

$$x_t = Px_{t-1} + Q_1 z_{1t} + R_1 u_t = Px_{t-1} + Qz_t \tag{39}$$

Q=$[Q_1, R_1]$, $z_t = [z'_{1t}, u'_t]'$.

Solution with m_1 shocks and m_2 measurement errors v_t is

$$x_t = Px_{t-1} + Q_1 z_{1t} + v_t = Px_{t-1} + Qz_t \tag{40}$$

Q=$[Q_1, I]$, $z_t = [z'_{1t} v'_t]'$.

• If $v_t \equiv Ru_t$ very hard to distinguish the two unless u_t makes P, Q_1 different.

Solution with m shocks is

$$x_t = Px_{t-1} + Qz_t \tag{41}$$

Q full matrix.