Learning and Monetary Policy

Lecture 1 – Introduction to Expectations and Adaptive Learning

George W. Evans (University of Oregon)

University of Paris X -Nanterre (September 2007)

J. C. Trichet: "Understanding expectations formation as a process underscores the strategic interdependence that exists between expectations formation and economics." (Zolotas lecture, 2005)

Ben S. Bernanke: "In sum, many of the most interesting issues in contemporary monetary theory require an analytical framework that involves learning by private agents and possibly the central bank as well." (NBER, July 2007).

Outline of Lectures

Lecture 1, Sept. 11, 10am - noon. Introduction to Expectations and Adaptive Learning. Convergence of least-squares learning to RE in natural-rate and cobweb models. Recursive algorithms, stochastic approximation and E-stability. Applications to simple monetary models.

Lecture 2, Sept. 11, 2 - 4pm. The New Keynesian Model of Monetary Policy: determinacy and stability under learning. The New Keynesian Model. Interest-rate rules and the possibility of multiple equilibria. Determinacy and stability under learning for Taylor-type interest-rate rules. Stability of sunspot equilibria under learning.

Lecture 3, Sept. 13, 10am - noon. Optimal Monetary Policy and Learning in the New Keynesian Model. Optimal discretionary policy and optimal policy with commitment. Stability under private-agent learning for alternative implementations of optimal policy. Structural parameter learning by policymakers. Robust policy under parameter uncertainty.

Lecture 4, Sept. 13, 2 - 4pm. Recent Applications of Learning to Monetary Policy. Perpetual learning, persistence, recurrent hyperinflations, liquidity traps and deflation, dynamic predictor selection and endogenous volatility.

Lecture 1 Outline

- Introduction to expectations and LS learning
- The Muth-Lucas model
- LS learning and E-stability
- Recursive LS, stochastic recursive algorithms and the E-stability principle
- Application to the Muth-Lucas model
- Other examples. Cagan model of inflation. Selection criterion. Sunspot equilibria.

Introduction

- Expectations play a key role in macroeconomics:
  (i) The private sector is forward-looking (e.g. investment, savings decisions)
  (ii) Forecasts (including private forecasts) of future inflation and output have a key role in monetary policy.

- Theories of expectation have evolved:
  Naive or static (1930s)
  Adaptive Expectations (1950s and 1960s)
  RE (1970s and 1980s). RE is the benchmark assumption,
  Learning (1990s $\rightarrow$)

- Since Lucas (1972, 1976) and Sargent (1973) the standard assumption in the theory of economic policy is rational expectations (RE). This assumes, for both private agents and policymakers,

  – knowledge of the correct form of the model

  – knowledge of all parameters, and

  – knowledge that other agents are rational & know that others know . . . .

- RE assumes too much and is therefore implausible. We need an appropriate model of **bounded rationality** What form should this take?

- My general answer is given by the **Cognitive Consistency Principle**: economic agents should be about as smart as (good) economists. Economists forecast economic variables using econometric techniques, so a good starting point: **model agents as "econometricians." (**There are other possibilities).

- Neither private agents nor economists at central banks do know the true model. Instead economists formulate and estimate models. These models are re-estimated and possibly reformulated as new data becomes available. Economists engage in *processes of learning* about the economy. This process **may or may** not converge to RE

Fundamental Issues Addressed by Learning:

- When private agents follow a learning rule there is the possibility that the REE of interest may exhibit **instability under learning**.

- In some models there are multiple equilibria under RE. Learning can then act as a **selection criteria**.

- The **learning dynamics** themselves may be of interest: either just the transitional dynamics or there may be persistent learning dynamics.

A Muth-Lucas-type Model

In this lecture we will focus on what is in many ways the simplest reduced form. For the detailed analysis of this model see EH2001, Chapter 2. The reduced form is:

$$p_t = \mu + \alpha E^*_{t-1} p_t + \delta' w_{t-1} + \eta_t. \tag{RF}$$

Here $E^*_{t-1} p_t$ denotes expectations of p_t formed at $t-1$, w_{t-1} is a vector of exogenous shocks observed at $t-1$, and η_t is an exogenous unobserved iid shock.

We assume $\alpha \neq 1$ and that w_t follows an exogenous stationary VAR process.

Muth example. The structural model consists of demand and supply equations:

$$\begin{aligned} d_t &= m_I - m_p p_t + v_{1t} \\ s_t &= r_I + r_p E^*_{t-1} p_t + r'_w w_{t-1} + v_{2t}, \end{aligned}$$

Assuming market clearing, $s_t = d_t$, yields the reduced form where $\mu = (m_I - r_I)/m_p$, $\delta = -m_p^{-1} r_w$ and $\alpha = -r_p/m_p$ and $\eta_t = (v_{1t} - v_{2t})/m_p$.

Note that $\alpha < 0$ if $m_p, r_p > 0$.

Lucas-type Monetary model. A simple Lucas-type model:

$$q_t = \bar{q} + \pi(p_t - E^*_{t-1}p_t) + \zeta_t,$$

where $\pi > 0$, and aggregate demand function is given by

$$\begin{aligned} m_t + v_t &= p_t + q_t, \\ v_t &= \mu + \gamma' w_{t-1} + \xi_t, \\ m_t &= \bar{m} + u_t + \rho' w_{t-1}. \end{aligned}$$

Here w_{t-1} are exogenous observables. The reduced form is again

$$\begin{aligned} p_t &= \mu + \alpha E^*_{t-1}p_t + \delta' w_{t-1} + \eta_t, \text{ where} \\ \alpha &= \pi(1+\pi)^{-1} \text{ and } \delta = (1+\pi)^{-1}(\rho + \gamma) \end{aligned}$$

In this example $0 < \alpha < 1$.

RATIONAL EXPECTATIONS

First consider the model under RE:

$$p_t = \mu + \alpha E_{t-1} p_t + \delta' w_{t-1} + \eta_t.$$

The model has a unique RE solution since

$$\begin{aligned} E_{t-1} p_t &= \mu + \alpha E_{t-1} p_t + \delta' w_{t-1} \longrightarrow \\ E_{t-1} p_t &= (1-\alpha)^{-1} \delta + (1-\alpha)^{-1} \delta' w_{t-1} \end{aligned}$$

Hence the unique REE is

$$\begin{aligned} p_t &= \bar{a} + \bar{b}' w_{t-1} + \eta_t, \text{ where} \\ \bar{a} &= (1-\alpha)^{-1} \delta \text{ and } \bar{b} = (1-\alpha)^{-1} \delta. \end{aligned}$$

LEAST-SQUARES LEARNING

Under learning, agents have the beliefs or perceived law of motion (PLM)

$$p_t = a + bw_{t-1} + \eta_t,$$

but a, b are unknown. At the end of time $t-1$ they estimate a, b by LS (Least Squares) using data through $t-1$, i.e. Then they use the estimated coefficients to make forecasts $E_{t-1}^* p_t$. Here the standard least squares (LS) formula are

$$\begin{aligned} \begin{pmatrix} a_{t-1} \\ b_{t-1} \end{pmatrix} &= \left(\sum_{i=1}^{t-1} z_{i-1} z_{i-1}' \right)^{-1} \left(\sum_{i=1}^{t-1} z_{i-1} p_i \right), \text{ where} \\ z_i' &= \begin{pmatrix} 1 & w_i' \end{pmatrix}. \end{aligned}$$

The timing is:

– End of $t-1$: w_{t-1} and p_{t-1} observed. Agents update estimates of a, b to a_{t-1}, b_{t-1} using $\{p_s, w_{s-1}\}_{s=1}^{t-1}$. Agents make forecasts

$$E_{t-1}^* p_t = a_{t-1} + b_{t-1}' w_{t-1}.$$

– Period t: (i) The shock η_t is realized, p_t is determined and w_t is realized. (ii) agents update estimates of a, b to a_t, b_t using $\{p_s, w_{s-1}\}_{s=1}^{t}$ and make forecasts

$$E_t^* p_{t+1} = a_t + b_t' w_t.$$

The system under learning is a fully specified dynamic system under learning.

Question: Will $(a_t, b_t) \to (\bar{a}.\bar{b})$ as $t \to \infty$?

Theorem: Consider model (RF) with $E^*_{t-1}p_t = a_{t-1} + b'_{t-1}w_{t-1}$ and with a_{t-1}, b_{t-1} updated over time using least-squares. If $\alpha < 1$ then $\begin{pmatrix} a_t \\ b_t \end{pmatrix} \rightarrow \begin{pmatrix} \bar{a} \\ \bar{b} \end{pmatrix}$ with probability 1. If $\alpha > 1$ convergence occurs with probability 0.

Thus the REE is stable under LS learning for both examples.

Example of an unstable REE: Muth model with $m_p < 0$ (Giffen good) and $|m_p| < r_p$.

E-STABILITY

Proving this theorem is not easy. However, there is an easy way of deriving the stability condition $\alpha < 1$ that is quite general. Start with the PLM

$$p_t = a + b' w_{t-1} + \eta_t,$$

and consider what would happen if (a, b) were fixed at some value possibly different from the RE values $(\bar{a}, \bar{b})$. The corresponding expectations are

$$E^*_{t-1} p_t = a + b' w_{t-1},$$

which would lead to the Actual Law of Motion (ALM)

$$p_t = \mu + \alpha(a + b' w_{t-1}) + \delta' w_{t-1} + \eta_t.$$

The implied ALM gives the mapping T: PLM $\longrightarrow$ ALM:

$$T \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \mu + \alpha a \\ \delta + \alpha b \end{pmatrix}.$$

The REE $\bar{a}, \bar{b}$ is a fixed point of T. Expectational-stability ("E-stability) is defined by the differential equation

$$\frac{d}{d\tau}\begin{pmatrix} a \\ b \end{pmatrix} = T\begin{pmatrix} a \\ b \end{pmatrix} - \begin{pmatrix} a \\ b \end{pmatrix}.$$

Here τ denotes artificial or notional time. $\bar{a}, \bar{b}$ is said to be E-stable if it is stable under this differential equation.

In the current case the T-map is linear. Component by component we have

$$\begin{aligned} \frac{da}{d\tau} &= \mu + (\alpha - 1)a \\ \frac{db_i}{d\tau} &= \delta + (\alpha - 1)b_i \text{ for } i = 1, ..., p. \end{aligned}$$

$$\frac{da}{d\tau} = \mu + (\alpha - 1)a$$
$$\frac{db_i}{d\tau} = \delta + (\alpha - 1)b_i \text{ for } i = 1, ..., p.$$

It follows that the REE is E-stable if and only if $\alpha < 1$. This is the stability condition, given in the theorem, for stability under LS learning.

Intuition: under LS learning the parameters a_t, b_t are slowly adjusted, on average, in the direction of the corresponding ALM parameters.

We will next outline the techniques used to prove the theorem.

RATIONAL VS. REASONABLE LEARNING

First, a remark. LS learning is boundedly rational but not fully rational. Under LS learning the true process followed by p_t is given by

$$\begin{aligned} p_t &= \mu + \alpha(a_{t-1} + b'_{t-1}w_{t-1}) + \delta' w_{t-1} + \eta_t, \text{ or} \\ p_t &= (\mu + \alpha a_{t-1}) + (\delta + \alpha b_{t-1})' w_{t-1} + \eta_t, \end{aligned}$$

The coefficients are thus time-varying, not constant.

However LS learning, though not fully rational, appears "reasonable" [Margaret Bray]. For $\alpha < 1$, since then $a_t, b_t \to \bar{a}, \bar{b}$, the agents' misspecification disappears asymptotically.

RECURSIVE FORMULATION

Letting $\phi_t = (a_t, b_t')'$, LS updating can be written recursively as

$$\begin{aligned} \phi_t &= \phi_{t-1} + t^{-1} R_t^{-1} z_{t-1}(p_t - \phi_{t-1}' z_{t-1}) \\ R_t &= R_{t-1} + t^{-1}(z_{t-1} z_{t-1}' - R_{t-1}), \end{aligned}$$

and under learning p_t is given by

$$\begin{aligned} p_t &= (\mu + \alpha a_{t-1}) + (\delta + \alpha b_{t-1})' w_{t-1} + \eta_t, \text{ or} \\ p_t &= T(\phi_{t-1})' z_{t-1} + \eta_t. \end{aligned}$$

Combining equations gives

$$\begin{aligned} \phi_t &= \phi_{t-1} + t^{-1} R_t^{-1} z_{t-1}(z_{t-1}'(T(\phi_{t-1}) - \phi_{t-1}) + \eta_t) \\ R_t &= R_{t-1} + t^{-1}(z_{t-1} z_{t-1}' - R_{t-1}). \end{aligned}$$

This is a Stochastic Recursive System (SRA), is an example of an SRA,

$$\theta_t = \theta_{t-1} + \gamma_t Q(t, \theta_{t-1}, X_t),$$

where θ_t is a vector of parameter estimates, X_t is the state vector and γ_t is a deterministic sequence of "gains". In X_t can depend on θ_{t-1}.

Here $\theta_{t-1} \leftrightarrow \phi_{t-1}, R_t$, $X_t \leftrightarrow z_t, z_{t-1}, \eta_t$ and $\gamma_t \leftrightarrow t^{-1}$.

The "stochastic approximation" approach associates an ordinary differential equation (ODE) with the SRA,

$$\frac{d\theta}{d\tau} = h(\theta(\tau)),$$

where $h(\theta)$ is obtained as

$$h(\theta) = \lim_{t \to \infty} EQ(t, \theta, X_t),$$

provided this limit exists. If X_t depends on θ_{t-1}, then instead

$$h(\theta) = \lim_{t \to \infty} EQ(t, \theta, \bar{X}_t(\theta)),$$

The stochastic approximation results are:

Under suitable assumptions, if $\bar{\theta}$ is a locally stable equilibrium point of the ODE then $\bar{\theta}$ is a possible point of convergence of the SRA. If $\bar{\theta}$ is not a locally stable equilibrium point of the ODE then $\bar{\theta}$ is not a possible point of convergence of the SRA, i.e. $\theta_t \to \bar{\theta}$ with probability 0.

Discuss "suitable assumptions" and "possible point of convergence". See Ch. 6 of EH (2001) for details.

APPLICATION TO THE MUTH-LUCAS MODEL

The SRA is

$$\begin{aligned} \phi_t &= \phi_{t-1} + t^{-1} R_t^{-1} z_{t-1}(z'_{t-1}(T(\phi_{t-1}) - \phi_{t-1}) + \eta_t) \\ R_{t+1} &= R_t + t^{-1} \frac{t}{t+1}(z_t z'_t - R_t). \end{aligned}$$

and we identify θ_t with the components of ϕ_t, R_{t+1}.

The regularity assumptions are satisfied. The ODE is computed to be

$$\begin{aligned} \frac{d\phi}{d\tau} &= R^{-1} M(T(\phi) - \phi) \\ \frac{dR}{d\tau} &= M - R \end{aligned}$$

where $M = E z_t z'_t$. The fixed point is $\phi = \bar{\phi} \equiv (\bar{a}, \bar{b}')'$ and $R = M$.

It is easy to see that $R(\tau) \to M$ globally. But then stability stability of the ODE reduces to stability of

$$\frac{d\phi}{d\tau} = T(\phi) - \phi.$$

But $\phi = \begin{pmatrix} a \\ b \end{pmatrix}$, so this is just the **E-stability differential equation**, and E-stability holds if and only if $\alpha < 1$.

Our conclusions: (i) stochastic approximation techniques can be used to prove the stated Theorem on convergence of LS learning.
(ii) E-stability governs convergence under LS learning.

Numerical simulation of learning in Muth model. $\mu = 5$, $\delta = 1$ and $\alpha = -0.5$. $w_t \overset{iid}{\sim} N(0, 1)$ and $\eta_t \overset{iid}{\sim} N(0, 1/4)$. Initial values $a_0 = 1, b_0 = 2$ and $R_0 =$ eye(2). Convergence to the REE $\bar{a} = 10/3$ and $\bar{b} = 2/3$ is rapid.

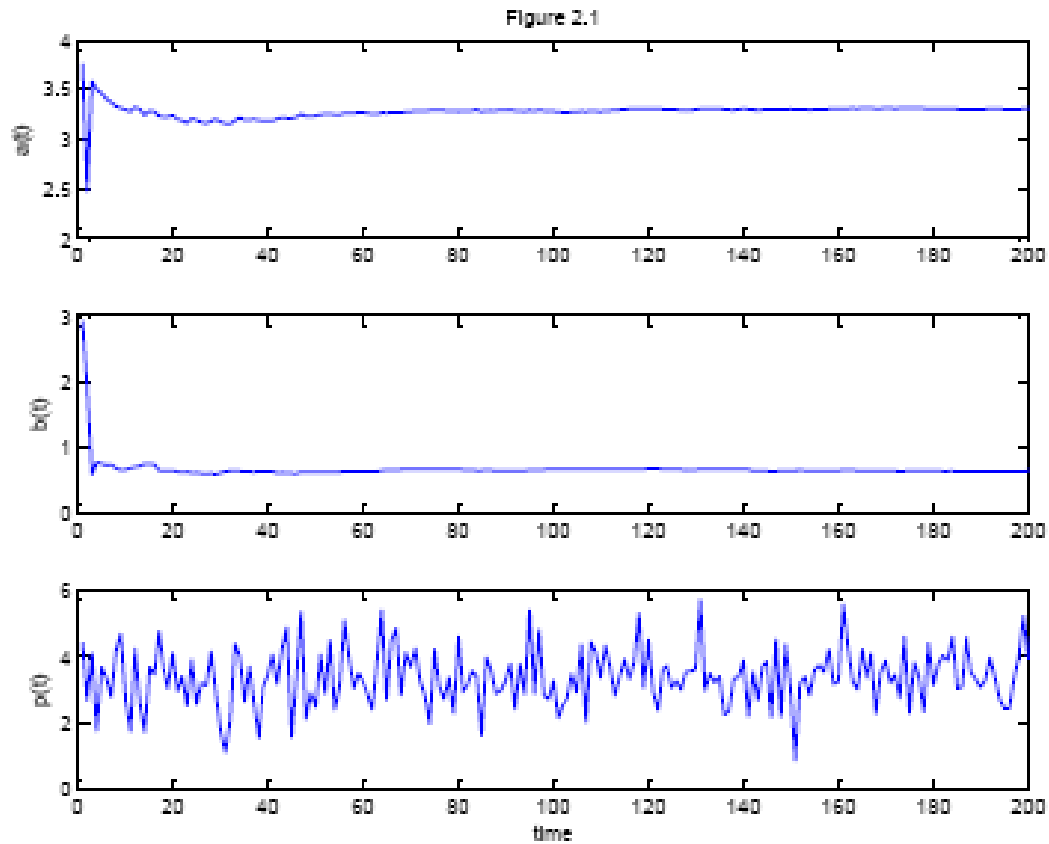


Figure 2.1
a(t)
b(t)
p(t)
time

THE E-STABILITY PRINCIPLE

– To study convergence of LS learning to an REE, specify a PLM with parameters ϕ. The PLM can be thought of as an econometric forecasting model. The REE is the PLM with $\phi = \bar{\phi}$.

– PLMs can take the form of ARMA or VARs or admit cycles or a dependence on sunspots.

– Compute the ALM for this PLM. This gives a map

$$\phi \to T(\phi),$$

with fixed point $\bar{\phi}$.

– E-stability is determined by local asymptotic stability of $\bar{\phi}$ under

$$\frac{d\phi}{d\tau} = T(\phi) - \phi.$$

The E-stability condition is that all roots of $DT(\bar{\phi}) - I$ have negative real parts. Equivalently, the real parts of all eigenvalues of $DT(\bar{\phi})$ must be less than 1.

– The E-stability principle: E-stability governs local stability of an REE under LS and closely related learning rules.

DISCUSSION OF E-STABILITY

– For some economic models the validity of the E-stability principle can be proved. In other cases it can be verified numerically.

– I regard the E-stability principle therefore as an operating hypothesis. The full extent of its validity remains to be determined. It appears to hold in a very wide range of economic models, e.g. RBC models, OG models with sunspots or cycles, New Keynesian models, open-economy macro models.

– The E-stability principle can be adapted to hold for misspecification issues, e.g. overparametrization (strong E-stability) or underparameterization (convergence to a "Restricted Perceptions Equilibrium").

– The E-stability principle also holds for "constant-gain" versions of LS in which there is incomplete convergence to an REE.

ADAPTIVE LEARNING VS. EDUCTIVE LEARNING

– **LS learning is adaptive**: it occurs in real time in response to forecast errors as data is accumulated.

– Another time of learning is "eductive." This occurs in "mental time" as a result of a reasoning process. In some models one can argue that if all agents know that Ep_t will be within a certain distance of its RE value, then in fact Ep_t will be strictly closer. Under common knowledge of rationality, one can then deduce that Ep_t **is** at its RE value.

– In such cases the REE is said to be **eductively stable**. See Guesnerie (1992, 2002), Evans and Guesnerie (1993, 2003, 2005), and others.

– In the cobweb model Guesnerie (1992) showed that with a homogeneous structure the condition for eductive stability is $|\alpha| < 1$. This condition can be obtained from "**iterative E-stability**," which requires that

$$\phi_{N+1} = T(\phi_N), \text{ for } N = 0, 1, 2, \ldots.$$

converges to the RE fixed point $\bar{\phi}$.

– Evans-Guesnerie showed more generally that iterative E-stability is a necessary condition for eductive stability. Iterative E-stability is stronger than E-stability: we need all roots of $DT(\bar{\phi})$ lie inside the unit circle.

– Eductive stability treats agents as economic theorists. It seems more appropriate in models with transparent structures. In the lectures I will focus on adaptive/LS learning.

Sargent-Wallace "ad hoc" model

There are many other univariate examples, e.g. the Sargent and Wallace (1975), AS-IS-LM model:

$$\begin{aligned}
q_t &= a_I + a_p(p_t - E^*_{t-1}p_t) + u_{1t}, \text{ where } a_p > 0, \\
q_t &= b_I + b_r(r_t - (E^*_{t-1}p_{t+1} - E^*_{t-1}p_t)) + u_{2t}, \text{ where } b_r < 0, \\
m &= c_I + p_t + c_q q_t + c_r r_t + u_{3t}, \text{ where } c_q > 0, \ c_r < 0.
\end{aligned}$$

The model can be solved to yield the reduced form of the price level

$$p_t = \alpha + \beta_0 E^*_{t-1}p_t + \beta_1 E^*_{t-1}p_{t+1} + v_t,$$

where $E_{t-1}v_t = 0$ and

$$\begin{aligned}
\beta_0 &= (a_p(1 + b_r c_q c_r^{-1}) + b_r)/(a_p(1 + b_r c_q c_r^{-1}) + b_r c_r^{-1}) \\
\beta_1 &= (1 - \beta_0)/(1 - c_r^{-1}).
\end{aligned}$$

$$p_t = \alpha + \beta_0 E^*_{t-1} p_t + \beta_1 E^*_{t-1} p_{t+1} + v_t.$$

For Sargent-Wallace $\beta_1 > 0$ and $\beta_0 + \beta_1 < 1$. The MSV solution is

$$\begin{aligned} p_t &= \bar{a} + v_t, \text{ where} \\ \bar{a} &= \alpha/(1 - \beta_0 - \beta_1). \end{aligned}$$

For the PLM $p_t = a + v_t$ the solution is stable under learning.

Taylor (1977) has a version with m_t in the AS curve that leads to indeterminacy. SSEs exist and can sometimes be stable under learning.

Cagan Model

The Cagan model of inflation is

$$m_t - p_t = -\gamma(E_t^* p_{t+1} - p_t) + \eta_t$$

where money supply m_t is exogenous. This can be put in the form

$$p_t = \beta E_t^* p_{t+1} + \kappa m_t + v_t,$$

where $0 < \beta = \gamma/(1+\gamma) < 1$ when $\gamma > 0$. In linearized OG models, $\gamma < 0$ and hence $\beta < 0$ or $\beta > 1$ is possible.

Other economic models fit this framework:
– Asset pricing with risk-neutrality
– PPP model of exchange rates

In some cases lagged p_t appears, e.g. in the Cagan model if

$$m_t = m + dp_{t-1} + e_t.$$

The general model

$$p_t = \beta E_t^* p_{t+1} + \delta p_{t-1} + v_t,$$

is of interest because of its one-step forward, one-step backward structure. Under RE the MSV solutions take the form

$$p_t = \bar{c} + \bar{a} p_{t-1} + \bar{k} v_t.$$

Here $\bar{a}$ satisfies

$$\beta a^2 - a + \delta = 0,$$

with roots

$$a_1 = (2\beta)^{-1} \left(1 - \sqrt{1 - 4\beta\delta}\right) \text{ and } a_2 = (2\beta)^{-1} \left(1 + \sqrt{1 - 4\beta\delta}\right)$$

If $|\beta + \delta| < 1$ then only the a_1 solution gives a stationary (non-explosive) solution. In some cases both solutions are stationary, as are sunspot solutions.

Using PLMs

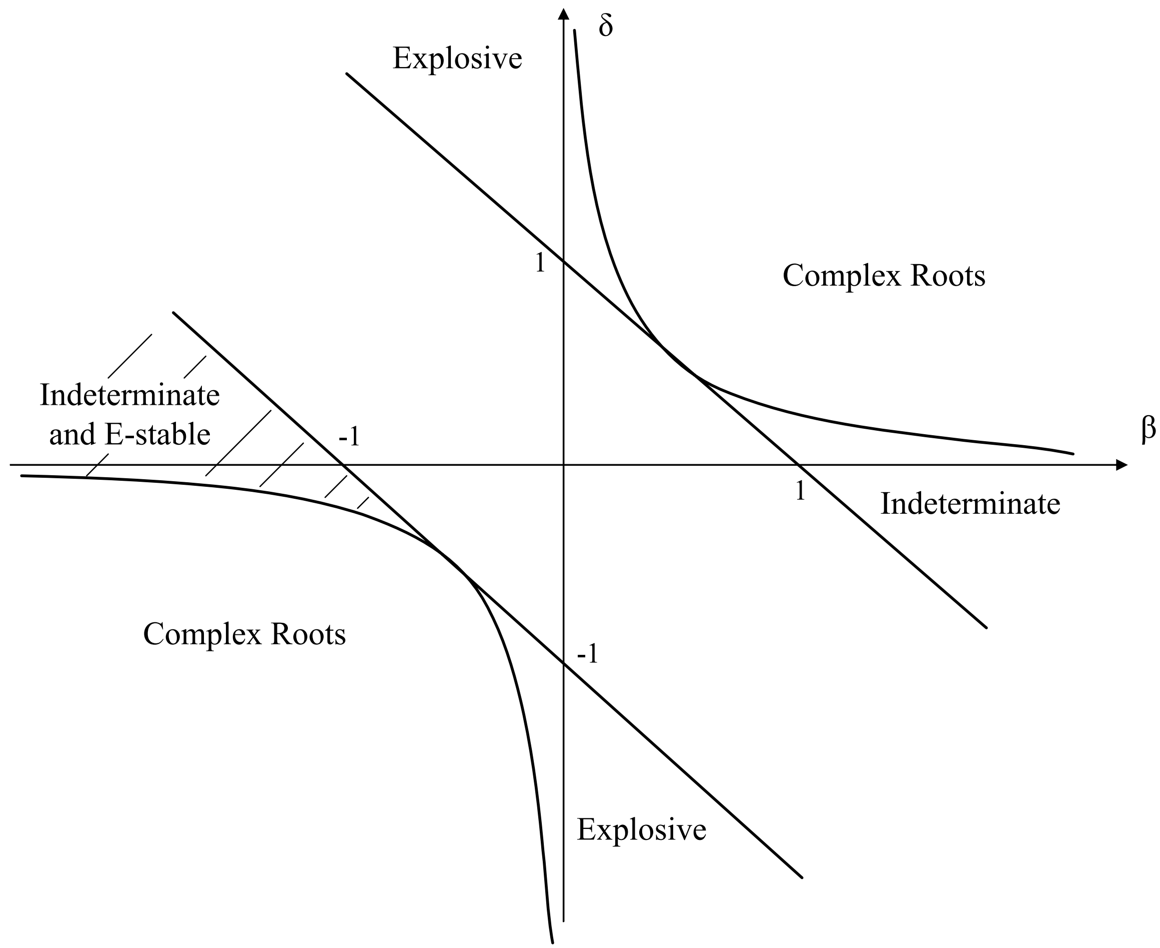
$$p_t = c + ap_{t-1} + kv_t$$

we can check whether under LS learning $(c_t, a_t, k_t) \rightarrow (\bar{a}, \bar{b}, \bar{k})$ locally.

If there are two stationary solutions of this form, E-stability can select between them. We can also look at the stability of SSEs (stationary sunspot equilibria)

$$p_t = c + ap_{t-1} + kv_t + d\xi_t,$$

where ξ_t is a suitable AR(1) sunspot. See Evans and McGough (2005a).


δ
Explosive
1
Complex Roots
Indeterminate
and E-stable
-1
β
1
Indeterminate
Complex Roots
-1
Explosive

CONCLUSIONS

- RE is very strong, unless supplemented by an account of how it can be reached by boundedly rational agents.

- The cognitive consistency principle suggests modeling economic agents as econometricians, forecasting using estimated models with parameter estimates that are updated over time.

- The E-stability principle governs converge

- In the Muth-Lucas model, we can show convergence to RE.

- More generally, multiple REE are possible. Stability under learning provides a selection criterion. Sometimes sunspot equilibria can be stable under learning.

- Next lecture: the New Keynesian model and monetary policy.