Adaptive Learning in Macroeconomics: some methodological issues

George W. Evans

University of Oregon and University of St. Andrews

FRB St. Louis Conference "Expectations in Dynamic Macroeconomic Models," August 2 - 4, 2012

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

Introduction
— Muth/Lucas model with LS learning. The E-stability principle
— Implications of learning for theory and policy

Methodological issues: overview of issues
— the planning horizon: short vs. long-horizon learning.
— structural knowledge: expected changes in policy and adaptive learning
— Examples.

Conclusions

The talk draws on Evans and Honkapohja "Learning as a Rational Foundation for Macroeconomics and Finance" in Frydman and Phelps, eds. (2013).

Introduction

– Macroeconomic models are usually based on optimizing agents in dynamic, stochastic setting and can be summarized by a **dynamic system,** e.g.

$$\begin{aligned} y_t &= Q(y_{t-1}, y^e_{t+1}, w_t) \\ \text{or } y_t &= Q\left(y_{t-1}, \left\{y^e_{t+1}\right\}_{j=0}^{\infty}, w_t\right) \end{aligned}$$

$y_t =$ vector of economic variables at time t (unemployment, inflation, investment, etc.), $y^e_{t+1} =$ expectations of these variables, $w_t =$ exogenous random factors at t. Nonstochastic models also of interest.

– The presence of **expectations** y^e_{t+1} makes macroeconomics inherently different from natural science. But **how are expectations formed?**

– Since Lucas (1972, 1976) and Sargent (1973) the standard assumption is **rational expectations** (RE).

– RE assumes too much knowledge & coordination for economic agents. We need a **realistic** model of **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, e.g.

- model agents like **economic theorists** – the **eductive** approach, or

- model them like **econometricians** – the **adaptive** (or evolutive) approach.

– In this talk I follow the adaptive approach. Agent/econometricians must select models, estimate parameters and update their models over time. I'll briefly mention connections to the eductive approach.

A Muth/Lucas-type Model

Consider a simple univariate reduced form:

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

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

Muth cobweb example. Demand and supply equations:

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

$s_t = d_t$, yields (RF) where $\alpha = -r_p/m_p < 0$ if $r_p, m_p > 0$.

Lucas-type monetary model. AS + AD + monetary feedback:

$$\begin{aligned} q_t &= \bar{q} + \pi(p_t - E^*_{t-1} p_t) + \zeta_t, \\ m_t + v_t &= p_t + q_t \text{ and } m_t = \bar{m} + u_t + \rho' w_{t-1} \end{aligned}$$

leads to yields (RF) with $0 < \alpha = \pi/(1+\pi) < 1$.

Rational Expectations vs. Least-Squares Learning

The model $p_t = \mu + \alpha E_{t-1} p_t + \delta' w_{t-1} + \eta_t$. has the **unique REE**

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

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

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

but a, b are unknown. At the end of time $t-1$ they estimate a, b by LS (Least Squares) using data through $t-1$. Then they use the estimated coefficients to make forecasts $E^*_{t-1} p_t$.

– 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

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

– **Temporary equilibrium at t:** (i) p_t is determined as

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

and w_t is realized. (ii) agents update estimates to a_t, b_t and forecast

$$E^*_t p_{t+1} = a_t + b'_t w_t.$$

The fully specified dynamic system under LS learning is written recursively as

$$\begin{aligned}
E^*_{t-1} p_t &= \phi'_{t-1} z_{t-1} \text{ where } \phi'_{t-1} = (a_{t-1}, b'_{t-1}) \text{ and } z'_{t-1} = (1, w_{t-1}) \\
p_t &= \mu + \alpha E^*_{t-1} p_t + \delta' w_{t-1} + \eta_t, \\
\phi_t &= \phi_{t-1} + t^{-1} R_t^{-1} z_{t-1} (p_t - \phi'_{t-1} z_{t-1}) \\
R_t &= R_{t-1} + t^{-1} (z_{t-1} z'_{t-1} - R_{t-1}),
\end{aligned}$$

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

Theorem (Bray & Savin (1986), Marcet & Sargent (1989)). Convergence to RE, i.e. $(a_t, b_t') \rightarrow (\bar{a}, \bar{b}')$ a.s. if $\alpha < 1$. If $\alpha > 1$ convergence with prob. 0.

Thus the REE is stable under LS learning both for Muth model ($\alpha < 0$) and Lucas model ($0 < \alpha < 1$), but is not stable if $\alpha > 1$. The stability condition can be obtained using the **E-stability principle** based on an associated ODE.

Instability arises for $\alpha > 1$ because economy under learning is **self-referential**.

LS learning is the most widely-used implementation of adaptive learning in stochastic models.

For a wide range of models **E-stability** has been shown to govern stability under LS learning, see Evans & Honkapohja (1992, 2001, etc.). The technique is general.

E-STABILITY

Proving the theorem relies on stochastic approximation theorems, but there is a simple way to obtain the stability condition $\alpha < 1$. Start with the PLM

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

and suppose (a, b) were fixed at some (possibly non-REE) value. Then

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

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 and it can be seen 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 parameter.

The E-Stability Principle

– The E-stability technique works quite generally.

– 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

$$\phi \rightarrow T(\phi),$$

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: eigenvalues of $DT(\bar{\phi})$ have real parts less than 1.

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

– E-stability can be used as a selection criterion in models with multiple REE.

– The techniques can be applied to multivariate linearized models, and thus to RBC, OLG, New Keynesian and DSGE models.

– Iterative E-stability, $\lim_{n\to\infty} T^n(\phi) = \bar{\phi}$, plays a role in eductive learning.

Constant gain learning dynamics

For **discounted LS** the "gain" t^{-1} is replaced by a (typically small) constant $0 < \gamma < 1$, e.g. $\gamma = 0.04$. Often called "constant gain" (or "perpetual") learning.

Especially plausible if agents are worried about structural change.

With constant gain in the Muth/Lucas and $\alpha < 1$ convergence of (a_t, b_t) is to a stochastic process around $(\bar{a}, \bar{b})$.

In the Cagan/asset-pricing model

$$p_t = \mu + \alpha E_t^* p_{t+1} + \delta w_t$$

constant gain learning leads to excess volatility, correlated excess return, etc.

General Implications of Learning Theory

- Can assess **plausibility** of RE based on stability under LS learning

- Use local stability under learning as a **selection criterion** in models with multiple REE

- **Persistent learning dynamics** that arise with modified or more general learning rules

- **Policy implications**: Policy should facilitate learning by private agents of the targeted REE.

Recent Methodological Issues

A variety of methodological issues have arisen:

- **Misspecification**. Like applied econometricians, agents may use misspecified models $\longrightarrow$ restricted perceptions equilibria, extension of E-stability principle.

- **Discounted LS & structural change.** Agents may be concerned about structural change and discount older data $\longrightarrow$ escape dynamics.

- **Heterogeneous expectations.** Can allow for heterogeneity of priors, econometric learning rules, inertia, forecasting models, etc.

- **Multiple forecasting models.** Dynamic predictor selection or Bayesian model averaging.

- **Degree of rationality.** Are agents fully rational or not (e.g. due to costs of optimizing or limited abilities)?

- **Planning horizon.** Infinitely-lived agents can engage in short-horizon decision making or use infinite-horizon learning.

- **Extent of structural knowledge.** Partial structural knowledge can be combined with adaptive learning.

In the remainder of this talk I will emphasize the last two issues: the planning horizon and structural knowledge. However, the issues listed are related.

The Planning Horizon

- In the Lucas-Muth model expectations are of prices one step ahead. Also true in the Cagan and in OG models with 2-period lives, in which

$$p_t = Q(p_{t+1}^e, w_t).$$

- However, in standard macro models (e.g. RBC or NK models) agents have long (infinite) lives and usually have corresponding planning horizons.

- Two main approaches have been used in agents with long lives: (i) "Euler-equation" or "shadow-price" learning, and (ii) "infinite-horizon" learning.

- Other approaches based on finite planning horizons are also possible.

Shadow-Price Learning

This discussion is based on "Learning to Optimize," Evans and McGough (2012). The idea is to precisely formulate the Euler-equation learning (EE-learning) idea, at the agent level, show how to set this up in a general way, and demonstrate that asymptotically it leads to optimal dynamic decision-making.

- Ours is a "bounded optimality" approach because agents are fully optimal only asymptotically.

- Consider the standard linear-quadratic regulator problem for a single agent: determine a sequence of controls u_t that solve, given the initial state x_0,

$$\begin{array}{ll} \text{max} & -E_0 \sum \beta^t \left(x_t' R x_t + u_t' Q u_t + 2 x_t' W u_t \right) \\ \text{s.t.} & x_{t+1} = A x_t + B u_t + C \varepsilon_{t+1}, \end{array}$$

A simple example is a linear-quadratic Robinson Crusoe economy. Under well-known conditions the sequence of controls are determined by

$$u_t = -Fx_t \text{ where } F = -\left(Q + \beta B'PB\right)^{-1}(\beta B'PA + W')$$

where P is obtained by analyzing Bellman's equation and satisfies

$$P = R + \beta A'PA - (\beta A'PB + W)\left(Q + \beta B'PB\right)^{-1}(\beta B'PA + W').$$

Solving this "Riccati equation" is generally only possible numerically. This requires a sophisticated agent with knowledge and computational skills.

- We replace RE and full optimality with (i) adaptive learning and (ii) bounded optimality, based on (iii) the Lagrangian approach.

- The FOCs from the Lagrangian

$$\begin{aligned}\mathcal{L} &= E_0 \sum \beta^t \{-x_t' R x_t - u_t' Q u_t - 2x_t' W u_t + \\ &\quad \lambda_t'(A x_{t-1} + B u_{t-1} + C\varepsilon_t - x_t)\} \text{ give}\end{aligned}$$

$$\begin{aligned}u_t &= -Q^{-1} W' x_t + (\beta/2) Q^{-1} B' E_t \lambda_{t+1} \\ \lambda_t &= -2Rx_t - 2W u_t + \beta A' E_t \lambda_{t+1},\end{aligned}$$

where λ_t is the vector of shadow-prices of the state variables. These, the transition equation and the TVC identify optimal decision-making.

- Assuming adaptive learning we replace (A, B) with (A_t, B_t), estimated and updated by RLS. Under bounded rationality we replace λ_t and $E_t \lambda_{t+1}$ with $\hat{E}_t \lambda_t^*$ and $\hat{E}_t \lambda_{t+1}^*$.

$$\begin{aligned}u_t &= -Q^{-1} W' x_t + (\beta/2) Q^{-1} B_t' \hat{E}_t \lambda_{t+1}^* \\ \hat{E}_t \lambda_t^* &= -2Rx_t - 2W u_t + \beta A_t' \hat{E}_t \lambda_{t+1}^*.\end{aligned}$$

These two equations are the heart of the **SP-learning** approach: (1) given estimates (A_t, B_t) and $\hat{E}_t \lambda^*_{t+1}$, agents know how to choose their control u_t. (2) given (x_t, u_t), A_t and $\hat{E}_t \lambda^*_{t+1}$, agents know how to compute their estimate $\hat{E}_t \lambda^*_t$ of the value of a unit of x_t today.

- Finally we specify how agents make forecasts $\hat{E}_t \lambda^*_{t+1}$. We again use adaptive learning. Under RE and optimal decision making $\lambda_t = \bar{H} x_t$ for some $\bar{H}$, so we assume agents use the PLM

$$\lambda^*_t = H x_t + \mu_t.$$

Agents do not know H, and at t use RLS to update their estimate to H_t, using a regression of $\hat{E}_s \lambda^*_s$ on x_s with data $s = 1, \ldots, t-1$. Then

$$\hat{E}_t \lambda^*_{t+1} = H_t (A_t x_t + B_t u_t).$$

- This fully describes the SP-learning as a recursive system.

- **Theorem** *Under standard assumptions, and assuming a suitable projection facility, then under SP-learning* (H_t, A_t, B_t) *converges to* $(\bar{H}, A, B)$ *almost surely.*

- This is a striking result: decisions converge asymptotically to the fully rational and fully optimal solution. By estimating shadow prices, we have converted an infinite-horizon problem into a two-period optimization problem.

Example: Robinson Crusoe economy

$$\max -E\sum_{t\geq 0}\beta^t\left((c_t-\hat{b})^2+\phi s_{t-1}^2\right)$$

$$\text{s.t.}\qquad s_t = A_1 s_{t-1} + A_2 s_{t-2} - c_t + \mu_{t+1}$$

Output is fruit/sprouting trees. Under SP-learning Bob estimates the SPs of new and old trees:

$$\lambda_{it}^* = a_{it} + b_{it}s_{t-1} + d_{it}s_{t-2},\ \text{for } i = 1,2,\ \text{and thus}$$

$$\hat{E}_t\lambda_{it+1}^* = a_{it} + b_{it}(A_{1t-1}s_{t-1} + A_{2t-1}s_{t-2} - c_t) + d_{it}s_{t-1},\ \text{for } i = 1,2.$$

These plus the FOC for the control

$$c_t = \hat{b} - 0.5\beta\hat{E}_t\lambda_{1t+1}^*.$$

determine $c_t, E_t\lambda_{1,t+1}^*, E_t\lambda_{2,t+1}^*$, given s_{t-1}, s_{t-2}.

The FOCs for the states give updated estimates of SPs

$$\begin{aligned}\hat{E}_t\lambda^*_{1t} &= -2\phi s_{t-1} + \beta A_{1t} E_t \lambda^*_{1t+1} + \beta E_t \lambda^*_{2t+1} \\ \hat{E}_t\lambda^*_{2t} &= \beta A_{2t} E_t \lambda^*_{1t+1},\end{aligned}$$

which allows us to use RLS update the SP equation coefficients.

For this example EE-learning is also possible (by substituting out the SPs)

$$\begin{aligned}c_t - \beta\phi s_t &= \Psi_t + \beta A_{1t}\hat{E}_t c_{t+1} + \beta^2 A_{2t}\hat{E}_t c_{t+2}, \\ \text{where } \Psi_t &= \hat{b}(1 - \beta A_{1t} - \beta^2 A_{2t})\end{aligned}$$

To implement EE-learning agents forecast using estimates of

$$c_t = a_3 + b_3 s_{t-1} + d_3 s_{t-2}.$$

SP-learning and EE-learning are not identical, but both are asymptotically optimal.

- SP learning can be applied to more general set-ups and in general equilibrium models.

- In special cases SP-learning reduces to Euler-equation learning, but SP-learning is more general.

- Advantage of SP-learning/EE-learning: agents need only solve 2-period optimization problems using one-step ahead forecasts of states and shadow prices.

Infinite-Horizon Learning

- An alternative to bounded optimality and SP-learning.

- While there are antecedents, e.g. Marcet and Sargent (1989) and Sargent (1993), this approach has been stressed by Preston IJCB (2005), JME (2006) and Eusepi and Preston AEJmacro (2010), in NK models, and by Eusepi and Preston AER (2010) for the RBC model.

- IH ('optimal') learning assumes at each t agents make fully optimal decisions given their forecasts for variables outside their control. This requires forecasts infinitely far into the future.

- IH-learning uses the “anticipated utility” approach described by Kreps (1998): agents make fully optimal decisions conditional on their forecasts, but do not take into account that their forecast rules will likely change over time.

- Advantage of IH-learning: agents are fully optimal given their forecasts, explicitly incorporating their TVC and any IBC.

IH-learning vs. EE-learning in the Ramsey Model

To illustrate consider a discrete-time non-stochastic Ramsey model

$$\max \; E_t^*\{\sum_{s=t}^{\infty} \beta^{s-t} \frac{c_s^{1-\sigma}}{1-\sigma}\} \text{ s.t. } a_{s+1} = w_s + r_s a_s - c_s - \tau_s, \text{ for all } s \geq t,$$

where $a_s = k_s + b_s$ and r_s is the real rate of return factor.

EE-learning

For now consider a balanced budget with constant g_t and no debt:

$$g_t = \tau_s = g, \; b_s = 0 \text{ and } \tau_{t+j}^e(t) = g.$$

The Euler equation is $c_t^{-\sigma} = \beta E_t^*(r_{t+1}c_{t+1}^{-\sigma})$. With point expectations

$$c_t = \beta^{-\frac{1}{\sigma}} \left(r_{t+1}^e(t)\right)^{-\frac{1}{\sigma}} c_{t+1}^e(t).$$

Under EE-learning agents form forecasts of $r_{t+1}^e(t)$ and of **their own consumption** next period, $c_{t+1}^e(t)$. The other equations are

$$\begin{aligned} w_t &= f(k_t) - k_t f'(k_t),\ r_t = 1 - \delta + f'(k_t) \text{ and} \\ k_{t+1} &= f(k_t) - c_t - g + (1-\delta)k_t. \end{aligned}$$

In a nonstochastic model, simple learning rules can be used

$$\begin{aligned} r_{t+1}^e(t) &= r_t^e(t-1) + \gamma(r_t - r_t^e(t-1)) \\ c_{t+1}^e(t) &= c_t^e(t-1) + \gamma(c_{t-1} - c_t^e(t-1)). \end{aligned}$$

This system converges to the steady state equilibrium. In stochastic versions we have convergence to RE under LS learning rules.

IH-learning

Under IH learning, agents fully optimize using their IBC and TVC. The Euler equation with point expectations implies $c^e_{t+j}(t) = c_t \beta^{\frac{j}{\sigma}} (\prod_{i=1}^{j} r^e_{t+i}(t))^{\frac{1}{\sigma}} \equiv c_t \beta^{\frac{j}{\sigma}} (D^e_{t,t+j}(t))^{\frac{1}{\sigma}}$. Substituting into the IBC of the household gives

$$
\begin{aligned}
c_t(1 + S^e_D(t)) &= r_t a_t + w_t - \tau_t + \sum_{j=1}^{\infty} (D^e_{t,t+j}(t))^{-1} (w^e_{t+j}(t) - \tau^e_{t+j}(t)), \\
\text{where } S^e_D(t) &\equiv \sum_{j=1}^{\infty} \beta^{j/\sigma} (D^e_{t,t+j}(t))^{\sigma^{-1}-1}.
\end{aligned}
$$

Given forecasts $w^e_{t+j}, r^e_{t+1}(t)$ and τ^e_{t+j}, this determines c_t. Other variables are given as before.

For the simplest version of adaptive learning in this nonstochastic setting ("steady state expectations")

$$\begin{aligned} r^e_{t+i}(t) &= r^e(t) \text{ where } r^e(t) = r^e(t-1) + \gamma(r_t - r^e(t-1)) \text{ and} \\ w^e_{t+i}(t) &= w^e(t) \text{ where } w^e(t) = w^e(t-1) + \gamma(w_t - w^e(t-1)). \end{aligned}$$

It can be shown numerically that the steady-state is stable under IH-learning.

For stochastic Ramsey models EE-learning and IH-learning can be implemented with suitable PLMs, and some analytical stability results are available.

An advantage of IH-learning is that it can be used to incorporate **structural information** about future changes known to agents. I consider examples that focus on known future policy changes.

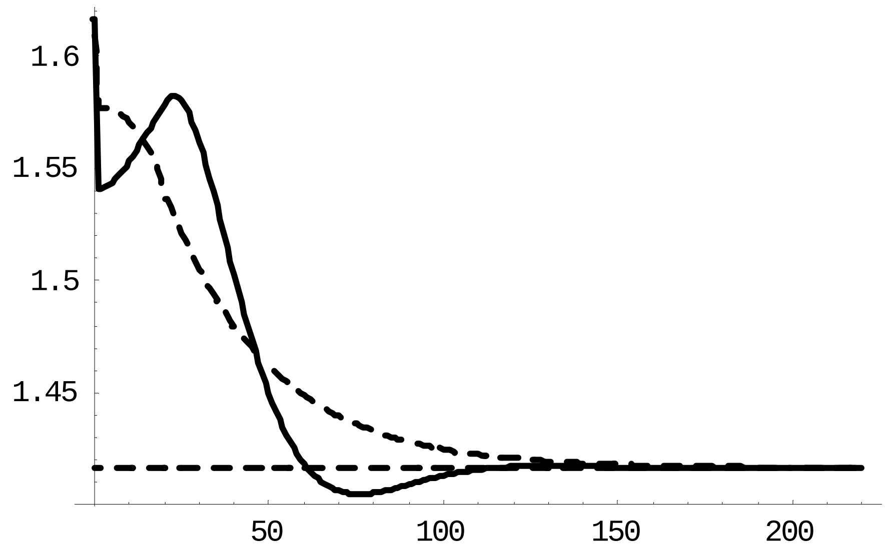
Anticipated Fiscal policy

IH-learning can be used to capture structural information, e.g. Eusepi & Preston AEJmacro (2010) emphasize that knowledge of the monetary policy rule helps stabilize the economy.

Evans, Honkapohja & Mitra JME (2009) incorporate anticipated future changes in fiscal policy into adaptive learning using the IH-learning approach.

A hallmark of RE is that announced future policies have an impact now. This also happens with IH learning. EHM (2009) show the impact in the Ramsey model of an announced future permanent increase in government spending.

$$g_t = g_0 = \tau_0 \text{ when } t < T_p \text{ and } g_t = g_1 = \tau_1 \text{ for } t \geq T_p.$$



c_t dynamics under learning (solid curve) and perfect foresight (dashed curve). Straight dashed line is new steady state for c. $T_p = 20$.

Immediate impact due to the understanding by agents that future taxes will be higher. Learning dynamics differ from RE because agents do not know GE effects and use adaptive learning to forecast w_{t+i} and r_{t+i}.

If same policy change were repeated many times, agents could eventually learn RE, but policy changes typically have unique features.

Ricardian Equivalence?

Evans, Honkapohja and Mitra (JMCB, forthcoming) "Does Ricardian Equivalence Hold when Agents are not Rational?"

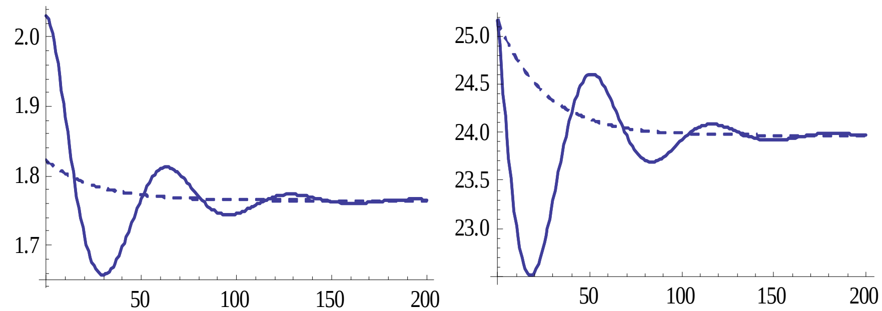
Suppose we drop the balanced budget assumption. Now agents must forecast taxes. Will Ricardian Equivalence hold under IH-learning? This may indeed hold if agents incorporate into their forecasts a key piece of structural information: the IBC of the government.

We examine the question for very general expectation formation mechanisms (of which LS learning is a special case).

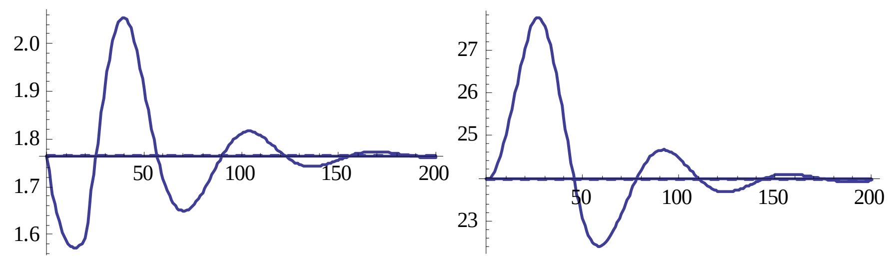
The key result is:

Proposition: Assume neither government spending nor expectations depend on current government financing variables (taxes and end-of-period debt). Then the Ramsey model exhibits Ricardian Equivalence: the sequence of consumption, capital, rates of return and wages along the path of equilibria with learning is independent of the government financing policy.

EHM gives examples of a temporary tax cut, followed later by an increase in taxes to cover the extra interest on debt. They find:
(i) that Ricardian Equivalence can hold even when under adaptive learning the learning paths are very different from RE paths, and
(ii) new reasons for possible failures of Ricardian Equivalence, despite incorporation of the govt. IBC into agent's lifetime budget constraint.



Ricardian equivalence. c_t (left) and k_t (right) under learning (solid line) and RE (dashed line). Initial $k_0 > \bar{k}$.



Failure of Ricardian Equivalence: c_t (left) and k_t (right) under learning for deficit financing (solid curve) and balanced budget (horizontal line) if learning rule depends on debt. $k_0 = \bar{k}$.

RBC models: business cycle fluctuations

Eusepi-Preston, AER (2010) look at equilibrium fluctuations in a an RBC model with IH-learning. The basics of the model are fairly standard. Markets are competitive and output of representative firm i is given by

$$\begin{aligned} Y_t^i &= (K_t^i)^\alpha (X_t H_t^i)^{1-\alpha}, \text{ where } 0 < \alpha < 1 \\ \ln(X_{t+1}/X_t) &= \gamma_t = \ln \bar{\gamma} + a_{t+1}, \text{ where } a_t \text{ is white noise.} \end{aligned}$$

Representative household j maximizes

$$\hat{E}_t^j \sum_{T=t}^{\infty} \beta^{T-1} \left[\ln C_T^j - \nu(H_T^j) \right].$$

Transform variables to stationary variables

$$y_t = Y_t/X_t, c_t = C_t/X_t, i_t = I_t/X_t, w_t = W_t/X_t \text{ and } k_t = K_t/X_{t-1}.$$

Log-linearizing, and substituting the Euler equation and the static FOC into the IBC yields the consumption rule.

Aggregating, the consumption function takes the log deviation form

$$\begin{aligned}\hat{c}_t &= \frac{1-\beta}{\epsilon_c}\left[\beta^{-1}\hat{k}_t + \bar{R}\hat{R}_t^K - \beta^{-1}\hat{\gamma}_t + \epsilon_w \hat{w}_t\right] \\ &\quad + \hat{E}_t \sum_{T=t}^{\infty} \beta^{T-1}\left[\frac{1-\beta}{\epsilon_c} - \beta\right]\beta\bar{R}\hat{R}_{T+1}^K + \hat{E}_t \sum_{T=t}^{\infty} \beta^{T-1}\frac{(1-\beta)}{\epsilon_c}\beta\epsilon_w\hat{w}_{T+1}.\end{aligned}$$

Under learning agents estimate and make forecasts using

$$\begin{aligned}\hat{R}_t^K &= \omega_0^r + \omega_1^r \hat{k}_t + e_t^r \\ \hat{w}_t &= \omega_0^w + \omega_1^w \hat{k}_t + e_t^w \\ \hat{k}_{t+1} &= \omega_0^k + \omega_1^k \hat{k}_t + e_t^k,\end{aligned}$$

The estimated parameters are updated each period using constant-gain learning ("perpetual" learning).

The basic findings for the calibrated model (compared to RE) are:
(i) The learning model delivers same output volatility with smaller technology shocks.
(ii) The learning model has more volatility in investment and hours.
(iii) The model captures persistence in investment and hours.

Adding IH-learning to the RBC model improves the fit to the data. The key mechanism: partially self-fulfilling shifts in expectations arising from technology shocks that generate temporary but persistent movements in estimated coefficients.

RBC models: changes in fiscal policy

Mitra, Evans and Honkapohja in "Policy change and learning in the RBC model" (2011) and "Fiscal policy and learning" (2012) consider announced changes in fiscal policy in a linearized, stochastic RBC model with IH learning.

Consider the impact of **announced temporary increase in** g financed by lump-sum taxes. What is the **output multiplier**?

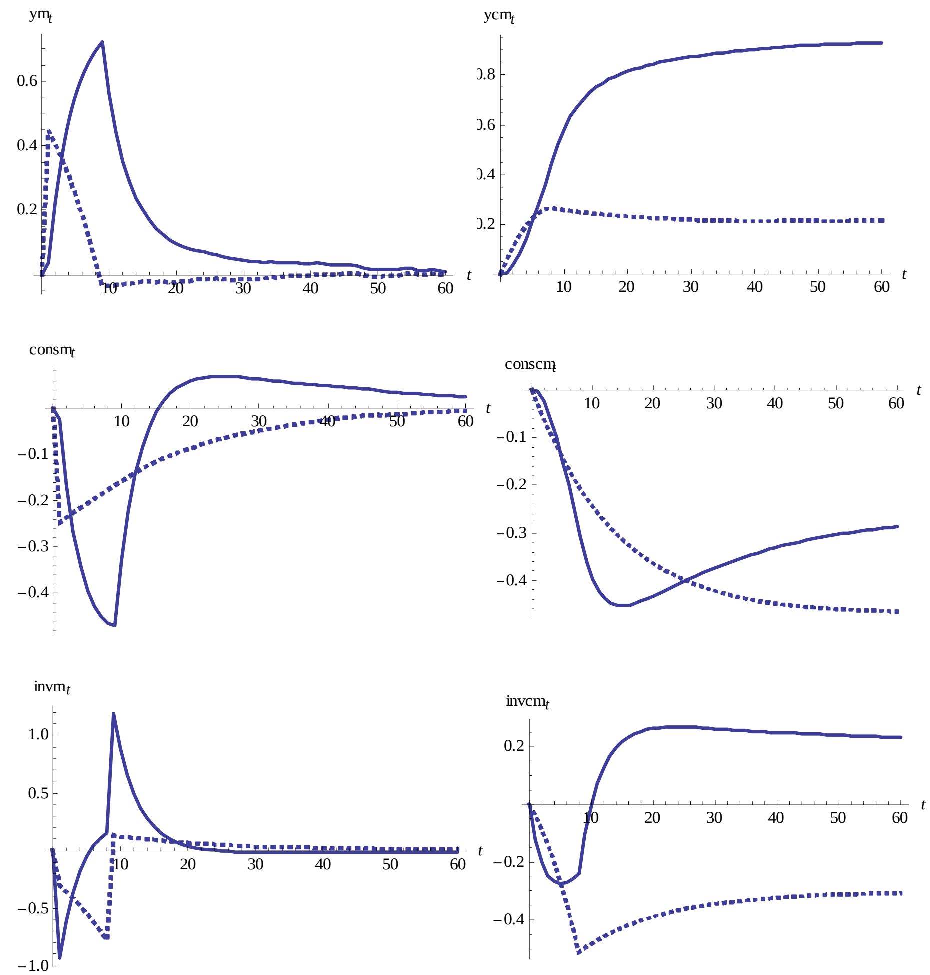
– We use an RBC model, not because of strong belief in it, but because the neoclassical mechanisms are one part of most DSGE models.
– Multipliers in RBC models are known under RE to be too small compared to empirical values of, e.g., Hall (2009) $0.7-1$ and Ramey (2011) $0.8-1.5$.

Our procedure under IH-learning is:

- Agents forecast wages and interest rates using a statistical model with parameters updated over time by constant-gain LS.
- To forecast future taxes agents use (credible) announcements of policy changes.

Main result: output multipliers are much higher than under RE. We illustrate using an 8 quarter 5% increase in g, credibly announced to be temporary.

Discounted cumulative output multiplier at five years over 0.8 under RE and $<$ 0.25 under RE. Big difference in investment multiplier.
In work in progress we are looking at NK models.



Multipliers (left) and cumulative multipliers (right), for y, c and i for temporary increase in g. Solid lines: learning and dashed lines: RE.

NK models: liquidity traps and fiscal policy

Benhabib, Evans and Honkapohja (2012), "Liquidity Traps and Expectation Dynamics: Fiscal Stimulus or Fiscal Austerity?" looks at IH learning in a Rotemberg-type NK model with multiple steady states due to the ZLB.

- Global analysis under adaptive learning in the nonlinear model. We show existence of deflation traps under IH-learning. Corridor of stability that includes targeted steady state.

- We consider versions both with and without Ricardian consumers.

- A well-designed temporary government spending stimulus will push the economy out of the deflation trap and to the intended steady state.

Back to Methodology

Eductive stability, RE and adaptive learning

- Eductive stability models agents as theorists. Suppose agents are hyper-rational, know the entire economic structure, and these are CK (common knowledge). If the REE can be deduced we say it is eductively stable.

- Not all REE are eductively stable: CK of structure and rationality may not be sufficient for coordination on RE – cobweb model with $\alpha < -1$.

- Evans, Guesnerie and McGough (2011) show the RBC model not strongly eductively stable.

- Failure of eductive stability opens the door for adaptive learning and hence learning dynamics (e.g. see Brock & Hommes (1997), Hommes (2011) for the cobweb model).

- Lack of full knowledge of structure, lack of full rationality, or lack of CK of rationality also open the door for adaptive learning.

SP-learning vs. IH-learning

Both SP-learning and IH-learning implement adaptive learning in models with long-lived agents. Both can converge to the REE.

- SP-learning only requires agents to make one-period forecasts and solve relatively simple two-period optimization problems, and it is computationally simple.

- Against this, SP-learning assumes a stationary environment. Extensions to incorporate knowledge about future structural change appear to require longer planning horizons.

- In IH-learning agents make optimal decisions, conditional on forecasts of future state variables, explicitly imposing any IBC and TVC. Known future policy changes can be incorporated.

- Against this, IH-learning neglects the additional parameter uncertainty in long-horizon forecasts, assumes agents can solve difficult optimization problems, and is computationally more demanding.

LS adaptive learning vs. Bayesian learning

- In the Muth-Lucas and Cagan models there is little difference between LS learning and Bayesian learning, which simply incorporates a prior on parameters.

- Even in Muth-Lucas, Bayesian learning is not "fully" rational in the sense that the subjective probability distribution is not correct during the learning transition..

- Bayesian learning can be incorporated in models with long-lived agents with short planning horizons, e.g. Adam & Marcet (JET, 2011).

- In models with longer planning horizons, Bayesian learning allows agents to take account of future parameter change, Cogley and Sargent (IER, 2008). Cogley-Sargent assumed finite lives and two-state exogenous Markov processes to make solving the problem feasible.

- For this example they find the anticipated utility approach of optimization using forecasts based on LS-learning to be approximately optimal.

The planning horizon and finite-horizon learning

- The tension between short and longer horizons is not new.

- For equity pricing Timmermann (REStud, 1996) examined both long-horizon "present-value" and short-horizon "self-referential" learning. See Chakraborty and Evans (JME, 2008) and Kim (JEDC, 2009) for an exchange rate example.

- Introspection and common sense suggest **finite-horizon learning**.

- Branch, Evans and McGough, "Finite-horizon learning," forthcoming in Sargent and J. Vilmunen, eds. (2012), show how to generalize both EE-learning and IH-learning to get finite-horizon learning, based on N-step versions of each.

- The planning horizon affects speed of convergence and other features. The planning horizon may be a key parameter that should be estimated.

Conclusions

- The adaptive learning approach models agents as econometricians in making forecasts.

- I've focused on two methodological issues: the extent of structural knowledge and the planning horizon.

- Adaptive learning agrees with the cognitive consistency principle if structural knowledge is imperfect, CK of rationality is unlikely or eductive stability fails.

- In decision-making, long-lived agents may plausibly use short or long horizons, depending on the setting.

- Short-horizon SP-learning or EE-learning will often work well: in stationary environments, they provide simple decision rules based on one-step ahead forecasts that converge asymptotically to fully optimal decisions.

- IH-learning can take account of policy commitments or announced paths of policy variables, e.g. temporary or permanent changes in g. Announced policy changes have impact effects as in RE.

- The planning horizon of agents will be a key parameter in applied models.