Frequentist evaluation of small DSGE models

Gunnar Bårdsen (Trondheim), Luca Fanelli (Bologna)

IV International Conference in Memory of Carlo Giannini,
Pavia, 25-26 March 2014

Motivations

- We want **'understand what happens'** with NK-DSGE models.
- NK-DSGE models are stylezed representation of an economy, hence they are **misspecified by definition**. They are typically estimated by using **Bayesian methods.**
- The only existing 'metric' to evaluate these models is the **DSGE-VAR approach** by Del Negro et al. (2007, JBES).

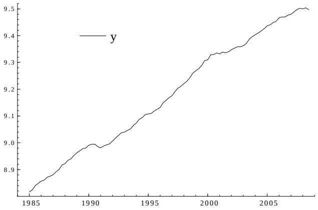
Motivations

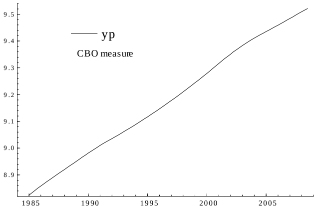
- Pesaran and Smith (2011, ManSch): DSGE →**straitjacket** !
- **Our starting point:** the scientific validity of a model should not be exclusively based on its logical coherence or its intellectual appeal, but also **on its capability of making empirical predictions that are not rejected by the data** (De Grauwe, 2010, PuChoice).
- **Our evaluation 'metric'**: **testing the restrictions NK-DSGE model imply on the data** .

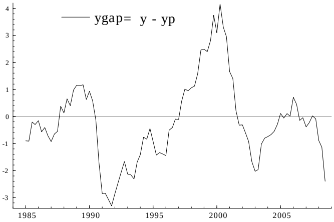
Objectives

- Purpose of this paper: **Evaluate** the NK-DSGE model using **classical/frequentist statistical tests**.
- The solution of a NK-DSGE model is a system that embodies
  - a set of recoverable **cointegration/common-trend restrictions**
  - a set of short-run **CER restrictions**.
  - **Canova, Finn and Pagan (1994, Book); Söderlind and Vredin (1996, JAE**).
- **Our contribution 1:** the two type of restrictions are **interrelated** and must be tested **jointly** through a **multiple hypothesis testing approach**.
- **Our contribution 2** (Bekaert and Hodrick, 2001, JoF)**:** we **reject too often also because of the poor small sample properties of the tests we use**. **Bootstrap methods !**

Why does non-stationarity matter ?

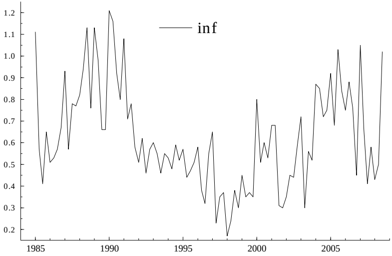


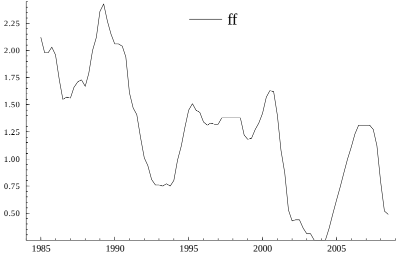




1985q1-2008q3

Why does non-stationarity matter ?





1985q1-2008q3

This paper's contribution: testing strategy

- We use an **approximating VAR** for the data.
- The suggested procedure involves computing **three tests:**

1. LR1: cointegration rank test ('one-shot version'/'sequential');
2. LR2: overidentifying cointegration relationships (β) test;
3. LR3: CER restrictions test.

- LR2 is run on condition that LR1 **does not reject** the cointegration rank;
- LR3 is run on condition that LR2 **does not reject** the overidentification cointegrating restrictions.
- **Joint test**: $\text{LR1}_{1\%} \rightarrow \text{LR2}_{2\%} \rightarrow \text{LR3}_{2\%}$

This paper's contribution: the NK-DSGE model goes to the bootstrap

- We consider also the **bootstrap analogue** of the 'LR1→LR2→LR3' sequence, where:

1. **LR1:** Cavaliere, Rahbek and Taylor (2012, Ecta);
2. **LR2:** Boswijk, Cavaliere, Rahbek and Taylor (2013, WP);
3. **LR3:** Chow and Moreno (2006, JMCB), adaptation from Fanelli and Palomba (2011, JAE).

This paper's contribution: think positive !

- The individual tests can be **used constructively** to rectify/modify, when possible, the baseline specification of the NK-DSGE model.
- If for instance LR1 and LR2 reject the common trend implications of the baseline specification, it is necessary to have in mind **an alternative framework** which e.g. accounts for the 'additional' stochastic trends. $\longrightarrow$ Example.
- Variations in the **common trend features** **have marked consequences** on the **structural parameters and CER.**

Connections to the literature

- **Canova, Finn and Pagan (1994, Book)** and **Söderlind and Vredin (1996, JAE)**: real business cycle models, **nonstationarity/cointegration** testable implications **other than short run CER**;
- **Fanelli (2008, OBES), Fukač and Pagan (2010, JAE)**, **single-equation 'limited-information' approach**;
- **Gorodnichenko and Ng (2010, JME)**: robust filters to both the model and the data. **No testing of the implications**. ML estimation is not 'practical' in their case.

Connections to the literature: the Bayesian DSGE-VAR approach

- Del Negro et al. (2007, JBES) also use a VAR approximation of the data.
- They impose - **without testing** - the common-trend/cointegration restrictions (**we use** LR1 and LR2).
- Prior distribution for the VAR parameters centred on the CER with dispersion governed by a scalar (hyper-)parameter, λ.
- **Small values of** λ: the VAR **is far** from the theoretical model; **large values of** λ: the VAR **is close** to the theoretical model. **No cutoff value is provided** (Christiano, 2007, JBES) !
- Our LR3 test **plays a role similar to** λ but **we do have a cutoff value** (which depends on the pre-fixed type I error)!

Outline

- Model
- Testable restrictions
- 'LR1→LR2→LR3' testing strategy:
  $\begin{cases} \text{all variable observed} \\ \text{observed \& unobserved} \end{cases}.$
- Asymptotic size.
- What if LR1 rejects ?
- Monte-Carlo experiment.
- Empirical illustration on U.S. quarterly data taking Benati and Surico's (2009) NK-DSGE model as reference model.
- Remarks for practitioners.

Reference NK-DSGE model

- Benati and Surico (2009, AER) model:

$$\textbf{AD:} \quad \tilde{y}_t = \gamma E_t \tilde{y}_{t+1} + (1-\gamma)\tilde{y}_{t-1} - \delta(i_t - E_t \pi_{t+1}) + \eta_{\tilde{y},t}$$

$$\textbf{NKPC:} \quad \pi_t = \frac{\varrho}{1+\varrho\varkappa} E_t \pi_{t+1} + \frac{\varkappa}{1+\varrho\varkappa}\pi_{t-1} + \kappa \tilde{y}_t + \eta_{\pi,t}$$

$$\textbf{Policy rule:} \quad i_t = \rho i_{t-1} + (1-\rho)(\varphi_\pi \pi_t + \varphi_y \tilde{y}_t) + \eta_{i,t}.$$

$$\textbf{Shocks:} \quad \eta_{a,t} = \rho_a \eta_{a,t-1} + u_{a,t} \ \ , \ \ u_{a,t} \sim WN\left(0, \sigma_a^2\right) \ \ , \ a = \tilde{y}, \pi, i$$

$$\theta := (\gamma,\ \delta,\ \varrho,\ \varkappa,\ \rho,\ \varphi_\pi,\ \varphi_y,\ \rho_{\tilde{y}},\ \rho_\pi,\ \rho_i,\ \sigma_{\tilde{y}}^2, \sigma_\pi^2, \sigma_i^2)'.$$

NK-DSGE model: representations

- **Compact representation**

$$B_0(\theta)W_t = B_f(\theta)E_t W_{t+1} + B_b(\theta)W_{t-1} + \eta_t^W$$

$$\eta_t^W = R_W(\theta)\eta_{t-1}^W + u_t^W \quad , \quad u_t^W \sim \mathrm{WN}(0_{p\times 1}, \Sigma_{W,u}(\theta)).$$

- Finite-order **reduced form solution**

$$W_t = \tilde{F}_1(\theta)W_{t-1} + \tilde{F}_2(\theta)W_{t-2} + \varepsilon_t^W \quad , \quad \varepsilon_t^W = \tilde{Q}(\theta)u_t^W$$

  where $\tilde{F}_1 = F_1(\theta)$, $\tilde{F}_2 = F_2(\theta)$ and $\tilde{Q} = Q(\theta)$ **embody the CER** (Binder and Pesaran, 1995, Book).

- ABC (and D's) representation: Hannan and Deistler (1988, Book) —*****— Giacomini (2013, Book).

NK-DSGE model: representations

- We interpret W_t as

$$W_t := H' Z_t \quad , \quad Z_t := \begin{pmatrix} W_t^o \\ W_t^u \end{pmatrix} \begin{array}{c} o \times 1 \\ (n-o) \times 1 \end{array}$$

- Z_t is the 'complete' $n \times 1$ vector of variables and H **selects the stationary elements/combinations of Z_t that enter the structural model**.

Our example

- In the Benati and Surico (2009, AER) model, $W_t := (\tilde{y}_t, \pi_t, i_t)'$ is $p \times 1$, $p=3$ and

$$Z_t = \begin{pmatrix} W_t^o \\ W_t^u \end{pmatrix} \begin{matrix} 3 \times 1 \\ 1 \times 1 \end{matrix} = \begin{pmatrix} y_t \\ \pi_t \\ i_t \\ -- \\ y_t^p \end{pmatrix} \qquad \tilde{y}_t := y_t - y_t^p$$

- Thus W_t can be thought of as being obtained through

$$\underset{W_t}{\begin{pmatrix} \tilde{y}_t \\ \pi_t \\ i_t \end{pmatrix}} = \underset{H'}{\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}} \underset{Z_t}{\begin{pmatrix} y_t \\ \pi_t \\ i_t \\ y_t^p \end{pmatrix}} \quad \text{Hp: } \begin{matrix} y_t^p = y_{t-1}^p + \eta_t^{y^p} \\ \text{technology shock} \end{matrix}$$

Complete specification

- Consider the $n \times 1$ vector $Z_t := (W_t^{o\prime}, W_t^{u\prime})'$
- **Assumption:** ΔW_t^u is covariance stationary.
- Re-formulate the NK-DSGE model with respect to Z_t:

$$A_0 Z_t = A_f E_t Z_{t+1} + A_b Z_{t-1} + \eta_t^Z$$

$$\eta_t^Z = R_Z \eta_{t-1}^Z + u_t^Z \quad , \quad u_t^Z \sim \text{WN}(0_{n\times 1}, \Sigma_{u,Z}).$$

All matrices depend on θ.

Testable restrictions I

- Z_t is I(1) **and cointegrated.**
- We know the **exact structure** of the cointegrating relationships, i.e. we know that $\beta_0' Z_t = H' Z_t \sim$I(0).
- **In our example:**

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}}_{r=3} \begin{pmatrix} y_t \\ \pi_t \\ i_t \\ y_t^p \end{pmatrix} \sim \text{I}(0) \quad \text{or} \quad \underbrace{\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{r=2} Z_t^o \sim \text{I}(0).$$

Mapping from I(1) to I(0) I

- Thus we can look for explicit mappings to I(0)

$$Y_t = \underbrace{\begin{pmatrix} \beta_0' \\ \tau'(1-L) \end{pmatrix}}_{\text{nonsingular}} Z_t = G(\beta_0, \tau, 1-L) Z_t$$

$$\beta_0 = H \ , \qquad \det(\tau' \beta_{0,\perp}) \neq 0$$

- In our example Y_t is given by

$$Y_t = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ - & - & - & --- \\ 0 & 0 & 0 & (1-L) \end{pmatrix} \begin{pmatrix} y_t \\ \pi_t \\ i_t \\ y_t^p \end{pmatrix} = \begin{pmatrix} W_t \\ \Delta y_t^p \end{pmatrix}.$$

Mapping from I(1) to I(0) I

- Given
$$A_0 Z_t = A_f E_t Z_{t+1} + A_b Z_{t-1} + \eta_t^Z$$
- Invert $Y_t = G(\beta_0, \tau, 1-L) Z_t$, obtaining $Z_t = G(\beta_0, \tau, 1-L)^{-1} Y_t$, and plug-in model:

$$A_0 G(\beta_0, \tau, 1-L)^{-1} Y_t = A_f G(\beta_0, \tau, 1-L)^{-1} E_t Y_{t+1} + A_b G(\beta_0, \tau, 1-L)^{-1} Y_{t-1} + \eta_t^Y$$

$$A_0^Y \ Y_t = A_f^Y \ E_t Y_{t+1} + A_b^Y \ Y_{t-1} + \eta_t^Y \quad , \quad \eta_t^Y = R_Y \ \eta_{t-1}^Y + u_t^Y$$

- If all model restrictions are met this system **involves only stationary variables**. It is the **balanced equilibrium-correction representation** of the NK-DSGE model.

Testable implications: the CER

- Given

$$A_0^Y Y_t = A_f^Y E_t Y_{t+1} + A_b^Y Y_{t-1} + \eta_t^Y \quad , \quad \eta_t^Y = R_Y \eta_{t-1}^Y + u_t^Y$$

the unique stable VAR solution is

$$Y_t = \tilde{\Phi}_1(\theta) Y_{t-1} + \tilde{\Phi}_2(\theta) Y_{t-2} + \varepsilon_t^Y \quad , \quad \varepsilon_t^Y = \tilde{\Psi}(\theta) u_t^Y$$

where the CER are given by:

$$(A_0^{Y,R} - A_f^Y \tilde{\Phi}_1)\tilde{\Phi}_1 - A_f^Y \tilde{\Phi}_2 + A_{b,1}^{Y,R} = 0_{n\times n}$$
$$(A_0^{Y,R} - A_f^Y \tilde{\Phi})\tilde{\Phi}_2 - A_{b,2}^{Y,R} = 0_{n\times n}$$
$$\tilde{\Sigma}_{Y,\varepsilon} = \tilde{\Psi}\ \Sigma_{Y,u} \tilde{\Psi}'$$

where $A_0^{Y,R} = (A_0^Y + R_Y A_f^Y)$, $A_{b,1}^{Y,R} = (A_b^Y + R_Y A_0^Y)$, $A_{b,2}^{Y,R} = -R_Y A_b^Y$.

Testable implications: summary

- To sum up:
  1. Z_t embodies the testable cointegration rank and cointegration matrix properties of the NK-DSGE model $\longrightarrow$ LR1 & LR2;
  2. we can map Z_t to Y_t (**if the long-run implications are met**);
  3. **we can test the short run CER** that the NK-DSGE system places on the VAR solution for Y_t $\longrightarrow$ LR3.

Testing strategy: LR1->LR2->LR3. The case of observables

- H_0: the NK-DSGE model is 'true'.
- Consider the **finite-order** VAR model for Z_t :

$$Z_t = \sum_{j=1}^{\ell} P_j Z_{t-j} + \mu d_t + \xi_t \quad , \quad \xi_t \sim \mathrm{WN}(0_{n\times 1}, \Sigma_\xi)$$

- and its **error-correction** counterpart

$$\Delta Z_t = \alpha\beta' Z_{t-1} + \sum_{j=1}^{\ell-1} \Theta_j \Delta Z_{t-1} + \mu d_t + \xi_t \quad , \quad \xi_t \sim \mathrm{WN}(0_{n\times 1}, \Sigma_\xi).$$

Testing strategy: LR1->LR2->LR3. The case of observables

- **LR1 [cointegration rank test]:** Compute the LR cointegration rank test. For instance, test that

  $$r = 3 \text{ in } Z_t{:=}(y_t, \pi_t, i_t, y_t^p)'$$

  Johansen's Trace test. We suggest the 'one-shot' version but we also use the 'sequential' version.
  If the rank is rejected, reject the NK-DSGE model, **otherwise go ahead**.
- **Bootstrap version:** (recursive design) iid bootstrap (Cavaliere *et al.* 2012, Ecta).

Testing strategy: LR1->LR2->LR3. The case of observables

- **LR2 [Overidentification cointegration restriction test]:** Compute LR test for the over-identification cointegration restrictions

$$\text{Hp:} \quad \beta = \beta_0 = H$$

$$r = 3 \quad , \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ \pi_t \\ i_t \\ y_t^p \end{pmatrix} = \begin{pmatrix} y_t - y_t^p \\ \pi_t \\ i_t \end{pmatrix} \sim \mathrm{I}(0).$$

- If the cointegration structure is rejected, reject the NK-DSGE model, **otherwise go ahead**.
- **Bootstrap version:** Boswijck *et al.* (2013, WP).

Testing strategy: LR1–>LR2–>LR3. The case of observables

- **LR3 [test for the CER]:** test the CER the NK-DSGE model places on the VAR representation for Y_t:

$$Y_t = \tilde{\Phi}_1(\theta) Y_{t-1} + \tilde{\Phi}_2(\theta) Y_{t-2} + \varepsilon_t^Y \ :$$

  1. Estimate VAR unrestrictedly and get LogLikH1 ;
  2. Estimate the VAR **subject to a numerical approximation of the CER** and get LogLikH0;
  3. $LR3_T = -2[\text{LogLikH0 - LogLikH1}]$.

- A grid search + quasi-Newton (BFGS) for θ helps in the log-likelihood maximization (Ox code).
- **Bootstrap version:** adapt Fanelli and Palomba (2011, JAE).
- We accept the NK-DSGE model if all three tests pass.

Testing strategy: LR1->LR2->LR3. The case of unobservables

- Z_t^o subset of Z_t, e.g. $Z_t^o := (y_t, \pi_t, i_t)' := W_t^o$.
- if $Z_t \sim VAR(\ell)$, Z_t^o ∼VARMA-type≈VAR(ℓ^*), ℓ^* relatively 'large';
- **LR1:** Stock and Watson (1988), Lütkepohl and Claessen (1997, JoE), Wagner (2010) "**Cointegration analysis with state-space models**"
- **LR2:** from the VEC for ΔZ_t^o, e.g. $\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} Z_t^o \sim$I(0).
- **LR3**: find the minimal state-space representation (Komunjer and Ng, 2011, Ecta) then apply the **Kalman filter,** see Guerron-Quintana *et al.* (2013, QE).

Asymptotic size

- The overall asymptotic probability of rejecting the NK-DSGE model is given by

$$\psi_\infty = \limsup_{T\to\infty} \psi_T$$

where

$$\psi_T = \mathbb{P}^{H_0}_{1,T}(LR1 \text{ rejects}) + \mathbb{P}^{H_0}_{1,2,T}(LR1 \text{ does not reject ; } LR2 \text{ rejects})$$
$$+ \mathbb{P}^{H_0}_{2,3,T}(LR2 \text{ does not reject ; } LR3 \text{ rejects})$$

- It can be easily proved that

$$\psi_\infty \leq \psi_{1,\infty} + \psi_{2,\infty} + \psi_{3,\infty} = \underset{\text{e.g. } 1\%}{\psi_1} + \underset{\text{e.g. } 2\%}{\psi_2} + \underset{\text{e.g. } 2\%}{\psi_3}$$

What if LR1 rejects ? Hypothesis 1

- The baseline specification of the Benati and Surico's (2009, AER) NK-DSGE model **predicts that the system must be driven by a single stochastic (technology) trend**:

$$r = 3 \quad , \quad \beta_0' Z_t = \begin{pmatrix} y_t - y_t^p \\ \pi_t \\ i_t \end{pmatrix} \sim \mathrm{I}(0).$$

- Suppose now that the tests LR1 and LR2 suggests that $r = \hat{r} = 2$, so that there are $n - \hat{r} = 2$ common stochastic trends:

$$\beta_0'(\nu) Z_t = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & -\nu & 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ \pi_t \\ i_t \\ y_t^p \end{pmatrix} = \begin{pmatrix} y_t - y_t^p \\ i_t - \nu \pi_t \end{pmatrix} \sim \mathrm{I}(0).$$

What if LR1 rejects ?

- To achieve a **balanced error-correction representation** of the system we need: $\nu = 1$, $\varkappa = 1$, $\varphi_\pi = 1$;

$$\tilde{y}_t = \gamma E_t \tilde{y}_{t+1} + (1-\gamma)\tilde{y}_{t-1} + \delta E_t \Delta\pi_{t+1} - \delta(i_t - \pi_t) + \eta_{\tilde{y},t}$$

$$\Delta\pi_t = \varrho E_t \Delta\pi_{t+1} + \left(\frac{\kappa}{1+\varrho}\right)\tilde{y}_t + (1+\varrho)\eta_{\pi,t}$$

$$(i_t - \pi_t) = (1-\rho)\varphi_y \tilde{y}_t - \rho\Delta\pi_t + \rho(i_{t-1} - \pi_{t-1}) + \eta_{i,t}$$

$$\Delta\tilde{y}_t = \Delta y_t + \eta^*_{y^p,t} \quad (\text{or } \Delta y^p_t = \eta_{y^p,t}).$$

Recall that **determinacy is a system property** and the ''standard Taylor principle" does not hold in this' hybrid' specification, see Lubik and Schorfheide (2004 AER) and Fanelli (2012 JoE).

What if LR1 rejects ? Hypothesis 2

- Take from Bekaert, Cho and Moreno (2006, JMCB)

$$\textbf{AD:} \quad y_t = \gamma E_t y_{t+1} + (1-\gamma) y_{t-1} - \delta(i_t - E_t \pi_{t+1}) + \eta_{\tilde{y},t}$$

$$\textbf{NKPC:} \quad \pi_t = \frac{\varrho}{1+\varrho\varkappa} E_t \pi_{t+1} + \frac{\varkappa}{1+\varrho\varkappa} \pi_{t-1} + \kappa \tilde{y}_t + \eta_{\pi,t}$$

$$\textbf{New rule:} \quad i_t = \rho i_{t-1} + (1-\rho)\varphi_\pi (E_t \pi_{t+1} - \pi_t^*) + (1-\rho)\varphi_y \tilde{y}_t + \eta_{i,t}$$

$$\pi_t^* = \frac{\varrho}{1+\varrho\varpi} E_t \pi_{t+1}^* + \frac{\varpi}{1+\varrho\varpi} \pi_{t-1}^* + (1 - \frac{\varrho}{1+\varrho\varpi} - \frac{\varpi}{1+\varrho\varpi}) \pi_t + \epsilon_{\pi^*,t}$$

- ϖ measures the extent to which the monetary authority anchors the inflation target to $\pi_t^{LR} = (1-\varrho)\sum_{j=0}^{\infty} \varrho^j E_t \pi_{t+j}$:
  1. with $\varpi = 0$, $\pi_t^* = \pi_t^{LR} + \epsilon_{\pi^*,t}$;
  2. with $\varpi = 1$, $\pi_t^* = \pi_{t-1}^* + \epsilon_{\pi^*,t}$, **stochastic trend !**

What if LR1 rejects ?

- If the approximation $\pi_t^* = \pi_{t-1}^* + \epsilon_{\pi^*,t}$ is **reasonable on the sample under** investigation, the

$$r = 2 \quad (n - r = 2)$$

$$\beta_0' Z_t = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} y_t \\ \pi_t \\ i_t \\ y_t^p \end{pmatrix} = \begin{pmatrix} y_t - y_t^p \\ i_t \end{pmatrix} \sim I(0).$$

Monte Carlo

- Setup: Benati and Surico (2009, AER)

$$\begin{aligned}
\tilde{y}_t &= \gamma E_t \tilde{y}_{t+1} + (1-\gamma)\tilde{y}_{t-1} - \delta(i_t - E_t \pi_{t+1}) + \eta_{\tilde{y},t} \\
\pi_t &= \omega_f E_t \pi_{t+1} + \omega_b \pi_{t-1} + \kappa \tilde{y}_t + \eta_{\pi,t} \\
i_t &= \rho i_{t-1} + (1-\rho)(\varphi_\pi \pi_t + \varphi_y \tilde{y}_t) + \eta_{i,t} \\
y_t^p &= y_{t-1}^p + \eta_t^{y^p}
\end{aligned}$$

$$\begin{pmatrix} \eta_{\tilde{y}} \\ \eta_\pi \\ \eta_i \\ \eta^{y^p} \end{pmatrix}_t = \begin{pmatrix} \rho_{\tilde{y}} & 0 & 0 & 0 \\ 0 & \rho_\pi & 0 & 0 \\ 0 & 0 & \rho_i & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \eta_{\tilde{y}} \\ \eta_\pi \\ \eta_i \\ \eta^{y^p} \end{pmatrix}_{t-1} + u_t^W \quad , \quad \Sigma_{W,u} = diag(\sigma^2_{\tilde{y}}, \sigma^2_\pi, \sigma^2_i, \sigma^2_{y^p})$$

- Parameters are calibrated from the estimated obtained by Benati and Surico (2009, AER) **on U.S. data, 'Great Moderation' sample**.

Monte Carlo results: finite sample size

H_0: NK-DSGE model is 'true'

	Empirical size	
Tests	$T = 100$	$T = 200$
$LR1_{\psi_1=0.01}$ ($r = 3$ vs I(0)):	0.006 (*0.006*)	0.009 (*0.009*)
$LR2_{\psi_2=0.02}$ ($\beta_0 = H \mid LR1$):	0.072(*0.022*)	0.043(*0.022*)
$LR3_{\psi_3=0.02}$ (CER $\mid$ $LR2$):	0.028 (*0.028*)	0.019 (*0.019*)
Overall rejection freq. $\hat{\psi} = \sum_{i=1}^{3} \hat{\psi}_i$	0.106 (*0.036*)	0.071 (*0.050*)

M=1000 simulated DGPs, B=399 for LR1 and LR2 and B=99 for LR3

Monte Carlo results: finite sample size with sequential Trace test in first step

H_0: NK-DSGE model is ‘true’

$LR1_{\psi_1=0.01,\text{seq}}$	$T = 100$	$T = 200$
$r = 0$	0.010 (*0.036*)	0 (*0*)
$r = 1$	0.353 (*0.445*)	0 (*0.001*)
$r = 2$	0.539 (*0.437*)	0.276 (*0.326*)
$r = 3$	0.094 (*0.078*)	0.712 (*0.663*)
$r = 4$ (stationary)	0.004 (*0.004*)	0.012 (*0.010*)
$LR1_{\psi_1=0.01}$ $(r = 3$ vs I(0)):	0.004 (*0.004*)	0.012 (*0.010*)
$LR2_{\psi_2=0.02}$ $(\beta_0 = H \mid LR1_{\text{seq}})$:	0.140 (*0.059*)	0.044 (*0.022*)
$LR3_{\psi_3=0.02}$ (CER $\mid$ $LR2$):	0.049 (*0.014*)	0.022 (*0.011*)

Illustration: results on U.S. quartely data, Great Moderation

- We work by assuming that potential output is proxied by the CBO official measure; $Z_t = (y_t, \pi_t, i_t, y_t^p)'$ **is observable.**
- 'Great Moderation' sample **1985.Q1-2008.Q2**
  Castelnuovo, Fanelli (2013, WP): do not reject determinacy.
- The 'LR1→LR2→LR3' testing strategy is applied at the **5%** nominal level of significance, taking 1%(LR1)-2%(LR2)-2%(LR3).

Results on U.S. quartely data

U.S. quarterly data **T=95**				
Tests		Trace	Asymptotic	Boot.
$LR1_{\text{seq}}$:	$r = 0$	107.10	0.000	0.000
	$r = 1$	32.33	0.024	0.071
	$r = 2$	15.07	0.056	0.248
$LR1_{\psi_1=0.01}$ ($r = 3$ vs I(0)) :	$r = 3$	2.43	0.119	0.491
$LR2_{\psi_2=0.02}$ ($\beta_0 = H \mid LR1$) :		11.665	0.009	0.040
$LR3_{\psi_3=0.02}$ (CER $\mid LR2$) :		17.94	0.022	0.80

Castelnuovo and Fanelli (2013, WP), idnetification-robust FIML inference

Results on U.S. quartely data

ML estimates of structural parameters

Parameters θ_s:	Interpretation	ML	90% IR-CIs
κ	NKPC, slope	0.083 (0.022)	0.048 (0.04 , 0.098)
ρ	Policy rule, smoothing term	0.573 (0.358)	0.67 (0.57 , 0.70)
$\varphi_{\widetilde{y}}$	Policy rule, react. to out. gap	0.073 (1.145)	0.92 (0.72 , 0.98)
φ_{π}	Policy rule, react. to inflation	5.37 (2.47)	5.44 (2.32 , 5.45)
ρ_i	Policy rule, disturbance persist.	0.810 (0.451)	0.79 (0.73 , 0.81)

See Castelnuovo and Fanelli (2013, WP) for **identification-robust** inference in this framework.

Remarks for practitioners

1. A frequentist testing-based evaluation approach **does not necessarily lead one to reject** the NK-DSGE model.
2. The empirical evaluation of a NK-DSGE model should be carried out by considering the **long run and short run restrictions JOINTLY**.
3. Bootstrap refinements in small samples are a good idea. **We reject too often because of small sample issues !** $LR2_{bootstrap} \rightarrow LR3_{bootstrap}$ is a reasonable idea when you have T=100 observations.
4. 'New' paradigms are now necessary, e.g. ZLB, fiscal/monetary policy mix; connection with financial markets.

Future extensions

1. Bootstrap evaluation of NK-DSGE models when we have a **state-space representation.**
2. **Wild bootstrap ??** to tackle heteroskedasticity of unknown form$\longrightarrow$ any advantage for IRFs confidence bands ?
3. How to deal with **medium/big systems** ?