Identifying Taylor Rules*

David Backus,† Mikhail Chernov,‡ and Stanley Zin§

May 23, 2010

Abstract

We show in some relatively simple examples how identification problems arise naturally in forward-looking models when agents observe shocks but econometricians do not. In familiar setting, long-term interest rates are a potentially useful source of additional information. ...

JEL Classification Codes: E43, E52, G12.

Keywords: monetary policy; interest rates??.

---

*Preliminary and incomplete: no guarantees of accuracy or sense. We welcome comments, including references to related papers we inadvertently overlooked. This started with our reading of John Cochrane's paper and subsequent conversations with Mark Gertler.

† Stern School of Business, New York University, and NBER; dbackus@stern.nyu.edu.

‡ London Business School, London School of Economics, and CEPR; mchernov@london.edu.

§ Stern School of Business, New York University, and NBER; stan.zin@nyu.edu.

1 Introduction

2 An example

We work through an example in which identification of model parameters can be illustrated in relatively simple settings. Given a model, the central issue is what agents and econometricians observe. When econometricians observe less than agents, there are inevitable issues with identification. Forward prices can be helpful here, since they provide the econometrician with additional information. The logic follows Hansen and Sargent (1980, 1991) and related work.

Consider this example of a forward-looking economic model:

$$y_t \quad = \quad \lambda E_t y_{t+1} + x_t.$$

Here y is an endogenous forward-looking variable, x is a shock, and λ is a parameter whose absolute value is less than one. We'll use two specifications for the shocks x. The first is an arbitrary scalar moving average,

$$x_t \quad = \quad \sum_{j=0}^{\infty} \chi_j w_{t-j} \quad = \quad \chi(L) w_t,$$

where x is stationary (square summable) and $\{w_t\} \sim \text{NID}(0, 1)$. The second is a state-space representation

$$x_{t+1} \quad = \quad A x_t + C w_{t+1}, \tag{1}$$

where A is stable (eigenvalues with absolute values less than one) and $\{w_t\} \sim \text{NID}(0, I)$. [dimensions?]

Under these conditions, y has a unique stationary solution. In the scalar moving average case, the solution is

$$y_t \quad = \quad \sum_{j=0}^{\infty} \lambda^j E_t x_{t+j} \quad = \quad \psi(L) w_t,$$

with

$$\psi_j \quad = \quad \sum_{i=0}^{\infty} \lambda^i \chi_{j+i}$$

and $j \geq 0$. See Appendix A. [Fix up notation, add vector process...]

We want to estimate the parameters, esp λ. Identification depends on what we observe. We assume throughout that the agents in this economy observe x and both agents and econometricians observe y. The question is whether we, the econometricians, can estimate the parameters. To see how this might work, suppose x is MA(1):

$$x_t = \chi_0 w_t + \chi_1 w_{t-1} = \chi_0 (w_t + \theta w_{t-1}).$$

Then $\psi_0 = \chi_0 + \lambda \chi_1$ and $\psi_1 = \chi_1$.

Here are some issues that come up:

- Suppose econometricians observe x and y and $\chi(L)$ is invertible (all the roots of $\chi(z) = 0$ are greater than one in absolute value). In the MA(1) case, this corresponds to $|\theta| < 1$. Then we can estimate χ from x and ψ (hence λ) from y.

  Example.

- Under the same conditions, suppose $\chi(L)$ is not invertible. Then the Wold representation for x (projection on past xs) is based on a smaller information set than the history of the xs. If agents and econometricians have the same information set, then it's the same situation as above, but using the Wold rep in place of the noninvertible MA. But if agents observe the ws and econometricians only observe x, then econometricians will misindentify the innovations. Which will lead to misidentification of λ.

  Example.

- Forward rates as sources of information. Suppose forward contracts are tied to future values of y and are priced at their conditional means: $f_t^j = E_t y_{t+j}$. This has the flavor of the term structure under the expectations hypothesis. Are observations of fs useful in identifying λ? [Given a spanning condition, forward rates should provide enough extra information to bail us out.]

- Vector process with partial observability. Suppose x is a vector with state-space representation

  $$x_{t+1} = A x_t + C w_{t+1}, \qquad (2)$$

  where A is stable (eigenvalues with absolute values less than one) and $\{w_t\} \sim \text{NID}(0, I)$. [dimensions?] Suppose $x_t = [x_{1t}, x_{2t}]$ has two components, and that agents observe both but econometricians only observe the first. Then we're back in the same situation as above...

3 Taylor rules in Cochrane's example

The simplest example in Cochrane (2007) consists of two equations:

$$i_t = r + E_t p_{t+1} + x_{1t} \qquad (3)$$

$$i_t = r + \tau p_t + x_{2t}, \qquad (4)$$

[Switch to vector version, which imposes invertibility on the overall system but not on subsets.]

a vector process for x given by (2). The variables are the nominal interest rate i and inflation p. Think of the first equation as a simple version of an Euler equation (EE) and the second as a Taylor rule (TR). We'll set the real interest rate $r = 0$ for now.

Solution. Equations (3) and (4) give us the stochastic difference equation

$$E_t p_{t+1} \;=\; \tau p_t - x_{1t} + x_{2t}. \tag{5}$$

We're looking for a stationary solution for p. If $|\tau| > 1$, there's a unique stationary solution of the form $p_t = \pi_1(L) w_{1t} + \pi_2(L) w_{2t}$. Eq (5) tells us the coefficients satisfy

$$\begin{aligned} \pi_{1j+1} &= \tau \pi_{1j} - \chi_{1j} \\ \pi_{2j+1} &= \tau \pi_{2j} + \chi_{2j} \end{aligned}$$

for $j \geq 0$. Consider the coefficients $\{\pi_{1j}\}$. Given an initial value π_{10}, the recursion generates coefficients

$$\begin{aligned} \pi_{11} &= \tau \pi_{10} - \chi_{10} \\ \pi_{12} &= \tau \pi_{11} - \chi_{11} \;=\; \tau^2 \pi_{10} - (\chi_{10} + \tau \chi_{11}) \\ \pi_{1j} &= \tau^j \left(\pi_{10} - \tau^{-1} \sum_{i=0}^{j-1} \tau^{-i} \chi_{1i} \right) \end{aligned}$$

For the coefficients to converge (a requirement of stationarity), we need $\pi_{10} = \tau^{-1} \sum_{j=0}^{\infty} \tau^{-j} \chi_{1j} = -\tau^{-1} \chi_1(\tau^{-1})$, which implies

$$\pi_{1j} \;=\; \tau^{j-1} \sum_{i=j}^{\infty} \tau^{-i} \chi_{1i}$$

for $j \geq 0$. The solution for the second component is the same except for the sign:

$$\pi_{2j} \;=\; -\tau^{j-1} \sum_{i=j}^{\infty} \tau^{-i} \chi_{2i}$$

Thus: given parameters $\{\tau, \chi_{ij}\}$ we can derive the coefficients $\{\pi_{ij}\}$ of the inflation process. [Comment: Related Hansen-Sargent formulas for the solution are given in an appendix.]

Identification. Suppose we observe inflation (p) and the interest rate (i), but not the shocks (x_1 and x_2). Can we identify the Taylor rule parameter (τ)? The difficulty is that the behavior of p and i combines the behavior of the shocks and the TR parameter, and it's not clear we can disentangle them. Here are some example.

- Special case 1 ($x_2 = 0$). Identification follows directly from estimating the Taylor rule (4), which is exact in this case. Both the lhs and rhs are observable.

  A little algebra shows how this works. Since the model has a univariate shock, we can estimate $\pi_1(L)$ from a long enough time series for p. From this process, we can infer the coefficients for expected inflation ($[\pi_1(L)/L]_+$ in Hansen-Sargent notation). The interest rate and the Euler equation then tell us x_1 and its coefficients $\chi_1(L)$. The coefficients of the lhs ($[\pi_1(L)/L]_+ + \chi_1(L)$) are τ times those of the rhs ($\pi_1(L)$):

$$w_{t-j}: \qquad \pi_{1j+1} + \chi_{1j} \;=\; \tau^j \sum_{i=j+1}^{\infty} \tau^{-i} \chi_{1i} + \chi_{1j} \;=\; \tau \pi_{1j}.$$

  [This holds by construction; verification here simply tells us we did the calculations right.]

- Special case 2 ($x_1 = 0$). This is the example Cochrane looks at. Again, we observe p and i. The former gives us the inflation coefficients π_2. The latter gives us no new information, since with $x_1 = 0$ expected inflation is implied by the inflation process. The result is that we can't isolate the roles of p_t and x_{2t} in the Taylor rule (4), hence can't estimate τ.

  Here's an example. Suppose x_2 is MA(1). Then p is MA(1), too. The inflation coefficients are

$$\begin{aligned} \pi_{20} &= -\tau^{-1}(\chi_{20} + \tau^{-1}\chi_{21}) \\ \pi_{21} &= -\tau^{-1}\chi_{21}. \end{aligned}$$

  We can estimate the π's from inflation data, but we can't disentangle the impact of the shock (the χ's) from the policy parameter (τ). That's true even if x_2 is white noise ($\chi_{2j} = 0$ for $j \geq 1$): the inflation coefficient is $\pi_{20} = -\tau^{-1}\chi_{20}$ and we can't separate the two components. Alternatively, let x_2 be AR(1), so that $\chi_{2j} = \varphi^j \chi_{20}$. In this case, $\pi_{2j} = -\varphi^j \chi_{20}/(\tau - \varphi)$, so inflation is AR(1) with the same autoregressive parameter. Inflation and the shock are perfectly correlated, so there's no way we can disentangle their effects in the Taylor rule. An inflation process can be reconciled with any choice of τ we like by adjusting χ_{20}.

- Shocks in both equations. One of the lessons here is getting extra information out of the interest rate (special case 1). If $x_1 = 0$ we can't do that. Another is that the TR shock x_2 shows up in p, which makes it difficult to separate their effects in the TR (special case 2). The question is how far we can go if we have shocks in both equations.

  Gertler's example shows how we might handle two shocks. Let the two state variables be AR(1):

$$x_{it} \;=\; \varphi_i x_{it-1} + w_{it}.$$

with $\varphi_2 = 0$ (we can generalize this later on). In either case, the state space is essentially (x_1, x_2), so we drop the infinite MA notation. The inflation process has the form $p_t = \pi_1 x_{1t} + \pi_2 x_{2t}$. Substituting into (5) and collecting terms gives us $\pi_1 = 1/(\tau - \varphi_1)$ and $\pi_2 = -1/\tau$. Observables are therefore

$$\begin{aligned} p_t &= [1/(\tau - \varphi_1)]x_{1t} - (1/\tau)x_{2t} \\ i_t &= [\tau/(\tau - \varphi_1)]x_{1t}. \end{aligned}$$

We can estimate φ_1 from the autocorrelation of i and τ from the ratio of the interest rate to expected inflation.

Here's a more formal approach to the same problem. The idea is to write the dynamics of the observables in terms of the dynamics of the shocks. Here if the shocks are a VAR, then so are the observables. The VAR for the observables can be estimated, so the question is whether we can use its estimated coefficients to uncover the TR parameter τ. Suppose the shocks follow

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

with $x = (x_1, x_2)$ and w vector white noise. In our example, $A = [\varphi_1, 0; 0, 0]$. Then the observables $y = (p, i)$ are a linear transformation of x: $y = Cx$. In our example,

$$C = \begin{bmatrix} 1/(\tau - \varphi_1) & -1/\tau \\ \tau/(\tau - \varphi_1) & 0 \end{bmatrix}.$$

Let's assume that C is invertible. Then in terms of observables, we also have a VAR, namely

$$y_{t+1} = CAC^{-1}y_t + CBw_{t+1} = A^* y_t + B^* w_{t+1}$$

Now to identification. We can estimate the autoregressive parameters A^* of the observables; the question is whether we can deduce τ from them. In the example,

$$A^* = \begin{bmatrix} 0 & \varphi_1/\tau \\ 0 & \varphi_1 \end{bmatrix},$$

so we can compute $\tau = a^*_{22}/a^*_{12}$. [This is tedious; we did it with Matlab's Maple toolbox.]

This line of thought can easily be generalized, although the expressions can get complicated. Let the expectational difference equation be

$$E_t p_{t+1} = \tau p_t - u_1^\top x_t + u_2^\top x_t,$$

where u_1 and u_2 are known vectors. (In our example, u_i picks out the ith element of x.) We guess $p_t = a^\top x_t$ and derive

$$a^\top = (u_1 - u_2)^\top (\tau I - A)^{-1}. \tag{6}$$

The observables are then

$$y_t \;=\; \begin{bmatrix} p_t \\ i_t \end{bmatrix} \;=\; \begin{bmatrix} a^\top \\ a^\top A + u_1^\top \end{bmatrix} x_t \;=\; C x_t.$$

We estimate A^* and then ask whether we can recover τ. By counting, you might guess that this won't work without some restrictions on A: we estimate 4 elements of A^*, which isn't enough to nail down the 4 elements of A plus τ. Clearly it works, as described, if all the elements of A but the first one are zero. Can we go beyond that? Suppose A is diagonal. This is horribly nonlinear, but the diagonal elements of A are the eigenvalues of A^*. [Remember: A and A^* are similar.] From there, we can find τ. For example [this courtesy of Matlab] the upper right element is

$$a_{12}^* \;=\; \frac{a_{11} - a_{22}}{\tau - a_{22}}.$$

As long as the eigenvalues aren't equal, we can find τ. [Problem: we don't know which eigenvalue is which element of A, so we get two possible estimates of τ.]

Could we make A triangular? Not clear. Also not clear whether we could use information on covariances: eg, assume B diagonal.

Extensions: (i) VAR(2), (ii) bond yields (EH), (iii) larger vector of x's.

4 Affine models

Which model? Essential? Or skip this?

Derive Wold decomp.

Show how Wold innovations differ.

Describe using a different information set.

5 A macro model

6 Conclusions

A Solving expectational difference equations

Scalar version. Here's a useful result from Hansen and Sargent (JEDC, 1980, p 14) and Sargent (*Macroeconomic Theory, 2e*, 1987, pp 303-304). An expectational difference equation with stationary forcing variable x generates a "geometric distributed lead":

$$\begin{aligned} y_t &= \lambda E_t y_{t+1} + x_t \\ &= \lambda E_t (\lambda E_{t+1} y_{t+2} + x_{t+1}) + x_t \\ &= \sum_{j=0}^{\infty} \lambda^j E_t x_{t+j}. \end{aligned}$$

If $x_t = \sum_{j=0}^{\infty} \chi_j w_{t-j} = \chi(L) w_t$, with w white noise, then what is y_t? A unique stationary solution $y_t = \psi(L) w_t$ exists if x is stationary and $|\lambda| < 1$, but what is $\psi(L)$?

Note how the distributed lead works. Conditional expectations of x have the form

$$E_t x_{t+j} = [\chi(L)/L^j]_+ w_t = \sum_{i=0}^{\infty} \chi_{j+i} w_{t-i}$$

(The subscript "+" means ignore negative powers of L.) Therefore the coefficient of w_{t-i} in the distributed lead is

$$\psi_i = \sum_{j=0}^{\infty} \lambda^j \chi_{i+j}.$$

This tells us, for example, that if x is MA(q), then so is y: if $\chi_j = 0$ for $j > q$, then $\psi_j = 0$ above the same limit.

There's a "lag notation" version that expresses the result in compact form. We're not sure whether it's all that useful for our purposes, but here it is. We're looking for a solution $y_t = \psi(L) w_t$ satisfying the expectational difference equation:

$$\psi(L) w_t = [\psi(L)/L]_+ w_t + \chi(L) w_t.$$

... [flesh this out]

See also Hansen and Sargent ("A note on Wiener-Kolmogorov prediction," ms, 1981).

Vector version. Here's a related result adapted from Ljungqvist and Sargent (*Recursive Macroeconomic Theory, 2e*, 2005, section 2.4). It extends the previous result to higher dimensional forcing processes that can be expressed as stationary vector autoregressions. Consider the system

$$\begin{aligned} y_t &= \lambda E_t y_{t+1} + u^\top x_t \\ x_{t+1} &= A x_t + B w_{t+1}, \end{aligned}$$

where u is an arbitrary vector and w is iid with mean zero and variance I. The solution in this case is

$$y_t \;=\; \sum_{j=0}^{\infty} \lambda^j u^\top E_t x_{t+j} \;=\; u^\top \sum_{j=0}^{\infty} \lambda^j A^j x_t \;=\; u^\top (I - \lambda A)^{-1} x_t.$$

[The last step follows from the matrix geometric series.]

There's a method of undetermined coefficients version of this. Guess $y_t = a^\top x_t$ for some vector a (we know the solution has this form from what we just did). Then the difference equation tells us

$$a^\top x_t \;=\; a^\top \lambda A x_t + u^\top x_t.$$

Collecting terms in x_t gives us $a^\top = u^\top (I - \lambda A)^{-1}$, as stated.

B Noninvertible moving averages

Consider the MA(1)

$$x_t \;=\; \chi_0 w_t + \chi_1 w_{t-1} \;=\; \chi_0 (w_t + \theta w_{t-1}).$$

Etc...

C VARs

Show if you have to project down, could get noninvertible...

Apply to state-space rep of noninvertible MA(1):

$$\begin{bmatrix} x_{t+1} \\ \chi_1 w_{t+1} \end{bmatrix} \;=\; \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_t \\ \chi_1 w_t \end{bmatrix} + \begin{bmatrix} \chi_0 \\ \chi_1 \end{bmatrix} w_{t+1}.$$

[??Do we need the χ_1 in the state vector?] See, for example, Hansen and Sargent (2005, Section 2.5.8).

Show how Kalman filter works if we only see x.

References

Backus, David K., and Stanley E. Zin, 1994, "Reverse engineering the yield curve," NBER Working Paper No. 4676, March.

Clarida, Richard, Jordi Gali, and Mark Gertler, 1999, "The science of monetary policy," *Journal of Economic Literature* 37, 1661-1707.

Cochrane, John H., 2007, "Identification with Taylor rules: A critical appraisal," manuscript, September.

Collin-Dufresne, Pierre, and Robert S. Goldstein, 2002, "Do bonds span the fixed income markets? Theory and evidence for unspanned stochastic volatility," *Journal of Finance* 57, 1685-1730.

Hansen, Lars Peter, and Thomas J. Sargent, 1980, "Forumulating and estimating dynamic linear rational expectations models," *Journal of Economic Dynamics and Control* 2, 7-46.

Hansen, Lars Peter, and Thomas J. Sargent, 1991, *Rational Expectations Econometrics*, San Francisco: Westview Press.

Hansen, Lars Peter, and Thomas J. Sargent, 2005, *Recursive Models of Dynamic Linear Economies* manuscript.

Hansen, Lars Peter, and Thomas J. Sargent, 2005, *Recursive Models of Dynamic Linear Economies* manuscript.

Rondina, Giacomo, and Todd B. Walker, 2009, "Information equilibria in dynamic economies," manuscript, November.

Sargent, Thomas J., 1987, *Macroeconomic Theory (2nd edition)* New York: Academic Press.