MAGYAR NEMZETI BANK MINI-COURSE
Lecture 6. FORESIGHT: THEORY AND ECONOMETRICS

Eric M. Leeper
Indiana University
September 2008
THE MESSAGES

- Draws on new work with Todd Walker & Shu-Chun Yang
- That focuses on *fiscal* foresight, but the issue is broader
- 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
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

- $\varepsilon_t$: vector of exogenous shocks agents observe
- $\varepsilon_{\tau,t}$: tax component
- $\Omega_t = \text{span}\{\varepsilon_t, \varepsilon_{t-1}, \ldots\}$ agents’ information set
- $\varepsilon^*_t$: 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 $\Omega_t$
Theory in Search of Application

- Hansen-Sargent pointed out problems of non-invertibility in general, abstract setting [see also Lippi-Reichlin]
- Our contribution
  - show how Sargent & Hansen’s insights apply to fiscal foresight
  - highlight & thoroughly study an important practical example of non-invertibility
  - use simple economic theory to exposit the issues explicitly
  - explore robustness to models & information flows
  - evaluate quantitative implications of non-invertibility
  - assess existing empirical work in light of the theory
  - lay out an econometric approach grounded in theory
No Consensus in Existing Empirical Work

• Results all over the map
• 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]
• Sources of the diverse results
  • weak instruments for fiscal foresight
  • no modeling of fiscal behavior that gives rise to foresight (information flows)
  • non-invertibility not directly confronted
ANECDOITAL EVIDENCE OF FISCAL FORESIGHT

- Large public finance literature [Auerbach-Slemrod]
- House & Shapiro: jobless recovery of 2001—tax phase-ins induced production delays
- Ramey: War dummies predict defense spending
## Capital Gains in Anticipation of TRA86

<table>
<thead>
<tr>
<th>Year</th>
<th>Long-term</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>1984</td>
<td>134.1</td>
<td>135.0</td>
</tr>
<tr>
<td>1985</td>
<td>164.9</td>
<td>167.0</td>
</tr>
<tr>
<td>1986</td>
<td>315.7</td>
<td>322.2</td>
</tr>
<tr>
<td>1987</td>
<td>—</td>
<td>137.4</td>
</tr>
<tr>
<td>1988</td>
<td>—</td>
<td>153.8</td>
</tr>
<tr>
<td>1989</td>
<td>—</td>
<td>145.6</td>
</tr>
<tr>
<td>1990</td>
<td>—</td>
<td>113.2</td>
</tr>
</tbody>
</table>

Capital Gains Realizations in Billions  
Source: Auerbach & Slemrod (1997)
Funds Rate in Anticipation of TRA86

Daily Federal Funds Rate
**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

\[
\frac{1}{C_t} = \alpha \beta E_t(1 - \tau_{t+1}) \frac{1}{C_{t+1}} \frac{Y_{t+1}}{K_t}
\]

\[
C_t + K_t = Y_t = A_t K_{t-1}^\alpha
\]
**Solving the Model**

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

\[
E_t k_{t+1} - \frac{1 + \alpha^2 \beta (1 - \tau)}{\alpha \beta (1 - \tau)} k_t + \frac{1}{\beta (1 - \tau)} k_{t-1}
\]

\[
= 1 - \frac{\alpha \beta (1 - \tau)}{\alpha \beta (1 - \tau)} \left( \frac{\tau}{1 - \tau} \right) E_t \hat{r}_{t+1} - \frac{1}{\alpha \beta (1 - \tau)} a_t
\]

or

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

with

\[
\gamma_0 = \frac{1 + \alpha^2 \beta (1 - \tau)}{\alpha \beta (1 - \tau)} > 0, \quad \gamma_1 = \frac{1}{\beta (1 - \tau)} > 0
\]

\[
\nu_0 = -\frac{1}{\alpha \beta (1 - \tau)} < 0, \quad \nu_1 = \frac{1 - \alpha \beta (1 - \tau)}{\alpha \beta (1 - \tau)} \left( \frac{\tau}{1 - \tau} \right) > 0
\]
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 \( t \)
Equilibrium capital accumulation obeys

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

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

- \( \theta \) plays central role in analysis
- Agent uses \( \theta \) 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

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

- 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

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

- More recent news is discounted (by \( \theta \)) 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 ⇒ projections equivalent to conditional expectations

• VAR is projection $P[x_t|x_{t-1}, x_{t-2}, ...]$

• If $x_{t-1}, x_{t-2}, ...$ reveals agents’ information set, then econometrician captures dynamics of economy

• Foresight implies $x_{t-1}, x_{t-2}, ...$ 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

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

- 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

\[ x_t = \left[ \frac{L - \theta}{1 - \rho L} \right] \left[ \frac{1 - \theta L}{L - \theta} \right] \left[ \frac{L - \theta}{1 - \theta L} \right] \varepsilon_t \]
\[ x_t = \left[ \frac{1 - \theta L}{1 - \rho L} \right] e_t \quad (2) \]
\[ e_t = \left[ \frac{L - \theta}{1 - \theta L} \right] \varepsilon_t \quad (3) \]

- (2) is invertible: current & past $x_t$ span same space as current & past $e_t$ (but *not* $\varepsilon_t$)
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
- \( \varepsilon_t \) 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 \( \varepsilon_t \) 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 \( \varepsilon_t \)
- Setting \( F = L^{-1} \), Blaschke factor has valid inverse in the forward operator \( F \)

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

- hence \( \varepsilon_t \) carries information about future \( e \)'s (and \( x \)'s)
ASIDE: REPRESENTATION THEORY

• A famous example

\[ y_t = w_t + 2w_{t-1}, \quad w_t \sim iidN(0, 1) \]

• 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

\[ y_t = 2(a_t + (1/2)a_{t-1}), \quad a_t \sim iidN(0, 1) \]

• 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

\[
(1 - \alpha L)k_t = -\kappa(L + \theta)e_{\tau,t}
\]

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

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

But \( \theta < 1 \), so not invertible in current and past capital

- *Is* invertible in current and future capital

\[
k_t = (\alpha^{-1} + \theta)k_{t+1} - \theta(\alpha^{-1} + \theta)k_{t+2} + \\
\theta^2(\alpha^{-1} + \theta)k_{t+3} - \cdots + \kappa e_{\tau,t}
\]
**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

\[
(1 - \alpha L)k_t = -\kappa(L + \theta)\left[\frac{1 + \theta L}{L + \theta}\right]\left[\frac{L + \theta}{1 + \theta L}\right]\varepsilon_{\tau,t}
\]

\[
= -\kappa(1 + \theta L)\varepsilon_{\tau,t}^*
\]

\[
= -(1 - \theta)\left(\frac{\tau}{1 - \tau}\right)\left\{\theta\varepsilon_{\tau,t-1}^* + \varepsilon_{\tau,t}^*\right\}
\]
**Econometrician’s Discounting**

- Econometrician does not discount news same way as agent
- Econometrician recovers current and past $\varepsilon^*_\tau$ *not* $\varepsilon_\tau$
- 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 $\theta \varepsilon^*_{\tau,t-1} + \varepsilon^*_\tau$
  - agents: $k_t$ depends on $\varepsilon_{\tau,t-1} + \theta \varepsilon_{\tau,t}$
**Impulse Response Functions: I**

![Graph showing the true response of k to Tax Shock](image-url)

- **True 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: \( \{\hat{\tau}_{t-j}, \hat{k}_{t-j}\}_{j=0}^\infty \)

• Does \( \{\hat{\tau}_{t-j}, \hat{k}_{t-j}\}_{j=0}^\infty \equiv \{\epsilon_{\tau,t-j}, \epsilon_{A,t-j}\}_{j=0}^\infty \)?

• No: Econometrician’s shocks convolute agents’ news

\[
\epsilon_{\tau,t}^* = a_1 \epsilon_{\tau,t-1} + a_2 \epsilon_{\tau,t-2} + a_3 \epsilon_{A,t-1} + a_4 \epsilon_{A,t-2}
\]

\[
\epsilon_{A,t}^* = b_1 \epsilon_{\tau,t} + b_2 \epsilon_{\tau,t-1} + b_3 \epsilon_{A,t} + b_4 \epsilon_{A,t-1}
\]

\(a’s\) 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: I**

True Response of $k$ to Tax Shock
Impulse Response Functions: II

Response of $k$ to Tax Shock
Econometrician’s Estimates: III

- Obvious how to align information sets with two-quarter foresight easy: include $\hat{\tau}_{t+2}$ in time $t$ VAR
  - true, but won’t always work
- Jazz-up tax rule to include automatic stabilizers
  $$\hat{\tau}_t = \varphi a_t + \varepsilon_{\tau,t-2}$$

- Agent’s no longer have perfect foresight $\Rightarrow$ forecast error due to forecasting technology
- This additional noise in tax rule implies econometrician’s inference especially poor with respect to technology
- Assume $a_t = \rho a_{t-1} + \varepsilon_{A,t}$, $\rho = 0.01$ (nearly i.i.d.) and $\varphi = 1$
Impulse Response Functions: III

Response of $k$ to Technology
**Impulse Response Functions: III**

Response of $k$ to Technology
ABCD Test for Invertibility

- Consider the system with 2 period foresight whose eqm is

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

- 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**

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

\[
x_{t+1} = Ax_t + Bw_{t+1}
\]

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

\[
y_{t+1} = Cx_t + Dw_{t+1}
\]

**•** 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, $D$ is singular and ABCD test cannot be applied

**•** Are straightforward ways to make $D$ non-singular
  - add an unanticipated contemporaneous shock to tax rule: $\varepsilon_{t+1}^u$
  - allow automatic stabilizers: $\tau_t = \phi y_t + \varepsilon_{\tau,t-q}$
Testing Economic Theory

- Present-value relations
  - Tests of fiscal sustainability ask if expected discounted present value of debt is zero (so government’s intertemporal budget constraint holds)
  - Campbell-Shiller test of PV impose restrictions on VAR
  - Test is misspecified with foresight (Type I Error)

- Reversal of Granger-causal ordering
  - Econometrician’s information set lags agents’
**Present-Value Relations**

- Government’s flow constraint (taking $E_{t-1}$ wrt agents’ info set $\Omega_{t-1} = \{\varepsilon_{1,t-j}, \varepsilon_{2,t-j}\}_{j=1}^{\infty}$)

\[
E(b_t|\Omega_{t-1}) = \beta^{-1}b_{t-1} - E(s_t|\Omega_{t-1})
\]

- Revenues & spending drive surpluses; 1 period of foresight
- Policy rule that ensures sustainability is

\[
st = \gamma b_{t-1} + \frac{\varepsilon_{1,t-1}}{1 - \rho_1 L} + \frac{\varepsilon_{2,t-1}}{1 - \rho_2 L}
\]

- Imposing TCV for debt: $\lim_{N \to \infty} \beta^N E(b_{t+N}|\Omega_{t-1}) = 0$

implies the eqm condition

\[
b_t = \sum_{j=1}^{\infty} \beta^j E(s_{t+j}|\Omega_{t-1}).
\]
Following Hansen-Roberds-Sargent, cross-equation restrictions that satisfy eqm condition are

\[
\begin{bmatrix}
  s_t \\
  b_t
\end{bmatrix} = \begin{bmatrix}
  \frac{LA(L)}{\beta[L^2A(L)-\beta^2A(\beta)]} & \frac{LC(L)}{\beta[L^2C(L)-\beta^2C(\beta)]}
\end{bmatrix} \begin{bmatrix}
  \varepsilon_{1,t} \\
  \varepsilon_{2,t}
\end{bmatrix}
\]

\[y_t = P(L)v_t\]

where \(A(L) = \frac{\beta^{-1}-\gamma}{(1-\rho_1L)(1-\gamma L)}\), and \(C(L) = \frac{\beta^{-1}-\gamma}{(1-\rho_2L)(1-\gamma L)}\).

Two observations from this expression:

- foresight \(\Rightarrow\) this is not invertible (a zero at \(L = 0\))
- cross-equation restrictions imposed on MAR are nonlinear
Present-Value Relations

- Campbell-Shiller derive PV restrictions on VAR rep instead of MA rep
- PV-constraint amounts to restrictions on the VAR coefficients
- Denote the invertible representation by $P^*(L)$ and write the corresponding VAR as

$$\begin{bmatrix} s_t \\ b_t \end{bmatrix} = A_{0}^{-1} A_{1}(L) \begin{bmatrix} s_{t-1} \\ b_{t-1} \end{bmatrix} + A_{0}^{-1} \begin{bmatrix} \varepsilon_{1t}^* \\ \varepsilon_{2t}^* \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} s_{t-1} \\ b_{t-1} \end{bmatrix} + \begin{bmatrix} w_{1t} \\ w_{2t} \end{bmatrix}$$

$$y_t = A^* y_{t-1} + w_t$$

Note that $A^*(L) = P(L)^{-1}$, implying that the coefficients of the VAR will not yield the correct cross-equation restrictions when there is foresight
Present-Value Relations

- ⇒ cross-equation restrictions on VAR coefficients

\[ a_{11} + a_{21} = 0, \quad a_{22} + a_{12} = \beta^{-1} \]

- Econometrician will test the restrictions

\[ a_{11} + a_{21} = \frac{\eta \rho_1 \rho_2 \beta A(\beta) C(\beta)}{\rho_2 C(\beta) - \rho_1 A(\beta)}, \quad a_{22} + a_{12} = \frac{A(\beta) \eta \rho_2 \rho_1 (C(\beta) - A(\beta))}{\beta (\rho_2 C(\beta) - \rho_1 A(\beta))} \]

where \( \eta = \left(1 + [A(\beta)C'(\beta)]^2\right)^{-1/2} \)

- Econometrician will incorrectly reject the null that PV holds
ROBUSTNESS OF NON-INVERTIBILITY

Parameter / Model Specifications

- Simple model with tax process

\[ \hat{\tau}_t = \rho_1 \hat{\tau}_{t-1} + \cdots + \rho_n \hat{\tau}_{t-n} + \varepsilon_{\tau,t-q} \]

\[ \alpha\beta(1 - \tau) < (1 + \rho_1)^{-1} \Rightarrow \text{non-invertibility for } q \geq 1 \]

- Tax rule

\[ \hat{\tau}_t^Z = \rho_Z \hat{\tau}_{t-1}^Z + \varphi_Z \hat{y}_t + \varepsilon_{\tau,t-4}^Z \text{ for } Z = L, K \]

Model Variations: elastic labor supply, variable utilization rates for capital and labor inputs, durable and non-durable consumption, habit formation in non-durable consumption, investment adjustment costs, and deliberation costs for durable goods \( \Rightarrow \) non-invertible for plausible parameters

- Non-invertibility very robust to alternative model specifications and parameter values

- Is it quantitatively important?
Robustness of Non-Invertibility

Alternative Information Flows

- Simple model with tax process

\[ \hat{\tau}_t = \psi \varepsilon_{\tau,t-2} + (1 - \psi) \varepsilon_{\tau,t-1} \]

\[ \psi \in (0, 1) \Rightarrow \text{no exact degree of foresight} \]

- Capital Dynamics

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

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

- Modeling of information flows critical: more work needed
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

\[ \hat{r}_t = e^u_t + \varepsilon_{t-q} \]

\( e^u_t \sim i.i.d. \Rightarrow \) 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
Legislative lags used to achieve identification

B-P admit identification is tenuous if foresight taken seriously

In our simple model, the VAR representation yields

\[ k_t = \alpha k_{t-1} + \eta^k_t, \]
\[ \hat{r}_t = -\kappa \delta^2 k_{t-1} + \kappa \alpha \delta^2 k_{t-2} + \eta^\tau_t. \]

where \( \eta^k_t = \delta^{-1} \varepsilon^*_A,t \) and \( \eta^\tau_t = \delta \varepsilon^*_{\tau,t} \)

Use \( \eta^\tau_{t+1} \) as instrument for agent’s news at \( t \)

But \( \eta^\tau_{t+1} = \delta \varepsilon^*_{\tau,t+1} = \delta [\delta \varepsilon_{\tau,t} + \kappa \varepsilon_A,t] \)

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

\[
\hat{\tau}_t = \rho(L)\hat{\tau}_{t-1} + \sum_{j=-P}^{P} \mu_j^C E_t y_{t+j}^C + \sum_{j=-M}^{M} \beta_j E_t g_{t+j}
\]

“RR endogenous”

\[
+ \sum_{j=-P}^{P} \mu_j^T E_t y_{t+j}^T + \sum_{j=-N}^{N} \gamma_j E_t s_{t+j-1}^B + \varepsilon_{\tau,t-q} + e_{\tau,t}
\]

“RR exogenous”
FORMALIZING THE NARRATIVE METHOD

- Specialize tax rule & information flows to

\[ \tau_t = \rho \tau_{t-1} + \mu^C y_t + \xi_{t-q} + e_{\tau,t}^u \]

where foresight is given by \( \xi_{t-q} \) and

\[ \xi_{t-q} = \mu^T y_{t-q-1} + \gamma s_{t-q-1}^B + \varepsilon_{\tau,t-q} \]

- 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

\[ X_t = CX_{t-1} + \sum_{i=0}^{24} D_i T_{t-i}^u + \sum_{i=0}^{24} F_i T_{t-i}^a + \sum_{i=1}^{6} G_i T_{t+i}^a + u_t \]
FORMALIZING THE NARRATIVE METHOD

• Alternative parametric interpretations of narrative method
  (a) taxes exogenous; transfers adjust
  \( \mu^C = 0, \mu^T = 0, \gamma_T = -0.1, \sigma_K = 0.025, \sigma_L = 0.02 \)
  (b) automatic stabilizers; taxes adjust
  \( \mu^C = 1, \mu^T = 0, \gamma_T = 0.05, \sigma_K = 0.025, \sigma_L = 0.02 \)
  (c) automatic stabilizers; response to trend; taxes adjust
  \( \mu^C = 1, \mu^T = 0.5, \gamma_T = 0.05, \sigma_K = 0.025, \sigma_L = 0.02 \)
  (d) (c) with higher relative variability of anticipated taxes
  \( \mu^C = 1, \mu^T = 5, \gamma_T = 0.05, \sigma_{Ka} = 0.0375, \sigma_{Ku} = 0.0125, \sigma_{La} = 0.03, \sigma_{Lu} = 0.01 \)

• Data and forecasts come from single coherent model
• If *ex-ante* efficacious, should nail true effects
CONS. RESPONSES TO LABOR TAXES

(a)
CONS. RESPONSES TO LABOR TAXES

(a)

(b)

(c)

(d)
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
Fiscal Foresight: Upshot

- Private agents discount recent news more heavily because it informs about taxes in the more distant future.
- The econometrician discounts in the usual way, down weighting older news relative to recent news.
- Agents and the econometrician employ different discounting patterns because the econometrician’s information set lags the agents’
- Fiscal foresight creates deep problems for econometric work.
- These problems are robust across many commonly used model variations.
- Confronting foresight is a necessary step toward detecting fiscal effects in macro time series.
Broader Implications & Future Work

- Analysis extends to other areas where information flows emphasized
  - News about future technological improvement [Beudry-Portier; Christiano, Ilut, Motto, and Rostagno; Jaimovich and Rebelo]
  - Foresight about government spending run ups [Ramey]
  - Inflation-targeting central banks that publish interest rate paths

- Future work will examine robustness of our results with respect to these other areas and assume various specifications of information flows
- Obtain quantitative consequences of non-invertibility in rich models
- Foresight as propagation mechanism in DSGE models
Broader Implications & Future Work

Tackling Econometric Issues

- Need deeper understanding of econometric issues: filtering, MLE, GMM, calibration, moment estimators
- Our approach
  1. Use spreads on asset yields with different tax treatment (e.g., munies vs. treasuries) to identify tax news
  2. Find robust sign restrictions on effects of tax foresight and impose restrictions on estimated VARMA
  3. Apply Ingram-Whiteman’s method to derive prior from a DSGE model and impose prior on estimated VARMA
  4. Specify & estimate full-blown DSGE model of tax policy
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

\[
x_{t+1} = Ax_t + Gw_{1t+1} \\
y_t = Cx_t + w_{2t}
\]  

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

\[
E \begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix} \begin{bmatrix} w_{1t+1} \\ w_{2t} \end{bmatrix}' = \begin{bmatrix} V_{1t} & V_{3t} \\ V_{3t}' & V_{2t} \end{bmatrix}
\]
ROOT FLIPPING VIA KALMAN FILTER

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

\[
\begin{align*}
\hat{x}_{t+1} &= A\hat{x}_t + Ka_t \\
y_t &= C\hat{x}_t + a_t
\end{align*}
\]

where \( K \) is the Kalman gain, \( \hat{x}_s \) is the optimal projection of \( x_s \) conditional on observing \( y_s, y_{s-1}, \ldots \), 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 \( \Sigma_t \) solves the matrix Ricatti equation

\[
\Sigma_{t+1} = A\Sigma_t A' + GV_1 t G' - (A\Sigma_t C' + GV_3 t)(C\Sigma_t C' + V_{2t})^{-1}(A\Sigma_t C + GV_{3t})'
\]
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 $\text{det}[zI - (A - KC)]$ are the eigenvalues of $A - KC$
The second representation that is useful to consider is the whitening filter

\[ \hat{x}_{t+1} = (A - KC)\hat{x}_t + Ky_t \]  
\[ a_t = y_t - C\hat{x}_t \]  

- A whitening filter takes a sequence of \( y \)'s and gives as output a sequence of \( a \)'s that are serially uncorrelated.
- We 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
  \[ x_t = -(\kappa \theta)^{-1} k_t - \varepsilon_{\tau,t}, \quad y_t = -(\kappa \theta)^{-1} k_t \]
  \[ x_{t+1} = \alpha x_t + [\alpha + \theta^{-1}] \varepsilon_{\tau,t} \]
  \[ y_t = x_t + \varepsilon_{\tau,t} \tag{8} \]

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 \(\lambda\) 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 \(\lambda\) 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 \(\Sigma_{t_0}\)
  - 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

\[-(\kappa \theta)^{-1} k_t = [1 + \frac{(\alpha + \theta)}{(L^{-1} - \alpha)}] a_t\]

\[= \left[ \frac{L^{-1} + \theta}{L^{-1} - \alpha} \right] a_t\]

\[= \left[ \frac{1 + \theta L}{1 - \alpha L} \right] a_t\]

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

- 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 \(a_t\)