EABCN Training School: Monetary-Fiscal Policy Interactions

Lecture 6. Foresight: Theory and Econometrics

Eric M. Leeper

Indiana University

September 2010

The Messages

- Draws on Leeper, Walker, Yang (2008,2009)
- It's about the problems that arise whenever agents' information sets do not align with the econometrician's information set
- Implications
  - standard econometric tools will fail
  - inferences can be very misleading
  - no quick fix to the problem
- Need to model information flows in a serious way
- This lecture will use *fiscal* foresight as the example
- Applies to rapidly growing literature on technology foresight

Fiscal Foresight: The Problem

- Legislative & implementation lags $\Rightarrow$ agents know changes in future tax rates before they are effective
- Agents act on the information before fiscal variables move
- Hard to build agents' information into econometric work
- Cannot extract "news" from current & past fiscal variables

Implications of Fiscal Foresight

- Agents' & econometrician's information sets misaligned
- Conventional econometric methods can fail to identify "news" correctly
- Fiscal foresight can create a non-invertible moving average in equilibrium data (first shown by Yang)
- Usual econometric tools can yield false inferences
  - impulse response functions, variance decompositions
  - tests of cross-equation restrictions
  - tests of present-value relations
- *All* identifications convoluted: fiscal & non-fiscal
  - dynamics wrong
  - shocks confounded

A Bit of Formalism

- ε_t: vector of exogenous shocks agents observe
- $\varepsilon_{\tau,t}$: tax component
- $\Omega_t = \text{span}\{\varepsilon_t, \varepsilon_{t-1}, \ldots\}$ agents' information set
- ε_t^*: vector of exogenous shocks econometrician identifies
- $\varepsilon_{\tau,t}^*$: tax component
- $\Omega_t^* = \text{span}\{\varepsilon_t^*, \varepsilon_{t-1}^*, \ldots\}$: econometrician's information set
- Fiscal foresight $\Rightarrow \Omega_t^*$ strictly smaller than Ω_t

No Consensus in Empirical Work

An anticipated tax cut

- has little or no effect [Poterba-Summers, Blanchard-Perotti, Romer-Romer]
- is expansionary in the short run [Mountford-Uhlig]
- is strongly contractionary in the short run [Mertens-Ravn, Branson-Fraga-Johnson, House-Shapiro]

Sources of the diverse results

- invalid instruments for fiscal foresight
- no modeling of fiscal behavior that gives rise to foresight (information flows)
- non-invertibility not directly confronted

Anecdotal Evidence of Fiscal Foresight

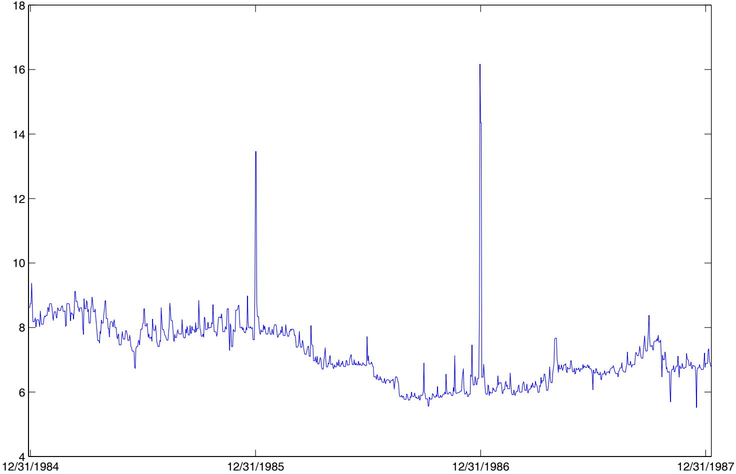
- Large public finance literature
- Branson, Fraga, & Johnson: Economic Recovery Tax Act of 1981 “purely anticipatory recession”
- House & Shapiro: jobless recovery of 2001—tax phase-ins induced production delays
- Ramey: War dummies predict defense spending

Capital Gains in Anticipation of TRA86

Year	Long-term	Total
1984	134.1	135.0
1985	164.9	167.0
1986	**315.7**	**322.2**
1987	—	137.4
1988	—	153.8
1989	—	145.6
1990	—	113.2

Capital Gains Realizations in Billions.
Source: Auerbach & Slemrod (1997)

Funds Rate in Anticipation of TRA86



Daily Federal Funds Rate

FUNDS RATE IN ANTICIPATION OF TRA86

"Despite repeated reserve injections, reserve needs persisted...The Treasury balance was enlarged beyond normal levels by heavy sales of nonmarketable debt to tax-exempt authorities which were engaged in large sales in the market to get ahead of proposed restrictive tax legislation."

–FRBNY Quarterly Review Spring 1986

Simple Illustrative Model

- Log preferences
- Inelastic labor supply
- Complete depreciation of capital
- Proportional tax levied against income [$T_t = \tau_t Y_t$]

Equilibrium conditions

$$\frac{1}{C_t} = \alpha\beta \mathbf{E_t}(\mathbf{1} - \boldsymbol{\tau}_{\mathbf{t+1}}) \frac{1}{C_{t+1}} \frac{Y_{t+1}}{K_t}$$

$$C_t + K_t = Y_t = A_t K_{t-1}^{\alpha}$$

Solving the Model

- Log linearize to get a second-order difference equation in k

$$\begin{aligned} & E_t k_{t+1} - \frac{1+\alpha^2\beta(1-\tau)}{\alpha\beta(1-\tau)} k_t + \frac{1}{\beta(1-\tau)} k_{t-1} \\ = & \frac{1-\alpha\beta(1-\tau)}{\alpha\beta(1-\tau)} \left(\frac{\tau}{1-\tau}\right) E_t \hat{\tau}_{t+1} - \frac{1}{\alpha\beta(1-\tau)} a_t \end{aligned}$$

or

$$E_t k_{t+1} - \gamma_0 k_t + \gamma_1 k_{t-1} = \nu_0 \varepsilon_{A,t} + \nu_1 E_t \hat{\tau}_{t+1}$$

with

$$\begin{aligned} \gamma_0 &= \frac{1+\alpha^2\beta(1-\tau)}{\alpha\beta(1-\tau)} > 0, \qquad \gamma_1 = \frac{1}{\beta(1-\tau)} > 0 \\ \nu_0 &= -\frac{1}{\alpha\beta(1-\tau)} < 0, \qquad \nu_1 = \frac{1-\alpha\beta(1-\tau)}{\alpha\beta(1-\tau)} \left(\frac{\tau}{1-\tau}\right) > 0 \end{aligned}$$

Solving the Model

- Solution satisfies saddlepath property; write the difference equation as

$$(B^{-2} - \gamma_0 B^{-1} + \gamma_1) E_t k_{t-1} = E_t z_t$$

  where $B^{-j} E_{t-1} k_t = E_{t-1} k_{t+j}$ for integer j and $z_t \equiv \nu_0 \varepsilon_{A,t} + \nu_1 E_t \hat{\tau}_{t+1}$
- Factor the quadratic as

$$(\lambda_1 - B^{-1})(\lambda_2 - B^{-1}) E_t k_{t-1} = E_t z_t$$

  so that $\gamma_1 = \lambda_1 \lambda_2$ and $\gamma_0 = \lambda_1 + \lambda_2$. Note that $\lambda_1 > 0$ and $\lambda_2 > 0$
- Select $\lambda_1 < 1$ and $\lambda_2 = [\beta(1-\tau)\lambda_1]^{-1} > 1$
- Operate on both sides of the equation with $(\lambda_2 - B^{-1})^{-1}$

$$(\lambda_1 - B^{-1}) E_t k_{t-1} = (\lambda_2 - B^{-1})^{-1} E_t z_t$$

Solving the Model

- Now
$$\frac{1}{\lambda_2 - B^{-1}} = \frac{1}{\lambda_2}\frac{1}{1-(1/\lambda_2)B^{-1}}$$
We shall use the facts that $\lambda_2^{-1} = \beta(1-\tau)\lambda_1$ and $[1-(1/\lambda_2)B^{-1}]^{-1} = \sum_{j=0}^{\infty}(\lambda_2 B)^{-j}$ to yield
$$k_t = \lambda_1 k_{t-1} - \beta(1-\tau)\lambda_1 \sum_{i=0}^{\infty}[\beta(1-\tau)\lambda_1]^i E_t z_{t+i}$$

- It turns out that $\lambda_1 = \alpha < 1$ & $\lambda_2 = [\alpha\beta(1-\tau)]^{-1} > 1$
- The solution for k_t is a function of the state at t: k_{t-1} and current and expected exogenous disturbances known at t

Model Solution

Equilibrium capital accumulation obeys

$$k_t = \alpha k_{t-1} + a_t - (1-\theta)\left(\frac{\tau}{1-\tau}\right)\sum_{\mathbf{i=0}}^{\infty} \boldsymbol{\theta}^{\mathbf{i}} \mathbf{E_t} \hat{\boldsymbol{\tau}}_{\tau,\mathbf{t+1+i}}$$

where $\theta = \alpha\beta(1-\tau) < 1$ and a_t is exogenous technology

- θ plays central role in analysis
- Agent uses θ to discount tax *rates* in usual way
- How does agent discount tax *news*?

Fiscal Foresight

- Need to specify information flows
- Start with simple flow
- Tax news arrives q periods before tax rates change

$$\hat{\tau}_t = \varepsilon_{\tau,t-q}$$

- Technology: $i.i.d.$ so $a_t = \varepsilon_{A,t}$
- Agent's information set at t consists of variables dated t and earlier, *including* $i.i.d.$ exogenous shocks

$$\Omega_t = \{\varepsilon_{A,t-j}, \varepsilon_{\tau,t-j}\}_{j=0}^{\infty}$$

- Agent at t has (perfect) knowledge of $\{\hat{\tau}_{t+q}, \hat{\tau}_{t+q-1}, \ldots\}$

Solution: Various Degrees of Foresight

$q = 0$ implies:

$$k_t = \alpha k_{t-1} + \varepsilon_{A,t}$$

$q = 1$ implies:

$$k_t = \alpha k_{t-1} + \varepsilon_{A,t} - (1-\theta)\left(\frac{\tau}{1-\tau}\right)\varepsilon_{\tau,t}$$

$q = 2$ implies:

$$k_t = \alpha k_{t-1} + \varepsilon_{A,t} - (1-\theta)\left(\frac{\tau}{1-\tau}\right)\left\{\varepsilon_{\tau,t-1} + \theta\varepsilon_{\tau,t}\right\}$$

$q = 3$ implies:

$$k_t = \alpha k_{t-1} + \varepsilon_{A,t} - (1-\theta)\left(\frac{\tau}{1-\tau}\right)\left\{\varepsilon_{\tau,t-2} + \theta\varepsilon_{\tau,t-1} + \theta^2\varepsilon_{\tau,t}\right\}$$

Discounting

$$k_t = \alpha k_{t-1} + \varepsilon_{A,t} - (1-\theta)\left(\frac{\tau}{1-\tau}\right)\left\{\boldsymbol{\varepsilon_{\tau,t-1} + \theta\varepsilon_{\tau,t}}\right\}$$

- More *recent* news is discounted (by θ) relative to more distant news. Why?
  - $\varepsilon_{\tau,t-1}$ affects $\hat{\tau}_{t+1}$
  - $\varepsilon_{\tau,t}$ affects $\hat{\tau}_{t+2}$
  - news that affects taxes farther into the future is discounted heaviest
- Foresight introduces moving-average terms into equilibrium
- Dynamic optimization implies more *recent* news more heavily discounted
- Seems perverse and creates econometric problems

VARs and Foresight

- Linear environment & Gaussian random variables $\Rightarrow$ projections equivalent to conditional expectations

- VAR is projection $P[\mathbf{x}_t|\mathbf{x}_{t-1}, \mathbf{x}_{t-2}, ...]$

- If $\mathbf{x}_{t-1}, \mathbf{x}_{t-2}, ...$ reveals agents' information set, then econometrician captures dynamics of economy

- Foresight implies $\mathbf{x}_{t-1}, \mathbf{x}_{t-2}, ...$ will not capture information set of agent in "typical" VAR

- Agent's information set at t: $\Omega_t = \{\varepsilon_{A,t-j}, \varepsilon_{\tau,t-j}\}_{j=0}^{\infty}$

  Econometrician's information set at t: $\Omega_t^* = \{k_{t-j}, z_{t-j}\}_{j=0}^{\infty}$

ASIDE: REPRESENTATION THEORY

- Univariate example: ARMA

$$x_t = \left[\frac{L-\theta}{1-\rho L}\right]\varepsilon_t, \qquad |\rho| \in (0,1), |\theta| \in (0,1) \tag{1}$$

- Given $|\theta| \in (0,1)$, (1) is not invertible $\Rightarrow$ linear space spanned by $\{x_{t-j}\}_{j=0}^{\infty}$ is not equal to the linear space spanned by $\{\varepsilon_{t-j}\}_{j=0}^{\infty}$
- To find space spanned by $\{x_{t-j}\}_{j=0}^{\infty}$, need to factor ARMA

$$x_t = \left[\frac{L-\theta}{1-\rho L}\right]\left[\frac{1-\theta L}{L-\theta}\right]\left[\frac{L-\theta}{1-\theta L}\right]\varepsilon_t$$

$$x_t = \left[\frac{1-\theta L}{1-\rho L}\right]e_t \tag{2}$$

$$e_t = \left[\frac{L-\theta}{1-\theta L}\right]\varepsilon_t \tag{3}$$

- (2) is invertible: current & past x_t span same space as current & past e_t (but *not* ε_t)

Aside: Representation Theory

- This points out that the agent's and the econometrician's information sets are different
- Agent observes $\{\varepsilon_t\}$
- Econometrician observes $\{x_t\}$
- Agent's information set larger than econometrician's
- ε_t called "non-fundamental" shocks because they produce a non-invertible representation
- e_t called "fundamental" shocks because they are associated with invertible (Wold) representation

Aside: Representation Theory

- $(1-\theta L)/(L-\theta)$ is a Blaschke factor
- From (2), $x_t = \left[\frac{1-\theta L}{1-\rho L}\right] e_t$
  - current & past ε_t sufficient for e_t
  - but inverse of Blaschke factor does not possess a valid expansion inside the unit circle in L due to the pole at $L = |\theta|$
  - hence, current & past e_t do not reveal ε_t
- Setting $F = L^{-1}$, Blaschke factor has valid inverse in the forward operator F

$$\left[\frac{F-\theta}{1-\theta F}\right] e_t = \varepsilon_t, \qquad \varepsilon_t = (L^{-1}-\theta)\sum_{j=0}^{\infty} \theta^j e_{t+j}$$

  - hence ε_t carries information about *future* e's (and x's)

Aside: Representation Theory

- A famous example

$$y_t = w_t + 2w_{t-1}, \qquad w_t \sim iidN(0,1)$$

- Define VAR innovations at $a_t = y_t - \hat{E}(y_t | y_{t-1}, y_{t-2}, \ldots)$
- MA representation of y_t is

$$y_t = 2(a_t + (1/2)a_{t-1}), \qquad a_t \sim iidN(0,1)$$

- IRF of first model, $(1, 2, 0, 0, \ldots)$ different from VAR, $(2, 1, 0, 0, \ldots)$: Why?
  - a_t belongs to linear space spanned by y_t:

$$a_t = \frac{1}{2}\sum_{j=0}^{\infty}\left(-\frac{1}{2}\right)^j y_{t-j}$$

  - and w_t belongs to the linear space spanned by future y_t:

$$w_t = \frac{1}{2}\sum_{j=0}^{\infty}\left(-\frac{1}{2}\right)^j y_{t+j+1}$$

Foresight & Non-Invertibility

- Econometrician's conditioning set: $\{k_{t-j}, a_{t-j}\}_{j=0}^{\infty}$
- Will drop a_t from equations since econometrician knows it
- Best case scenario
- Solution with 2-period foresight

$$(1-\alpha L)k_t = -\kappa(L+\theta)\varepsilon_{\tau,t}$$

- Does $\{k_{t-j}\}_{j=0}^{\infty} \equiv \{\varepsilon_{\tau,t-j}\}_{j=0}^{\infty}$?
- Invertibility requires $|\theta| > 1$: yields conv. seq. in past k

$$\left[\frac{1-\alpha L}{\mathbf{1+\boldsymbol{\theta}^{-1}L}}\right]k_t = -\kappa\theta\varepsilon_{\tau,t}$$

But $\theta < 1$, so not invertible in current and past capital
- *Is* invertible in current and *future* capital

$$k_t = (\alpha^{-1}+\theta)k_{t+1} - \theta(\alpha^{-1}+\theta)k_{t+2} + \\ \theta^2(\alpha^{-1}+\theta)k_{t+3} - \cdots + \kappa\varepsilon_{\tau,t}$$

ECONOMETRICIAN'S ESTIMATES: I

- So $\{k_{t-j}\}_{j=0}^{\infty} \neq \{\varepsilon_{\tau,t-j}\}_{j=0}^{\infty}$
- Need to find the econometrician's information set: $\{k_{t-j}\}_{j=0}^{\infty} \equiv$???
- Wold representation for capital

$$
\begin{aligned}
(1-\alpha L)k_t &= \underbrace{-\kappa(L+\theta)\left[\frac{1+\theta L}{L+\theta}\right]}_{} \underbrace{\left[\frac{L+\theta}{1+\theta L}\right]\varepsilon_{\tau,t}}_{} \\
&= \qquad \boldsymbol{-\kappa(1+\theta \mathbf{L})} \qquad\qquad \boldsymbol{\varepsilon^*_{\tau,\mathbf{t}}} \\
&= -(1-\theta)\left(\frac{\tau}{1-\tau}\right)\left\{\theta\varepsilon^*_{\tau,t-1}+\varepsilon^*_{\tau,t}\right\}
\end{aligned}
$$

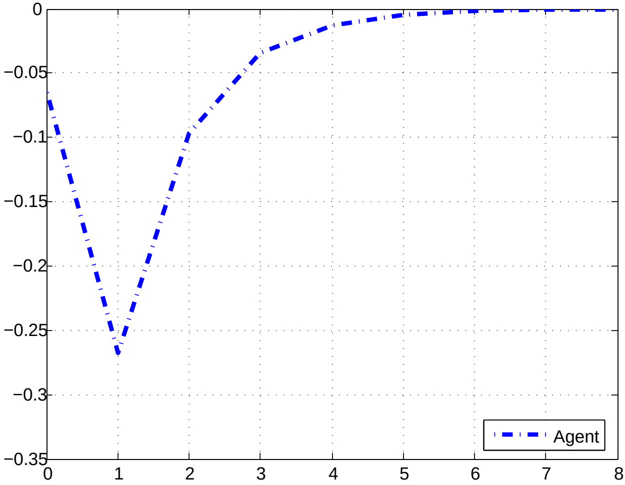
Econometrician's Discounting I

- Econometrician does not discount news same way as agent
- Econometrician recovers current and past ε^*_τ *not* ε_τ
- Econometrician's innovations are "old news"

$$\varepsilon^*_{\tau,t} = \theta\varepsilon_{\tau,t} + (1-\theta^2)\varepsilon_{\tau,t-1} - \theta(1-\theta^2)\varepsilon_{\tau,t-2} + \theta^2(1-\theta^2)\varepsilon_{\tau,t-3} + \cdots$$

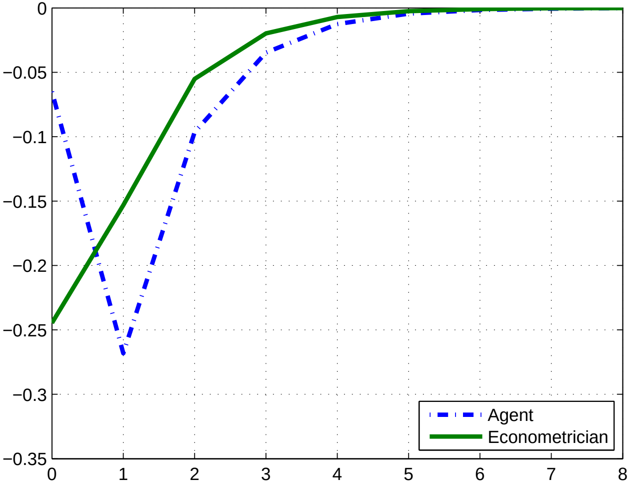
- Econometrician discounts innovations incorrectly because information set lags agents'
  - econometrician: k_t depends on $\theta\varepsilon^*_{\tau,t-1} + \varepsilon^*_{\tau,t}$
  - agents: k_t depends on $\varepsilon_{\tau,t-1} + \theta\varepsilon_{\tau,t}$

Impulse Response Functions: I



Response of k to Tax Shock

Impulse Response Functions: I



Response of k to Tax Shock

Econometrician's Estimates: II

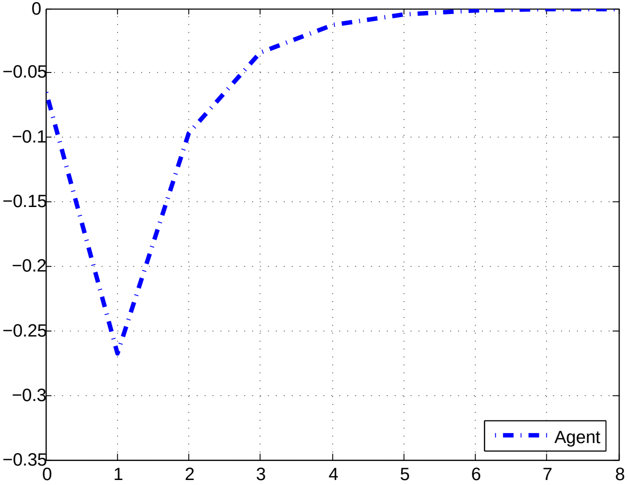
- Make more plausible assumption that econometrician does not observe technology: $\{\hat{\tau}_{t-j}, k_{t-j}\}_{j=0}^{\infty}$
- Does $\{\hat{\tau}_{t-j}, k_{t-j}\}_{j=0}^{\infty} \equiv \{\varepsilon_{\tau,t-j}, \varepsilon_{A,t-j}\}_{j=0}^{\infty}$?
- No: Econometrician's shocks convolute agents' news

$$\varepsilon^*_{\tau,t} = a_1\varepsilon_{\tau,t-1} + a_2\varepsilon_{\tau,t-2} + a_3\varepsilon_{A,t-1} + a_4\varepsilon_{A,t-2}$$
$$\varepsilon^*_{A,t} = b_1\varepsilon_{\tau,t} + b_2\varepsilon_{\tau,t-1} + b_3\varepsilon_{A,t} + b_4\varepsilon_{A,t-1}$$

a's and b's are functions of model parameters

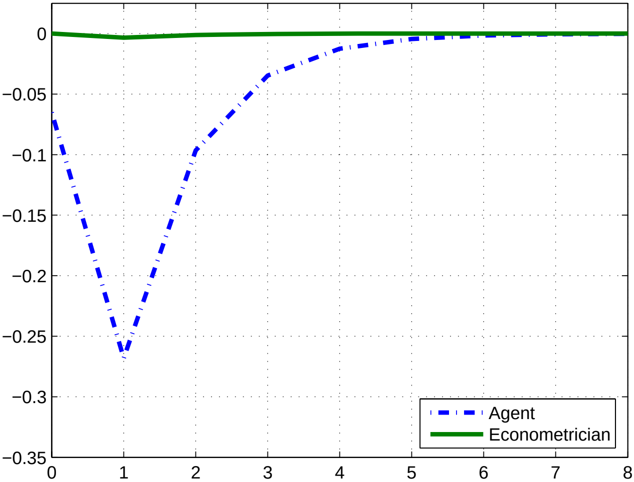
- Econometrician gets effects of *both* taxes and technology wrong
- Conclude taxes don't matter; everything driven by technology

Impulse Response Functions: II



True Response of k to Tax Shock

Impulse Response Functions: II



Response of k to Tax Shock

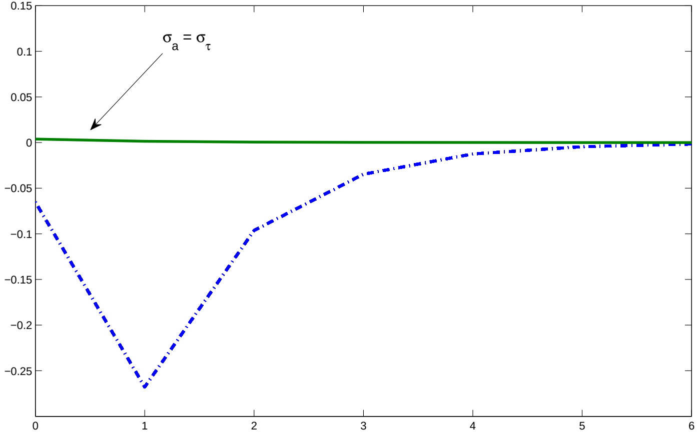
Impulse Response Functions: III

- Have shown that depending on econometrician's information set, may infer that higher expected taxes
  - are contractionary (Mountford & Uhlig)
  - have little effect (Poterba, Blanchard & Perotti, Romer & Romer)

- Now show may infer higher expected taxes are expansionary (Mertens & Ravn, House & Shapiro)

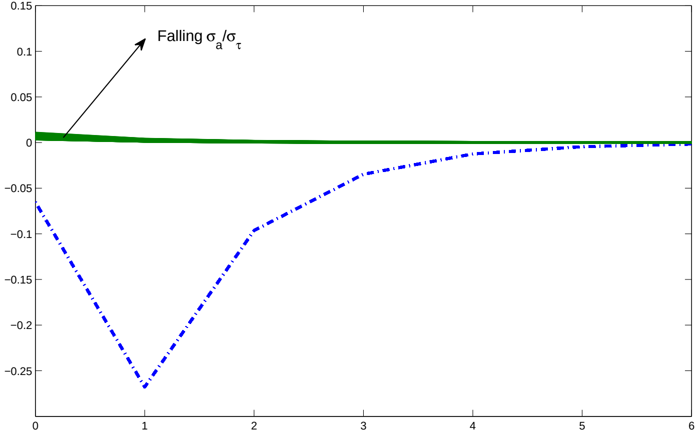
- Consider effects of variations in σ_a/σ_τ (relative volatility of technology and taxes)
  - alters the signal-extraction problem econometrician faces
  - as $\sigma_a/\sigma_\tau \to 0$, problem gets worse
  - infer higher expected taxes are increasingly expansionary

Impulse Response Functions: III



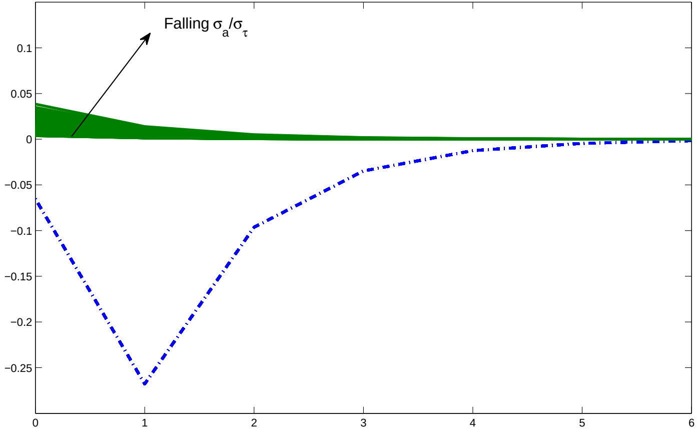
Responses of k to Higher Expected Taxes

IMPULSE RESPONSE FUNCTIONS: III



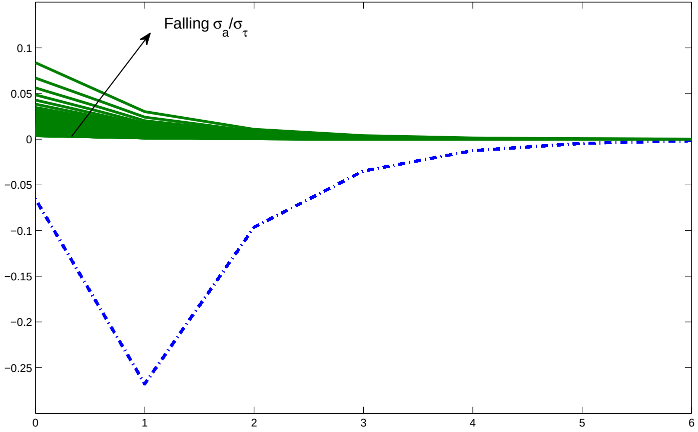
Responses of k to Higher Expected Taxes

Impulse Response Functions: III



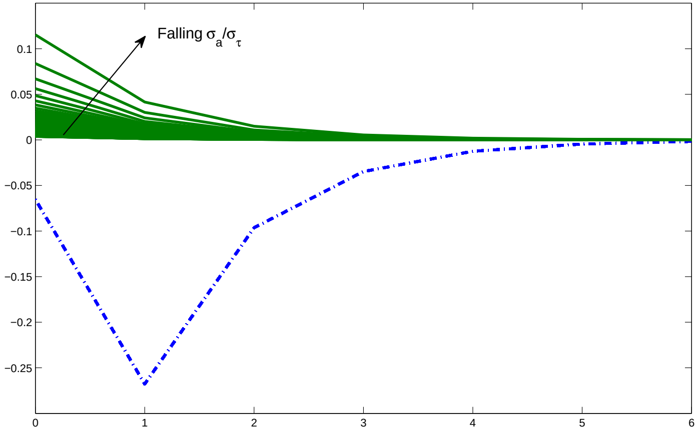
Responses of k to Higher Expected Taxes

Impulse Response Functions: III



Responses of k to Higher Expected Taxes

IMPULSE RESPONSE FUNCTIONS: III



Responses of k to Higher Expected Taxes

ABCD Test for Invertibility

- Consider the system with 2 period foresight whose eqm is

$$\begin{bmatrix} \tau_t \\ k_t \end{bmatrix} = \begin{bmatrix} L^2 & 0 \\ -\frac{\kappa(L+\theta)}{1-\alpha L} & \frac{1}{1-\alpha L} \end{bmatrix} \begin{bmatrix} \varepsilon_{\tau,t} \\ \varepsilon_{A,t} \end{bmatrix}$$

- We showed by directly computing the roots of the MA term that this is not invertible
- Fernandez-Villaverde, Rubio-Ramirez, Sargent, Watson propose a simple test of invertibility for a system written in state-space form
- This test will give identical results as checking roots of MA

ABCD Test: State-Space Form

$$\begin{bmatrix} \tau_{t+1} \\ k_{t+1} \\ \varepsilon_{\tau,t+1} \\ \varepsilon_{\tau,t} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & \alpha & -\theta & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \tau_t \\ k_t \\ \varepsilon_{\tau,t} \\ \varepsilon_{\tau,t-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -\kappa\theta & 1 \\ 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \varepsilon_{\tau,t+1} \\ \varepsilon_{A,t+1} \end{bmatrix}$$

$$x_{t+1} = Ax_t + Bw_{t+1}$$

$$\begin{bmatrix} \tau_{t+1} \\ k_{t+1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & \alpha & -\theta & 0 \end{bmatrix} \begin{bmatrix} \tau_t \\ k_t \\ \varepsilon_{\tau,t} \\ \varepsilon_{\tau,t-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -\kappa\theta & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{\tau,t+1} \\ \varepsilon_{A,t+1} \end{bmatrix}$$

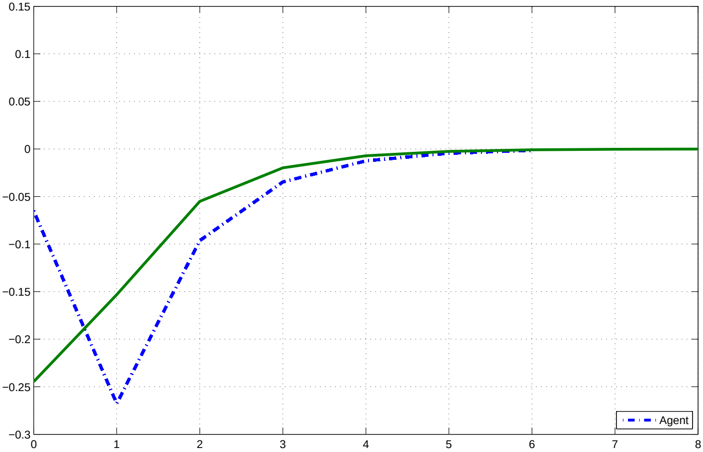
$$y_{t+1} = Cx_t + Dw_{t+1}$$

- If D^{-1} exists, then the system is invertible if and only if all the eigenvalues of $A - BD^{-1}C$ are inside the unit circle
- With foresight and no unanticipated part to taxes, D s singular and ABCD test cannot be applied
- Are straightforward ways to make D non-singular
  - add an unanticipated contemporaneous shock to tax rule: e_t^u
  - allow automatic stabilizers: $\tau_t = \phi y_t + \varepsilon_{\tau,t-q}$

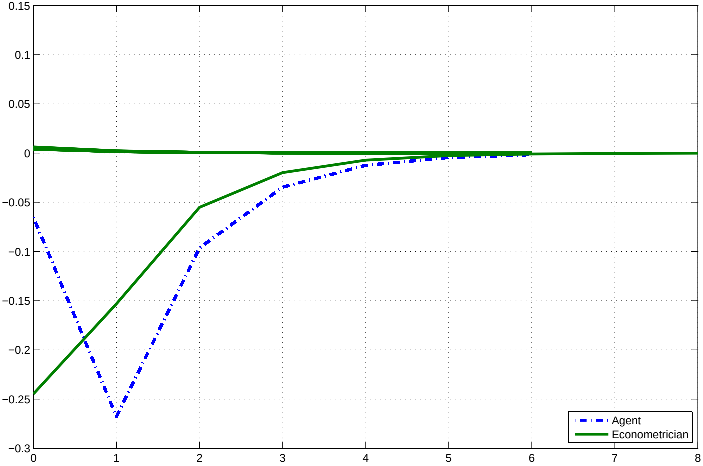
INTERMISSION

- As vary econometrician's information set
  - by adding or subtracting data
  - by changing signal-extraction problem
- Can obtain *any* inference about effects of news about higher future taxes
- Connects to findings in empirical literature

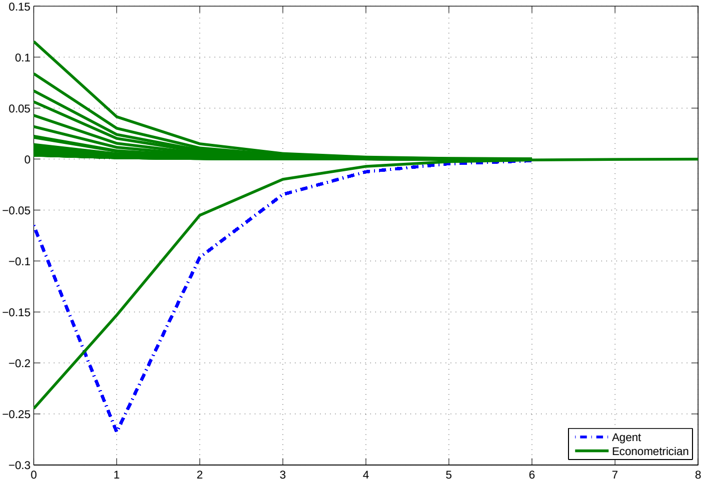
ANTICIPATED TAX HIKE: CONTRACTIONARY



Anticipated Tax Hike: No Effect



Anticipated Tax Hike: Expansionary



ROBUSTNESS

1. Do results from simple model carry over to more complex settings?
   - extend model specification—a “serious” model
   - generalize information flows

2. How do model elements alter the effects of foresight?
   - internal propagation mechanisms alter econometric errors in important ways

3. How sensitive are results to alternative assumptions about information flows?
   - relax rigid assumptions above and in “news” literature

Information Flows—Intuition

- Tax process

$$\hat{\tau}_t = \rho_1 \hat{\tau}_{t-1} + \cdots + \rho_n \hat{\tau}_{t-n} + \varepsilon_{\tau,t-2}$$

$\alpha\beta(1-\tau) < (1+\rho_1)^{-1} \;\Rightarrow$ non-invertibility.

- Tax process

$$\hat{\tau}_t = \psi \varepsilon_{\tau,t-2} + (1-\psi)\varepsilon_{\tau,t-1} \qquad \psi \in (0,1)$$

- Capital Dynamics

$$k_t = \alpha k_{t-1} + \varepsilon_{A,t} - \left(\frac{(1-\theta)\tau}{1-\tau}\right)\{[1-\psi(1-\theta)]\varepsilon_{\tau,t} + \psi \varepsilon_{\tau,t-1}\}$$

- If more recent news receives the heavier discount, $1-\psi(1-\theta) < \psi$, then the equilibrium will be non-invertible.

A Serious Model

- Includes elastic labor supply, variable utilization rates for capital inputs, durable and non-durable consumption, habit formation in non-durable consumption, investment adjustment costs, deliberation costs for durable, capital and labor taxes goods

- Calibrated to US data 1947:Q1 to 2008:Q2

- Tax Rules: Generalizing information flows

$$\widehat{\tau}_t^i = \rho_i \widehat{\tau}_{t-1}^i + q_i \widehat{s}_{t-1}^B + \mu_i \widehat{Y}_t + \sum_{j=0}^{4} \phi_j \varepsilon_{i,t-j}$$

  with $\sum_j \phi_j = 1$ for $i = K, L$

  - ϕ's are weights that imply trajectories of expected tax rates
  - determined by technology of tax choice

Information Flows and Tax Multipliers

A class of information flows—the ϕ's in $\sum_j \phi_j \varepsilon_{i,t-j}$

1. Use serious model and randomly draw ϕ's according to:

$$\phi_4 \sim 1 - \beta(1.5, 5); \qquad \phi_1, \phi_2, \phi_3 \sim U[-0.1, 0.1]$$

2. ϕ_4 is left skewed between 0 and 1 with mean of 0.77 and standard deviation of 0.15 following work on tax information flows by Yang (2008)
   - *very* conservative period of foresight
   - ϕ's reflect average degree of foresight

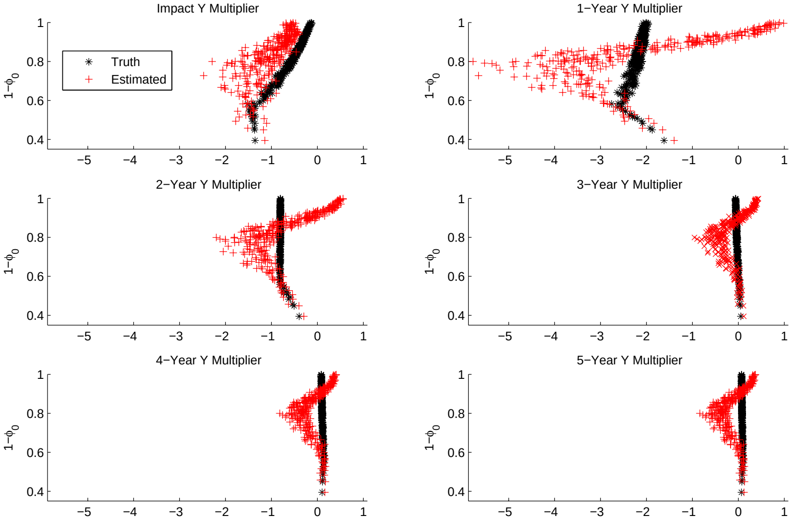
3. Need not imply non-fundamental representation

4. Estimate identified VAR and calculate dynamic multipliers
   - obtain distribution for degree of foresight
   - derive distribution for errors in inference

Information Flows and Tax Multipliers

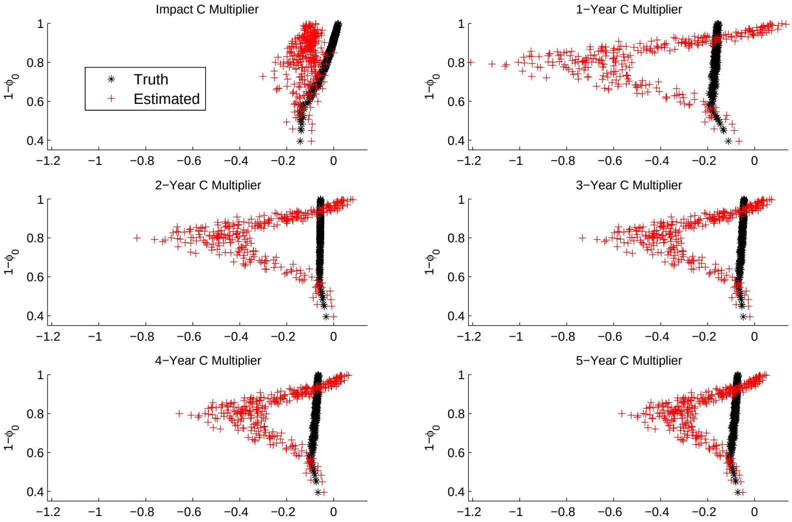
- Move beyond $0 - 1$ treatment of noninvertibility
- Ask if errors of inference are *quantitatively* important
- Empirical estimates of multipliers from anticipated tax changes are all over the map:
  1. Poterba, Blanchard & Perotti, Romer & Romer: ≈ 0
  2. Mountford & Uhlig: strongly negative
  3. Mertens & Ravn: strongly positive

Information Flows and Tax Multipliers



K-Tax Output Multipliers: Actual & Estimated

INFORMATION FLOWS AND TAX MULTIPLIERS



K-Tax Consumption Multipliers: Actual & Estimated

SOLUTIONS: THE USUAL SUSPECTS

- Expand the econometrician's information set
  - add financial variables [Sims, Beaudry-Portier]
  - add revenue forecasts [Romer-Romer]
  - need strong identifying assumptions in either case
- Impose sign restrictions to identify a VAR
  - theory supports *any* response of economic activity to anticipated tax increase
- Estimate a DSGE model
  - fine; conditional on getting information flows right [Blanchard-L'Huillier-Lorenzoni, Christiano et al., Schmitt-Grohe/Uribe]

Broader Implications

- Analysis extends to other areas where information flows emphasized
  - News about future technological improvement [Beaudry-Portier; Christiano-Ilut-Motto-Rostagno; Jaimovich-Rebelo; Schmitt-Grohe/Uribe]
  - Foresight about government spending run ups [Ramey-Shapiro, Ramey]
  - Inflation-targeting central banks that publish interest rate paths [Laseen-Linde-Svensson]
  - Distinction between "authorization" (e.g., 2009) and actual "outlays" (through 2019), especially for government infrastructure spending [Leeper-Walker-Yang]

Two Empirical Lines of Attack

- *Ex-post*: estimate conventional VARs
  - creative identification of “anticipated taxes”
  - Sims, Blanchard-Perotti, Mountford-Uhlig, Yang
- *Ex-ante*: reject VARs
  - identify foresight through narrative method
  - Ramey-Shapiro, Ramey, Romer-Romer, Mertens-Ravn
- Both lines seek instruments for foresight (future k)
- We assess the methods
  - formalize “narrative” approach
  - theory leads to skepticism about the methods

Ex-Post Approach

- Estimate VAR ignoring foresight *and then* impose identification restrictions to deal with foresight
- Tend to conclude foresight is second-order [Mountford-Uhlig and Blanchard-Perotti]
- But if foresight not handled properly, variation due to anticipated shocks gets attributed to unanticipated shocks
  - consider the following tax rule in the simple model

$$\hat{\tau}_t = \mathbf{e_t^u} + \varepsilon_{t-q}$$

    $e_t^u \sim i.i.d. \ \Rightarrow$ no effect on dynamics of k

  - econometrician who estimates VAR $\{a_t, k_t\}$ and ignores foresight attributes *all* dynamics of anticipated shock to unanticipated shock

Blanchard-Perotti

- Legislative lags used to achieve identification
- B-P admit identification is tenuous if foresight taken seriously
- In our simple model, the VAR representation yields

$$k_t = \alpha k_{t-1} + \eta_t^k,$$
$$\hat{\tau}_t = -\kappa\delta^2 k_{t-1} + \kappa\alpha\delta^2 k_{t-2} + \eta_t^\tau.$$

  where $\eta_t^k = \delta^{-1}\varepsilon^*_{A,t}$ and $\eta_t^\tau = \delta\varepsilon^*_{\tau,t}$
- Use η_{t+1}^τ as instrument for agent's news at t
- But $\eta_{t+1}^\tau = \delta\varepsilon^*_{\tau,t+1} = \delta[\delta\varepsilon_{\tau,t} + \kappa\varepsilon_{A,t}]$
- Instrument is a mongrel shock, confounding tax news and technology

MOUNTFORD-UHLIG

- Ambitious—identify several shocks: taxes, spending, monetary policy, business cycle
- Use sign restrictions to address fiscal foresight
- Impose zero restrictions on response of fiscal variables over period of foresight: tax revenues cannot move for q periods
- Delivers eccentric result that output falls while tax revenues do not change $\Rightarrow$ implicitly injects a sequence of *unanticipated* tax-rate shocks
- Sign-restrictions also likely to incorrectly other shocks, on whom the tax identification is conditional

Ex-Ante Approach

- Reject VARs *ex-ante* as unable to align information sets
- Foresight implies agents' news a function of current & *future* data
- In the simple analytical example

$$\kappa\varepsilon_{\tau,t} = k_t - (\alpha^{-1} + \theta)k_{t+1} + \theta(\alpha^{-1} + \theta)k_{t+2}$$
$$-\theta^2(\alpha^{-1} + \theta)k_{t+3} - \cdots$$

- *Ex-ante* approach uses changes in revenue forecasts due to legislation to instrument for $\{k_{t+j}\}$
- Richer model: $\varepsilon_{\tau,t}$ a linear combo of all responses of endogenous variables
- Romer-Romer use narrative to classify forecasted revenue changes as "endogenous" or "exogenous"
- Need to interpret narrative method

Formalizing the Narrative Method

- To reflect multiplicity of motivations for tax policy in Romers' narrative

$$\begin{aligned}\hat{\tau}_t &= \rho(L)\hat{\tau}_{t-1} + \underbrace{\sum_{j=-P}^{P} \mu_j^C E_t y_{t+j}^C + \sum_{j=-M}^{M} \beta_j E_t g_{t+j}}_{\text{“}RR\ endogenous\text{”}} \\ &\quad + \underbrace{\sum_{j=-P}^{P} \mu_j^T E_t y_{t+j}^T + \sum_{j=-N}^{N} \gamma_j E_t s_{t+j-1}^B + \varepsilon_{\tau,t-q} + e_{\tau,t}^u}_{\text{“}RR\ exogenous\text{”}}\end{aligned}$$

Formalizing the Narrative Method

- Specialize tax rule & information flows to

$$\tau_t = \rho\tau_{t-1} + \mu^C y_t + \xi_{t-q} + e^u_{\tau,t}$$

where foresight is given by ξ_{t-q} and

$$\xi_{t-q} = \mu^T y_{t-q-1} + \gamma s^B_{t-q-1} + \varepsilon_{\tau,t-q}$$

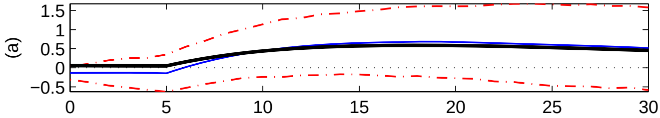
- Embed various specifications of tax behavior in DSGE model with capital and labor tax rates
- Simulate data & forecasted revenues
- Estimate VARs with forecasted revenues on right-side

$$X_t = CX_{t-1} + \sum_{i=0}^{24} D_i T^u_{t-i} + \sum_{i=0}^{24} F_i T^a_{t-i} + \sum_{i=1}^{6} G_i T^a_{t+i} + u_t$$

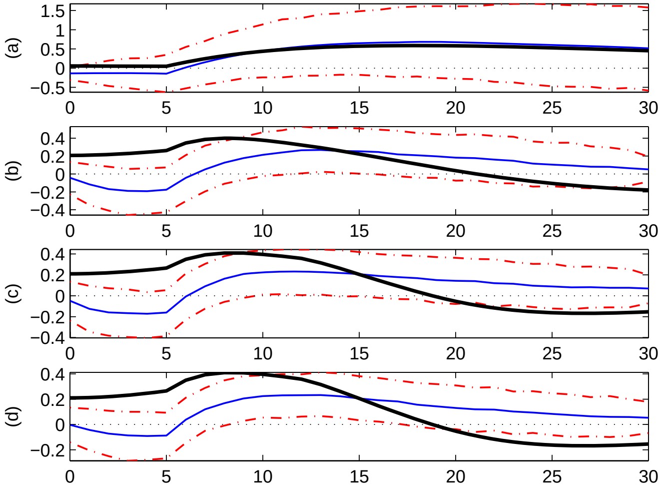
FORMALIZING THE NARRATIVE METHOD

- Alternative parametric interpretations of narrative method
  - (a) taxes exogenous; transfers adjust ($\mu^C = 0, \mu^T = 0, \gamma_T = -.1, \sigma_K = .025, \sigma_L = .02$)
  - (b) automatic stabilizers; taxes adjust ($\mu^C = 1, \mu^T = 0, \gamma_\tau = .05, \sigma_K = .025, \sigma_L = .02$)
  - (c) automatic stabilizers; response to trend; taxes adjust ($\mu^C = 1, \mu^T = .5, \gamma_\tau = .05, \sigma_K = .025, \sigma_L = .02$)
  - (d) (c) with higher relative variability of anticipated taxes ($\mu^C = 1, \mu^T = 5, \gamma_\tau = .05, \sigma_{Ka} = .0375, \sigma_{Ku} = .0125, \sigma_{La} = .03, \sigma_{Lu} = .01$)
- Data and forecasts come from single coherent model
- If *ex-ante* efficacious, should nail true effects

CONS. RESPONSES TO LABOR TAXES



Cons. Responses to Labor Taxes



Summary of Ex-Ante Approach

- *Ex-ante* approach may perform well or poorly: Conditional on how narrative approach formalized
- Narrative method of identification is not uniquely reproducible
- Different reasonable formalizations produce different conclusions
- *Ex-ante* approach does not model information flows: The more exogenous the forecasted revenues, the better the performance
- Connection between policy behavior and agents' information left implicit
- Difficult to integrate identification scheme into efforts to estimate DSGE models

Appendix: Root Flipping via Kalman Filter

- A very cool result: instead of using Blaschke factors, as in the paper, could obtain econometrician's information set—the VAR—using Kalman filter
- This may be surprising
- Usually we use the Kalman filter to get best linear prediction in models with latent variables
- But with non-invertibility induced by foresight, Kalman filter will not align agents and econometrician's info sets
- Kalman filter will, however, correctly recover the *econometrician's* info set

Root Flipping via Kalman Filter

- Deriving the fundamental (or invertible) representation is referred to as "root flipping"
- Two ways to flip roots: Blaschke factors & Kalman filter
- Consider the following state space representation

$$x_{t+1} = Ax_t + Gw_{1t+1} \quad (4)$$
$$y_t = Cx_t + w_{2t}$$

where $[w'_{1,t+1}, w'_{2t}]$ is a white noise vector with covariance matrix

$$E\begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix}\begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix}' = \begin{bmatrix} V_{1t} & V_{3t} \\ V'_{3t} & V_{2t} \end{bmatrix} \quad (5)$$

ROOT FLIPPING VIA KALMAN FILTER

- Two useful representations can be derived from the Kalman filter
- The first is an "innovations representation"

$$\hat{x}_{t+1} = A\hat{x}_t + Ka_t \tag{6}$$
$$y_t = C\hat{x}_t + a_t$$

where K is the Kalman gain, $\hat{x}_s$ is the optimal projection of x_s conditional on observing $y_s, y_{s-1}, ...$, and a_t is the innovation in predicting y_t linearly from observing current and past y's
- The covariance matrix of the innovations is given by

$$Ea_t a_t' = C\Sigma_t C' + V_{2t}$$

where Σ_t solves the matrix Ricatti equation

$$\Sigma_{t+1} = A\Sigma_t A' + GV_{1t}G' - (A\Sigma_t C' + GV_{3t})(C\Sigma_t C' + V_{2t})^{-1}(A\Sigma_t C + GV_{3t})'$$

ROOT FLIPPING VIA KALMAN FILTER

- Can now see the invertibility condition clearly; using the lag operator L and solving innovations rep for y_t gives

$$y_t = [I + C(L^{-1}I - A)^{-1}K]a_t$$

- In order for y_t (the observables) to span the same linear space as the innovations, the zeroes of the determinant:

$$\text{det}[I + C(zI - A)^{-1}K] = \frac{\text{det}[zI - (A - KC)]}{\text{det}(zI - A)} = 0, \Rightarrow \ |z| < 1$$

  cannot be outside the unit circle
  - $z = L^{-1}$ here; this condition is equivalent to the condition in LWY
  - the zeros of det$[zI - (A - KC)]$ are the eigenvalues of $A - KC$

ROOT FLIPPING VIA KALMAN FILTER

- The second representation that is useful to consider is the whitening filter

$$\begin{aligned} \hat{x}_{t+1} &= (A - KC)\hat{x}_t + Ky_t \\ a_t &= y_t - C\hat{x}_t \end{aligned} \tag{7}$$

  - a whitening filter takes a sequence of y's and gives as output a sequence of a's that are serially uncorrelated
  - can see why the invertibility condition is crucial
  - if the eigenvalues of $A - KC$ are not all inside the unit circle, then (7) is not a stationary process

ROOT FLIPPING VIA KALMAN FILTER

- Consider the following ARMA process from LWY

$$k_t = \alpha k_{t-1} - \kappa\{\varepsilon_{\tau,t-1} + \theta\varepsilon_{\tau,t}\}$$

- A state space formulation for this process is given by $x_t = -(\kappa\theta)^{-1}k_t - \varepsilon_{\tau,t}$, $y_t = -(\kappa\theta)^{-1}k_t$

$$x_{t+1} = \alpha x_t + [\alpha + \theta^{-1}]\varepsilon_{\tau,t}$$
$$y_t = x_t + \varepsilon_{\tau,t} \tag{8}$$

and $V_1 = V_2 = V_3$. We assume that the initial state is known ($\Sigma_{t_0} = 0$), which implies that the Kalman gain is given by $K_t = G$ for all t and then

$$A - KC = \alpha - \alpha - \theta^{-1} = |\theta^{-1}| > 1$$

The root of $z - A + KC$ is outside the unit circle and the innovations of (8) ($\varepsilon_{\tau,t}$) do not span the same space as the observables and therefore (8) cannot be the innovations representation

Root Flipping via Kalman Filter

- Time invariance of the Kalman filter requires two primary assumptions:
  1. The pair (A', C') is stabilizable. A pair (A, C) is stabilizable if $y'C = 0$ and $y'A = \lambda y'$ for some complex number λ and some complex vector y implies that $|\lambda| < 1$ or $y = 0$
  2. The pair (A, G) is detectable. The pair (A, G) is detectable if $G'y = 0$ and $Ay = \lambda y$ for some complex number λ and some complex vector y implies that $|\lambda| < 1$ or $y = 0$
- Given (1) & (2), iterations on the matrix Ricatti equation converge as $t \to \infty$, starting from any semi-positive definite matrix Σ_{t_0}
  - This then implies a time invariant Kalman gain K
  - And $A - KC$ is a stable matrix with eigenvalues less than unity in modulus

Root Flipping via Kalman Filter

- Returning to the example, both assumptions hold for the non-invertible ARMA process
- Both conditions yield $\alpha = \lambda < 1$ or $y = 0$—implying there exists a K such that $A - KC$ is less than one
- Note that the Ricatti equation and time-invariant solution are given by

$$\sigma_{t+1} = \alpha^2 \sigma_t + (\alpha + \theta^{-1})^2 + \frac{(\alpha\sigma_t + \alpha + \theta^{-1})^2}{1 + \sigma_t}$$

$$\sigma_\infty = \frac{1 - \theta^2}{\theta^2}$$

This gives the Kalman gain as

$$K = (A\Sigma_\infty C' + GV_{3t})(C\Sigma_\infty + V_{2t})^{-1} = \frac{\alpha\sigma_\infty + \alpha + \theta^{-1}}{1 + \sigma_\infty} = \alpha + \theta$$

and now $A - KC = -\theta < 1$

Root Flipping via Kalman Filter

- The representation gives the innovation as

$$
\begin{aligned}
-(\kappa\theta)^{-1}k_t &= [1 + \frac{(\alpha+\theta)}{(L^{-1}-\alpha)}]a_t \\
&= \left[\frac{L^{-1}+\theta}{L^{-1}-\alpha}\right]a_t \\
&= \left[\frac{1+\theta L}{1-\alpha L}\right]a_t
\end{aligned}
$$

$$(1-\alpha L)k_t = -\kappa(1+\theta L)\theta a_t$$

- This is equivalent to equation (13) of LWY
- It implies that $\theta a_t = \varepsilon_t^*$, so the impulse response function using the Kalman filter must be normalized by the standard deviation of a_t