Notes on Maximum Likelihood Estimation of Dynamic General Equilibrium Models

A cookbook recipe for empirical analysis of DGE models.

1. Specify preferences, production functions, budget constraints, etc...
2. Take first-order conditions.
3. Solve for steady state.
4. Log-linearize questions describing the equilibrium.
5. System reduction. Write evolution of the economy in state space form.
6. Estimation.

References.

- Lectures: 2008 NBER Summer Institute (http://www.nber.org/minicourse_2008.html) by James Stock and Mark Watson.
- Readings:
  - King et al. (2002) "Production, Growth and Business Cycles: Technical Appendix." (recipe steps 1-5)
  - Uhlig: "A toolkit for analyzing nonlinear dynamic stochastic models easily" (recipe steps 1-5)
  - Fabio Canova: *Methods for Applied Macroeconomic Research*

Specify preferences, production functions, budget constraints,

- Representative consumer maximizing:

$$\mathbb{E}\sum \beta^t \left[\frac{b}{e}\log\left[ac_{mt}^e + (1-a)\,c_{nt}^e\right] + (1-b)\log\left(1-h_{mt}-h_{nt}\right)\right]$$

  - $e \equiv \frac{\sigma-1}{\sigma}$ related to elasticity of sub.
- Production in home and market sectors:

$$y_m = s_m k_m^\theta h_m^{1-\theta} \quad ; \quad y_n = s_n k_n^\eta h_n^{1-\eta}$$

$$y_m = c_m + \underbrace{i_m}_{k_{t+1}-(1-\delta)k_t} \quad ; \quad y_n = c_n$$

$$k = k_n + k_m$$

- Evolution of productivity:

$$s_i = \rho \log s_{i,t-1} + \varepsilon_{it}$$

  - $\gamma \equiv$ correlation between ε_{mt} and ε_{nt}

Take first-order conditions

Market work hours

$$\frac{1-b}{1-h_{mt}-h_{nt}} = abs_{mt}\left(1-\theta\right)\left(\frac{k_t}{h_{mt}}\right)^{\theta}\frac{\left(c_{mt}\right)^{\sigma-1}}{ac_{mt}^{\sigma}+\left(1-a\right)c_{nt}^{\sigma}}$$

Home work hours

$$\frac{1-b}{1-h_{mt}-h_{nt}} = \left(1-a\right)bs_{nt}\frac{\left(c_{nt}\right)^{\sigma-1}}{ac_{mt}^{\sigma}+\left(1-a\right)c_{nt}^{\sigma}}$$

Euler equation:

$$\frac{\left(c_{mt}\right)^{\sigma-1}}{ac_{mt}^{\sigma}+\left(1-a\right)c_{nt}^{\sigma}} = \beta\frac{\left(c_{m,t+1}\right)^{\sigma-1}}{ac_{m,t+1}^{\sigma}+\left(1-a\right)c_{n,t+1}^{\sigma}} \times\left[s_{m,t+1}\theta\left(\frac{k_{t+1}}{h_{n,t+1}}\right)^{\theta-1}+1-\delta\right]$$

Production functions and market clearing:

$$y_t = c_{mt}+i_t; \quad y_t = s_{mt}\left(h_{mt}\right)^{1-\theta}\left(k_t\right)^{\theta}\ ;\ c_{nt} = s_{nt}h_{nt}$$

Log linearize

Write $\hat{z}$ as the percent deviation of the z from its steady state value.

Market work hours

$$\frac{s_{h_m}\hat{h}_{mt} + s_{h_n}\hat{h}_{nt}}{1 - s_{h_m} - s_{h_n}} = \hat{s}_{mt} + \theta\left(\hat{k}_t - \hat{h}_{mt}\right) + \left[\sigma\left(1-a\right) - 1\right]\hat{c}_{mt} - \sigma\left(1-a\right)\hat{c}_{nt}$$

Home work hours

$$\frac{s_{h_m}\hat{h}_{mt} + s_{h_n}\hat{h}_{nt}}{1 - s_{h_m} - s_{h_n}} = \hat{s}_{nt} - \sigma a\hat{c}_{mt} + \left[a\sigma - 1\right]\hat{c}_{nt}$$

Euler equation

$$\begin{aligned}&\left[\sigma\left(1-a\right) - 1\right]\hat{c}_{mt} - \sigma\left(1-a\right)\hat{c}_{nt} \\ &= \left[\sigma\left(1-a\right) - 1\right]\hat{c}_{m,t+1} - \sigma\left(1-a\right)\hat{c}_{n,t+1} \\ &+ \left(1 - \beta + \beta\delta\right)\left[\hat{s}_{m,t+1} + \left(\theta - 1\right)\left(\hat{k}_{t+1} - \hat{h}_{m,t+1}\right)\right]\end{aligned}$$

Production functions and market clearing:

$$\hat{y}_t = s_{i/y}\hat{\imath}_t + s_{c/y}\hat{c}_{mt}; \quad \hat{y}_t = \theta\hat{k}_t + \left(1-\theta\right)\hat{h}_{mt} + \hat{s}_{mt}\ ; \ \hat{c}_{nt} = \hat{s}_{nt} + \hat{h}_{nt}$$

Write evolution of the economy in state space form

Using the methods in King et al. (2002) and Canova (Chapter 2), we can write out the economy as:

$$\begin{pmatrix} \hat{k}_{t+1} \\ \hat{s}_{n,t+1} \\ \hat{s}_{m,t+1} \end{pmatrix} = \underbrace{\begin{pmatrix} F_1 & F_2 & F_3 \\ 0 & \rho & 0 \\ 0 & 0 & \rho \end{pmatrix}}_{F} \begin{pmatrix} \hat{k}_t \\ \hat{s}_{n,t} \\ \hat{s}_{m,t} \end{pmatrix} + \underbrace{\begin{pmatrix} 0 & 0 \\ \sigma_n^2 & \gamma\sigma_n\sigma_m \\ \gamma\sigma_n\sigma_m & \sigma_m^2 \end{pmatrix} \begin{pmatrix} \varepsilon_{nt} \\ \varepsilon_{mt} \end{pmatrix}}_{\eta_t}$$

Knowing the joint evolution of $\hat{k}$, $\hat{s}_{nt}$, $\hat{s}_{mt}$, we can write...

$$\begin{pmatrix} \hat{h}_{nt} \\ \hat{h}_{mt} \end{pmatrix} = \underbrace{\begin{pmatrix} H_1 & H_2 & H_3 \\ H_4 & H_5 & H_6 \end{pmatrix}}_{H} \begin{pmatrix} \hat{k}_t \\ \hat{s}_{tn} \\ \hat{s}_{tm} \end{pmatrix} + \underbrace{\begin{pmatrix} \sigma_{\text{mes}}^2 & 0 \\ 0 & \sigma_{\text{mes}}^2 \end{pmatrix}}_{\Sigma_{\text{mes}}} \begin{pmatrix} \varepsilon_{nt}^{\text{mes}} \\ \varepsilon_{mt}^{\text{mes}} \end{pmatrix}$$

F and H are functions of the model's parameters $(\sigma, \delta, \beta, a, b, \theta)$

Notation and set-up (from Mark Watson's lecture)

Write

- $s_t = \left\{ \hat{s}_{n,t},\ \hat{s}_{m,t},\ \hat{k}_t \ \right\}$, $S_t = \{s_1, s_2, ... s_t\}$,
- $y_t = \left\{ \hat{h}_{nt},\ \hat{h}_{mt} \ \right\}, Y_t = \{y_1, y_2, ... y_t\}$.

Prediction of s_t given Y_{t-1}

$$\begin{aligned} f\left(s_t | Y_{t-1}\right) &= \int f\left(s_t,\ s_{t-1} | Y_{t-1}\right) ds_{t-1} \\ &= \int f\left(s_t |\ s_{t-1},\ Y_{t-1}\right) f\left(s_{t-1} | Y_{t-1}\right) ds_{t-1} \\ &= \int f\left(s_t |\ s_{t-1}\right) f\left(s_{t-1} | Y_{t-1}\right) ds_{t-1} \end{aligned}$$

Prediction of y_t given Y_{t-1}

$$f\left(y_t | Y_{t-1}\right) = \int f\left(y_t | s_t\right) f\left(s_t | Y_{t-1}\right) ds_t$$

Notation and set-up (from Mark Watson's lecture)

Write

- $s_t = \left\{ \hat{s}_{n,t},\ \hat{s}_{m,t},\ \hat{k}_t \ \right\},\ S_t = \{s_1, s_2, ... s_t\},$
- $y_t = \left\{ \hat{h}_{nt},\ \hat{h}_{mt} \ \right\}, Y_t = \{y_1, y_2, ... y_t\}.$

Prediction ("filtering") of s_t given Y_t

$$f\left(s_t | Y_t\right) = f\left(s_t | y_t, Y_{t-1}\right) = \frac{f\left(y_t | s_t, Y_{t-1}\right) f\left(s_t | Y_{t-1}\right)}{f\left(y_t | Y_{t-1}\right)}$$

$$= \frac{f\left(y_t | s_t\right) f\left(s_t | Y_{t-1}\right)}{f\left(y_t | Y_{t-1}\right)}$$

Useful facts about multivariate normal distribution (from Mark Watson's lecture)

Suppose

$$\begin{pmatrix} u \\ v \end{pmatrix} \text{ is } \mathcal{N}\left(\begin{pmatrix} \mu_u \\ \mu_v \end{pmatrix}, \begin{pmatrix} \Sigma_{\mathrm{uu}} & \Sigma_{\mathrm{uv}} \\ \Sigma_{\mathrm{vu}} & \Sigma_{\mathrm{vv}} \end{pmatrix}\right)$$

Then

$$u|v \text{ is } \mathcal{N}\left(\mu_{u|v}, \Sigma_{u|v}\right), \text{ where}$$

$$\mu_{\mathrm{u|v}} = \mu_{\mathrm{u}} + \Sigma_{\mathrm{uv}} \left(\Sigma_{\mathrm{vv}}\right)^{-1} \left(v - \mu_{\mathrm{v}}\right)$$

$$\Sigma_{\mathrm{u|v}} = \Sigma_{\mathrm{uu}} - \Sigma_{\mathrm{uv}} \left(\Sigma_{\mathrm{vv}}\right)^{-1} \Sigma_{\mathrm{vu}}$$

The Kalman Filter (from Mark Watson's lecture)

Prediction of s_t given Y_t

$$f\left(s_t|Y_t\right) = f\left(s_t|y_t, Y_{t-1}\right) = \frac{f\left(y_t|s_t, Y_{t-1}\right) f\left(s_t|Y_{t-1}\right)}{f\left(y_t|Y_{t-1}\right)}$$
$$= \frac{f\left(y_t|s_t\right) f\left(s_t|Y_{t-1}\right)}{f\left(y_t|Y_{t-1}\right)}$$

If $s_{t-1}|Y_{t-1}$ is Normal$(s_{t-1|t-1}, P_{t-1|t-1})$.
Then

$$\begin{pmatrix} s_t \\ y_t \end{pmatrix} |Y_{t-1} \text{ is } \mathcal{N}\left(\begin{pmatrix} s_{t|t-1} \\ \mu_{t|t-1} \end{pmatrix}, \begin{pmatrix} P_{t|t-1} & P_{t|t-1}H' \\ HP_{t|t-1} & HP_{t|t-1}H' + \Sigma_{\text{mes}} \end{pmatrix}\right)$$

But how to compute $s_{t|t-1}$, $\mu_{t|t-1}$, $P_{t|t-1}$?

The Kalman Filter (from Mark Watson's lecture)

Model:

$$y_t = Hs_t + \varepsilon_{\text{mes}}; \quad s_t = Fs_{t-1} + \eta_t$$

- Prediction of y_t and mean squared error of y_t, given s_{t-1}
  - $\mathbb{E}[s_t|s_{t-1}] = Fs_{t-1|t-1} \equiv s_{t|t-1}$
  - $\text{Var}[s_t|s_{t-1}] = FP_{t-1|t-1}F' + \Sigma_\eta \equiv P_{t|t-1}$
  - $\mathbb{E}[y_t|s_{t-1}] = Hs_{t|t-1} \equiv \mu_{t|t-1}$
  - $\text{Var}[y_t|s_{t-1}] = HP_{t|t-1}H' + \Sigma_{\text{mes}} \equiv \Sigma_{t|t-1}$
- Updating (after observing y_t). Use "useful facts" interpreting v as y_t and u as s_t (both conditioned on Y_{t-1})
  - $K_t \equiv P_{t|t-1}H'\Sigma_{t|t-1}^{-1}$
  - $\mathbb{E}[s_t|y_t, s_{t-1}] = s_{t|t-1} + K_t\left(y_t - \mu_{t|t-1}\right) \equiv s_{t|t}$
  - $\text{Var}[s_t|y_t, s_{t-1}]$

$$\begin{aligned} &= P_{t|t-1} - HP_{t|t-1}\left(\Sigma_{t|t-1}\right)^{-1}P'_{t|t-1}H' \\ &= P_{t|t-1} - HP_{t|t-1}\Sigma_{t|t-1}^{-1}P'_{t|t-1}H' \\ &= P_{t|t-1} - P_{t|t-1}H'\Sigma_{t|t-1}^{-1}P'_{t|t-1} = (I - K_t)P_{t|t-1} \equiv P_{t|t} \end{aligned}$$

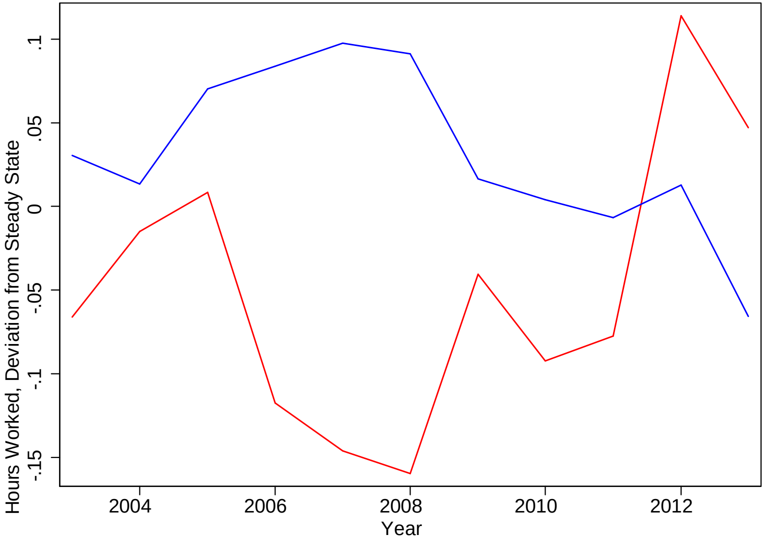
The log likelihood

- Write the log-likelihood as:

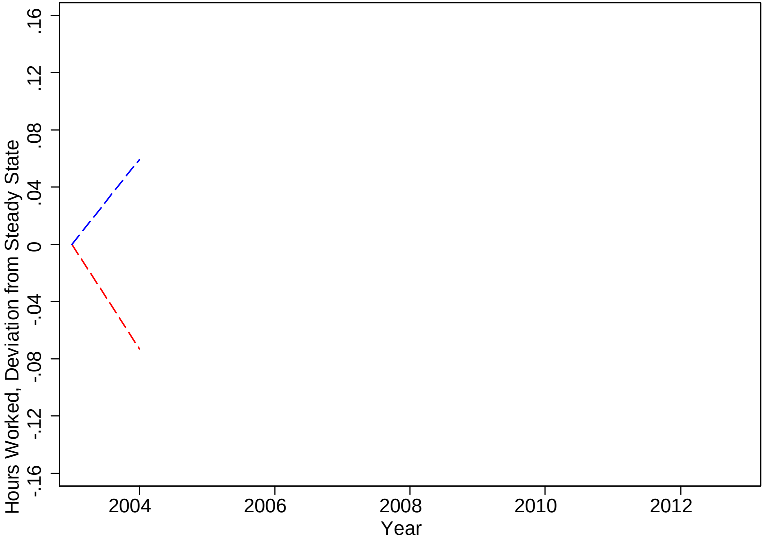
$$\log f\left(Y_t\right) = \sum_{t=1}^{T} \log f\left(y_t | S_{t-1}\right)$$
$$= -\frac{2T}{2}\log\left(2\pi\right) - \frac{1}{2}\sum_{t=1}^{T}\{\log|\Sigma_{t|t-1}^{-1}|$$
$$+ \left(y_t - \mu_{t|t-1}\right)' \Sigma_{t|t-1}^{-1} \left(y_t - \mu_{t|t-1}\right)\}$$

- Both $\Sigma_{t|t-1}$ and $\mu_{t|t-1}$ are products of the Kalman filter (functions of H, F, Σ_{mes}, and $\Sigma_s \equiv \begin{pmatrix} \sigma_n^2 & \gamma\sigma_n\sigma_m \\ \gamma\sigma_n\sigma_m & \sigma_m^2 \end{pmatrix}$)

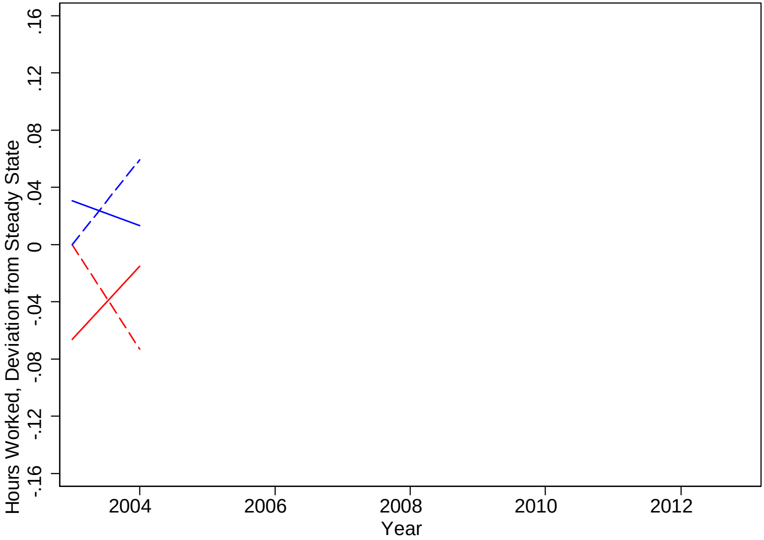
Raw Data: Hours worked in Florida



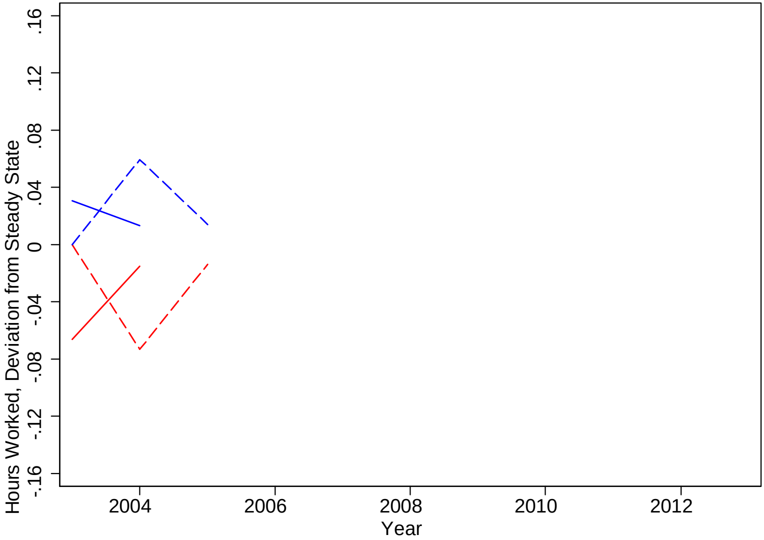
Hours Worked and its Predicted Values



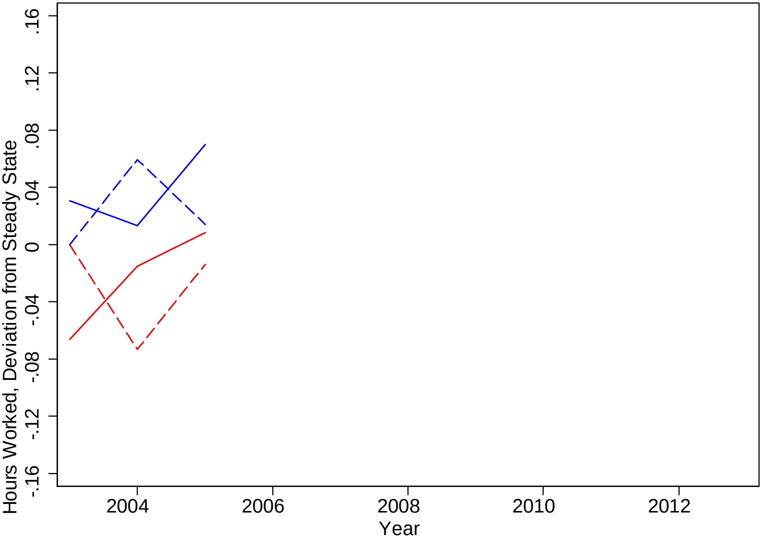
Hours Worked and its Predicted Values



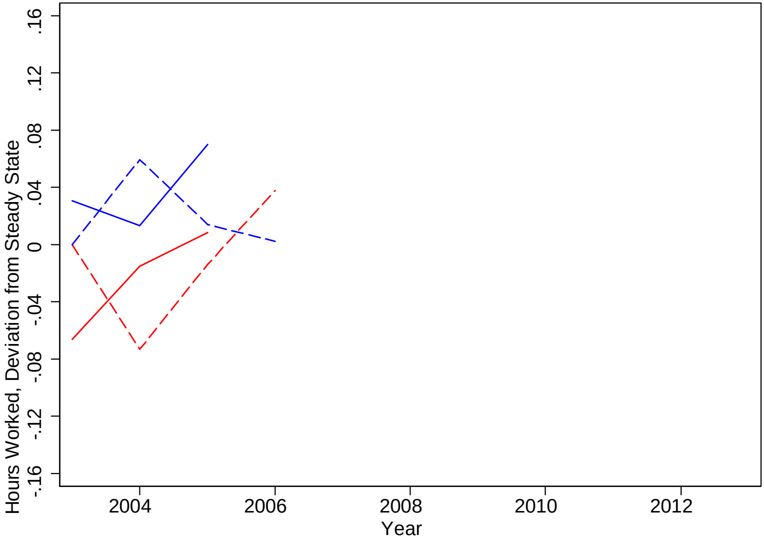
Hours Worked and its Predicted Values



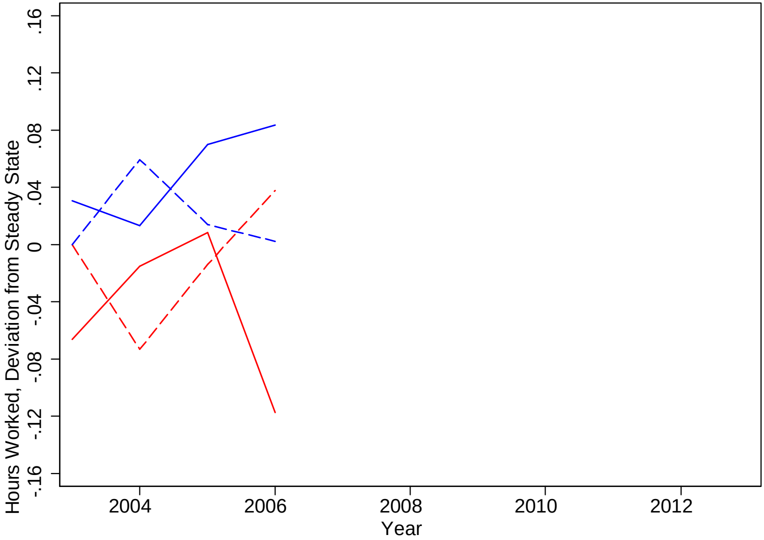
Hours Worked and its Predicted Values



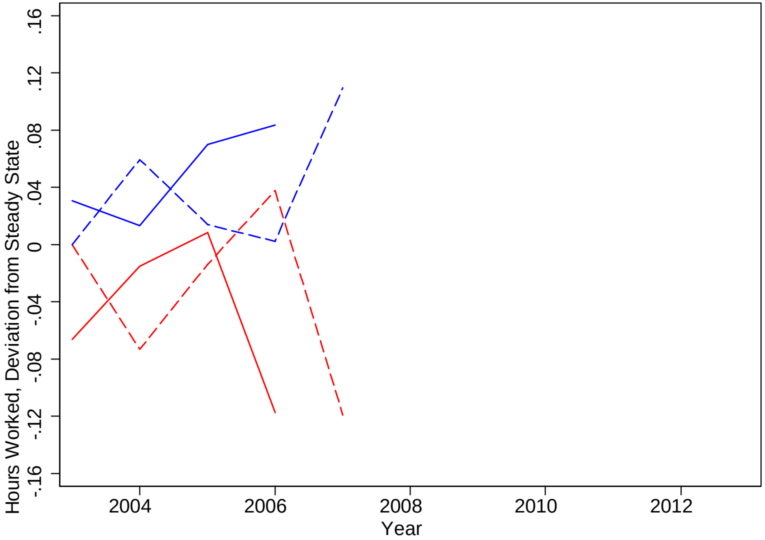
Hours Worked and its Predicted Values



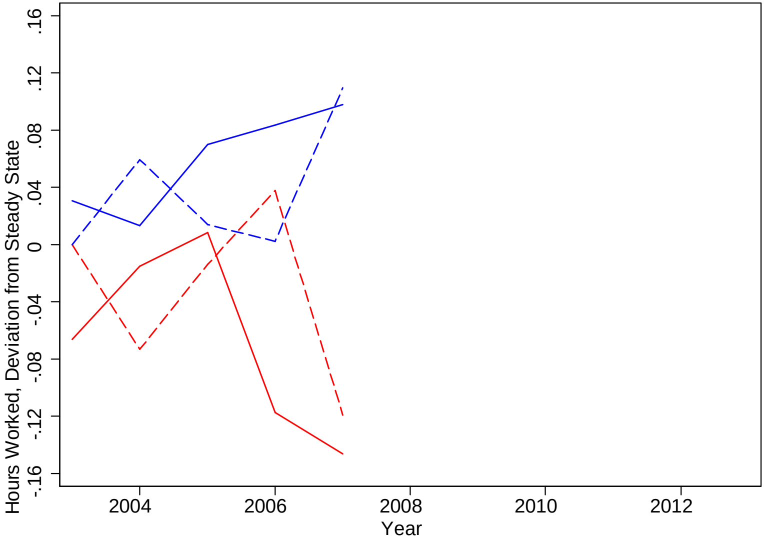
Hours Worked and its Predicted Values



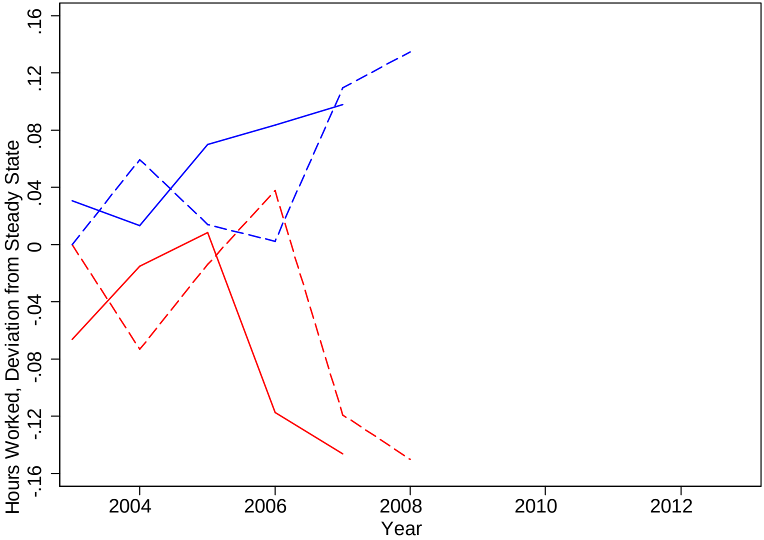
Hours Worked and its Predicted Values



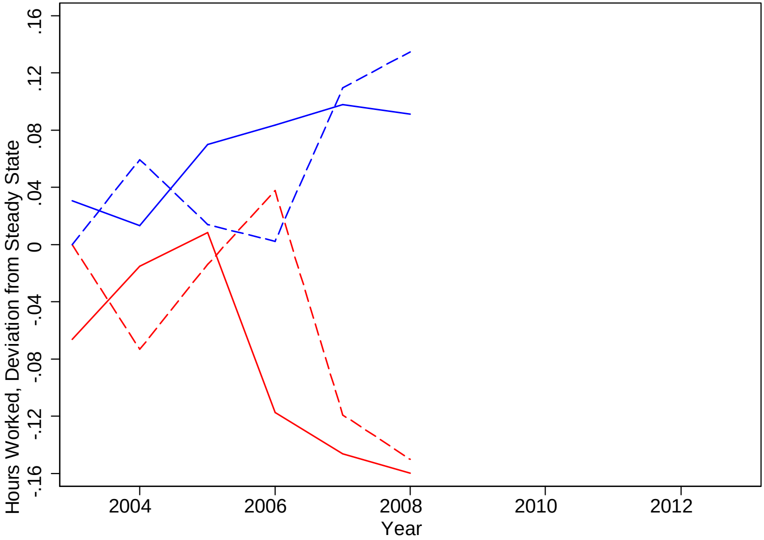
Hours Worked and its Predicted Values



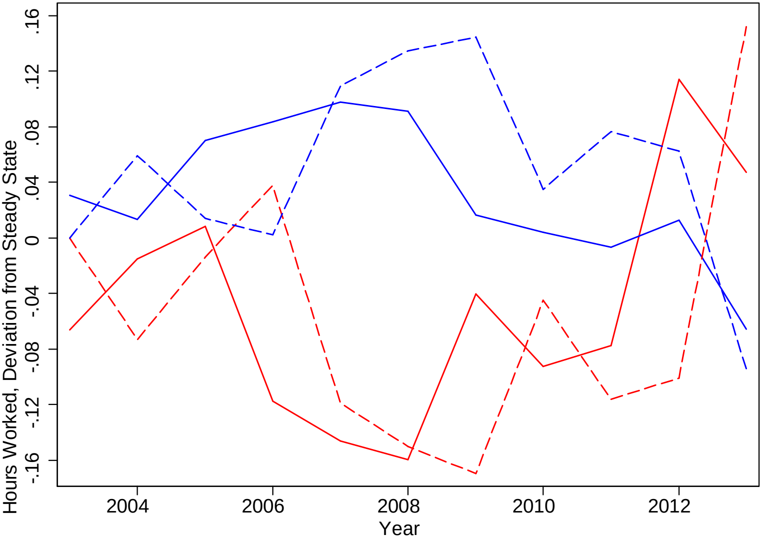
Hours Worked and its Predicted Values



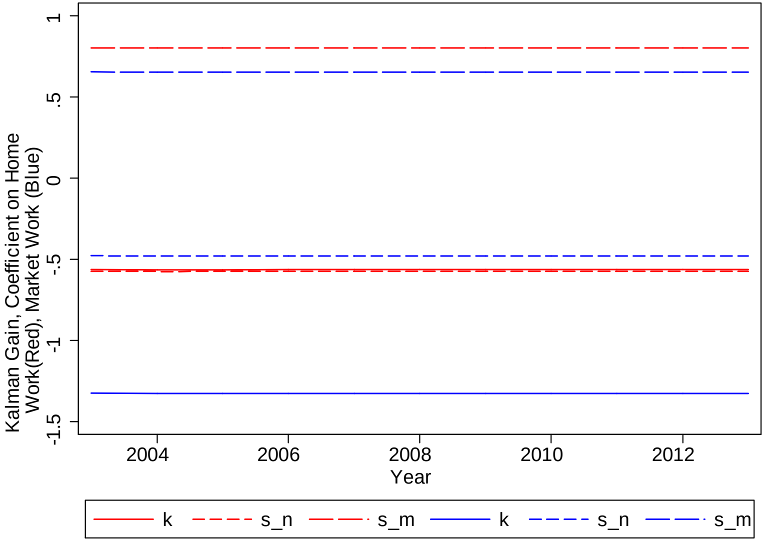
Hours Worked and its Predicted Values



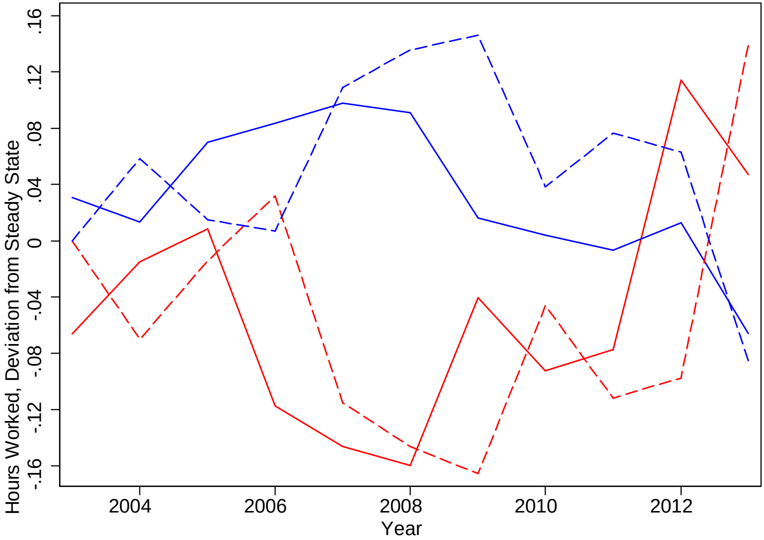
Hours Worked and its Predicted Values



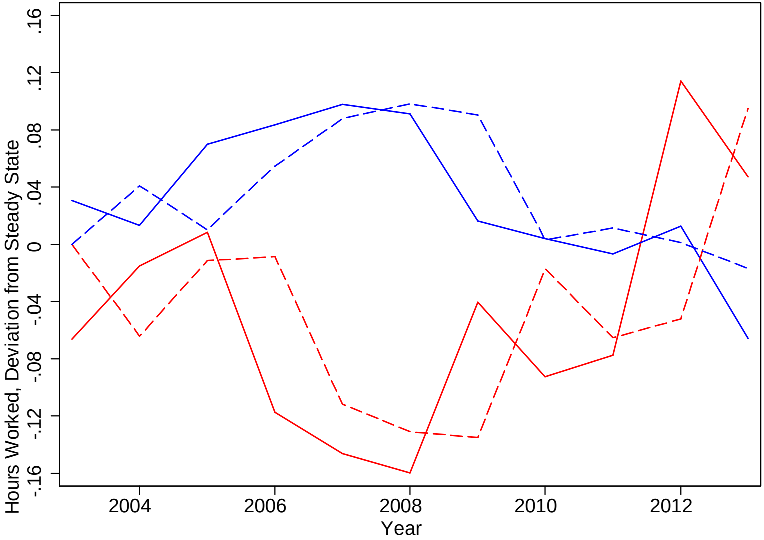
The Kalman Gain



Higher value of σ



Lower value of σ



Other Issues

- Kalman Smoothing
  - There is unused, potentially useful, information left on the table after the Kalman Filtering exercise

$$
\begin{aligned}
f(s_t|Y_T) &= \int f(s_t, s_{t+1}|Y_T)\, ds_{t+1} \\
&= \int f(s_t|s_{t+1},\ Y_T)\, f(s_{t+1}|Y_T)\, ds_{t+1} \\
&= \int f(s_t|s_{t+1},\ Y_t)\, f(s_{t+1}|Y_T)\, ds_{t+1} \\
&= \int \frac{f(s_{t+1}|s_t)\, f(s_t|Y_t)}{f(s_{t+1}|Y_t)} f(s_{t+1}|Y_T)\, ds_{t+1} \\
&= f(s_t|Y_t) \cdot \int f(s_{t+1}|s_t)\, \frac{f(s_{t+1}|Y_T)}{f(s_{t+1}|Y_t)} ds_{t+1}
\end{aligned}
$$

Other Issues

- Dependence on initial conditions
  - Potentially the estimate of $\mathbb{E}[s_t|s_{t-1}]$ is sensitive to our initial guess of $\mathbb{E}[s_1|s_0]$. Solution: Exclude first few periods from the log likelihood.
- Local identification (Iskrev 2010)
  - Because of linear-Gaussian assumption, all information about the parameters are embedded within the first and second moments of the $\{y_t\}$.
  - Check for identification: Compute the Jacobian matrix of the first and second moments of $\{y_t\}$ with respect to the parameters we are trying to estimate $(\sigma, \gamma, \sigma_{\varepsilon_n}, \sigma_{\varepsilon_m})$
- How does the de-trending method affect the estimated parameters?