Learning to Optimize: Theory and Applications

George W. Evans

University of Oregon and University of St Andrews

Bruce McGough

University of Oregon

WAMS, December 12, 2015

Outline

– Introduction

– Shadow-price learning for dynamic programming problems

– Learning to optimize in an LQ problem

– Value function learning & Euler-equation learning

– Applications: Investment model; Crusoe; Ramsey model; NK model

– Conclusions

Introduction

- In microfounded models we assume agents are rational in two ways:
  – they form forecasts optimally (they are endowed with RE)
  – they make choices by maximizing their objective function

- RE may be implausibly demanding. The adaptive (e.g. least-squares) learning approach is a natural bounded-rationality response to this critique. See, e.g., Marcet & Sargent (1989), Evans & Honkapohja (2001).

- Under least-squares learning, agents can learn over time to coordinate on an REE in "self-referential models" if it is "E-stable." Interesting learning dynamics can emerge.

- That agents are endowed with the solution to dynamic optimization problems is equally implausible: it may take time to learn to optimize.

- Boundedly optimal decision-making is a natural complement to boundedly rational forecasting. It obeys the "cognitive consistency principle."

- Our implementation, which we call shadow-price learning, complements and extends least-squares learning in expectation formation.

- Using shadow-price learning agents can learn over time to solve their dynamic stochastic optimization problem.

- Again, interesting learning dynamics can emerge.

Literature on agent-level learning and decision-making

– Cogley and Sargent (IER, 2008). Bayesian decision-making with learning.

– Adam and Marcet (JET, 2011). "Internal rationality."

– Preston (IJCB, 2005). Eusepi & Preston (AEJmacro 2010), 'Anticipated utility' and infinite-horizon decisions. Evans, Honkapohja & Mitra (JME, 2009).

– Evans and Honkapohja (ScandJE, 2006) and Honkapohja, Mitra and Evans (2013). "Euler-equation learning." Howitt and Özak (2014).

– Watkins (1989). Q-learning. 'Quality value' of state-action pairs. Typical applications are to models with finite states and actions.

– Marimon, McGrattan and Sargent (JEDC, 1990). Classifier systems. Lettau and Uhlig (AER, 1999).

Shadow-price learning

We now introduce our approach – Shadow-price (SP) learning. Consider a standard dynamic programming problem

$$
\begin{aligned}
V^*(x_0) &= \max E_0 \sum_{t \geq 0} \beta^t r(x_t, u_t) \\
\text{subject to } x_{t+1} &= g(x_t, u_t, \varepsilon_{t+1})
\end{aligned}
$$

and $\bar{x}_0$ given, with $u_t \in \Gamma(x_t) \subseteq \mathbb{R}^m$ and $x_t \in \mathbb{R}^n$.

Linear-Quadratic (LQ) special case:

$$
\begin{aligned}
r(x_t, u_t) &= x_t' R x_t + u_t' Q u_t + 2 x_t' W u_t \\
g(x_t, u_t, \varepsilon_{t+1}) &= A x_t + B u_t + C \varepsilon_{t+1}.
\end{aligned}
$$

Examples:

1. Lucas-Prescott-Sargent model of investment. Market demand is

$$p_t = \alpha_0 - \alpha_1 y_t + v_t,$$

where v_t is AR(1) stationary and observable. Firm ω's problem is

$$\begin{aligned}
\max_{I_t(\omega)} \quad & \hat{E}(\omega) \sum_{t \geq 0} \beta^t \left(p_t y_t(\omega) - (J + q_t(\omega)) I_t(\omega) - \frac{\gamma}{2} I_t(\omega)^2 \right) \\
k_t(\omega) &= (1-\delta) k_{t-1}(\omega) + \mu I_t(\omega) + (1-\mu) I_{t-1}(\omega), \text{ with } 0 < \delta \leq 1 \\
y_t(\omega) &= k_t(\omega)^\alpha \text{ where } 0 < \alpha \leq 1, \\
p_t &= a_0 + a_1 p_{t-1} + a_2 v_{t-1} + \varepsilon_t^p,
\end{aligned}$$

where $q_t(\omega) = q(\omega) + \varepsilon_t^q(\omega)$. Firm's problem is LQ for $\alpha = 1$.

Two cases: (i) Single agent problem for given price process.

(ii) Market equilibrium with p_t given by demand and $y_t = \int_\Omega y_t(\omega) d\omega$.

2. Robinson Crusoe & Ramsey models of capital accumulation.

(i) Robinson Crusoe LQ problem:

$$\max -E\sum_{t\geq 0}\beta^t\left((c_t-b^*)^2+\phi s_{t-1}^2\right)$$
$$\text{s.t.}\qquad s_{t+1}=A_1 s_t+A_2 s_{t-1}-c_t+\mu_{t+1}.$$

(ii) General equilibrium Ramsey problem:

$$\max\qquad E\sum\beta^t U(c_t(\omega))$$
$$s_{t+1}(\omega)=(1+r_t)s_t(\omega)+w_t-c_t(\omega).$$

Competitive firms use technology $y=zk^\alpha$ where $\log z_t$ is AR(1) stationary. w_t, r_t given by marginal products. Representative agent with $k_t=s_t=s_t(\omega)$.

3. New Keynesian model with heterogeneous agents.

Shadow-price learning

Return to consideration of the standard dynamic programming problem

$$\begin{aligned} V^*(x_0) &= \max E_0 \sum_{t \geq 0} \beta^t r(x_t, u_t) \\ \text{subject to } x_{t+1} &= g(x_t, u_t, \varepsilon_{t+1}) \end{aligned}$$

and $\bar{x}_0$ given, with $u_t \in \Gamma(x_t) \subseteq \mathbb{R}^m$ and $x_t \in \mathbb{R}^n$. The state x_t includes an intercept. Our approach is based on the corresponding Lagrangian

$$\mathcal{L} = E_0 \sum_{t \geq 0} \beta^t \Big(r(x_t, u_t) + \lambda_t^{*\prime} \left(g(x_{t-1}, u_{t-1}, \varepsilon_t) - x_t \right) \Big).$$

Our starting point is the FOC and envelope condition

$$\begin{aligned} \lambda_t^* &= r_x(x_t, u_t)' + \beta E_t g_x(x_t, u_t, \varepsilon_{t+1})' \lambda_{t+1}^* \\ 0 &= r_u(x_t, u_t)' + \beta E_t g_u(x_t, u_t, \varepsilon_{t+1})' \lambda_{t+1}^*. \end{aligned}$$

In **SP learning** we replace λ_t^* with λ_t, the perceived shadow price of the state x_t, and we treat these equations as **behavioral**.

To implement this we need forecasts. In line with the adaptive learning literature $x_{t+1} = g(x_t, u_t, \varepsilon_{t+1})$ is often assumed **unknown** and is **approximated** by

$$x_{t+1} = Ax_t + Bu_t + C\varepsilon_{t+1},$$

where unknown parameter estimates are updated over time using RLS, i.e. recursive LS. Agents must also forecast λ_{t+1}. We assume that they believe the dependence of λ_t on x_t can be approximated by

$$\lambda_t = Hx_t + \mu_t,$$

where estimates of H are updated over time using RLS.

To implement SP learning, given x_t and estimates A, B, H the agent sets u_t to satisfy

$$r_u(x_t, u_t)' = -\beta B' \hat{E}_t \lambda_{t+1}, \text{ since } B = \partial x_{t+1}/\partial u_t, \text{ and}$$

$$\hat{E}_t \lambda_{t+1} = H\left(Ax_t + Bu_t\right)$$

for u_t and $\hat{E}_t \lambda_{t+1}$. The FOC r_u equation may in general be nonlinear.

The x_t FOC (envelope condition) gives a value for λ_t

$$\lambda_t = r_x(x_t, u_t)' + \beta A' \hat{E}_t \lambda_{t+1}, \text{ since } A = \partial x_{t+1}/\partial x_t,$$

which is used next period to update the estimate of H. This equation has an asset price interpretation.

At $t+1$ RLS is used to update estimates of A and/or B and H in

$$x_{t+1} = Ax_t + Bu_t + C\varepsilon_{t+1} \text{ and } \lambda_t = Hx_t + \mu_t,$$

This fully defines SP learning as a recursive system.

Advantages of SP learning as a model of boundedly optimal decision-making:

- The **pivotal role of shadow prices** λ_t, central to economic decisions.

- $\hat{E}_t \lambda_{t+1}$ and transition dynamics $B = \partial x_{t+1} / \partial u_t$ measure the **intertemporal trade-off** which determines actions u_t.

- **Simplicity**. Agents each period solve a **two-period problem** – an attractive level of sophistication.

- Incorporates **recursive LS updating** of A, B, H, the hallmark of **adaptive learning**, but extended to include forecasts of shadow prices.

- As we will see, although our agents are **boundedly optimal**, in a LQ setting they **become fully optimal asymptotically**.

- SP learning **can be incorporated into** standard **DSGE models**.

We also outline two alternative implementations of SP learning:

– Value function learning: value function estimated instead of shadow prices.

– Euler equation learning: closely related to SP-learning in special cases.

SP learning is related to the other approaches in the literature:

– Like Q-learning and classifier systems, it builds off of Bellman's equation.

– Like Internal Rationality we do not impose RE.

– As with anticipated utility/IH agents neglect parameter uncertainty.

– Like Euler-equation learning, it is sufficient to forecast one step ahead.

– Like anticipated utility/IH & Internal Rationality, an agent-level approach

SP-learning has simplicity, generality and economic intuition, and can be embedded in general equilibrium models with heterogeneous agents..

Learning to Optimize in an LQ set-up

- We now specialize the dynamic programming set-up to be the standard linear-quadratic set-up, which has been extensively studied and widely applied. In this set-up we can obtain our asymptotic convergence result.

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

$$\begin{aligned} V^*(x_0) = \max \quad & -E_0 \sum \beta^t \left(x_t' R x_t + u_t' Q u_t + 2 x_t' W u_t \right) \\ s.t. \quad & x_{t+1} = A x_t + B u_t + C \varepsilon_{t+1}. \end{aligned}$$

  We make standard assumptions on R, Q, W, A, B: LQ.1 (concavity), LQ.2 (stabilizability) and LQ.3 (detectability).

- Under LQ1 – LQ3 the optimal controls are given by

$$u_t = -F^* x_t \text{ where } F^* = \left(Q + \beta B' P^* B\right)^{-1} (\beta B' P^* A + W')$$

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

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

Also $V^*(x) = -x' P^* x - \beta (1-\beta)^{-1} \text{tr}(\sigma_\varepsilon^2 P^* C C')$.

- Solving the "Riccati equation" for P^* generally only possible numerically. This requires a sophisticated agent with a lot of knowledge and computational skills. Our agents follow a simpler boundedly optimal procedure.

- Our approach replaces RE and full optimality with adaptive learning and bounded optimality, based on shadow prices.

- For LQ models the true transition equation is linear and the optimal shadow price equation is linear

- The SP-learning system can be written recursively as:

$$
\begin{aligned}
x_t &= Ax_{t-1} + Bu_{t-1} + C\varepsilon_t \\
\mathcal{R}_t &= \mathcal{R}_{t-1} + \gamma_t \left(x_t x_t' - \mathcal{R}_{t-1} \right) \\
H_t' &= H_{t-1}' + \gamma_t \mathcal{R}_{t-1}^{-1} x_{t-1} \left(\lambda_{t-1} - H_{t-1} x_{t-1} \right)' \\
A_t' &= A_{t-1}' + \gamma_t \mathcal{R}_{t-1}^{-1} x_{t-1} \left(x_t - Bu_{t-1} - A_{t-1} x_{t-1} \right)' \\
u_t &= F^{SP}(H_t, A_t, B) x_t \\
\lambda_t &= T^{SP}(H_t, A_t, B) x_t \\
\gamma_t &= t^{-1} \text{ or } \gamma_t = \kappa (t+N)^{-1}
\end{aligned}
$$

In this formulation A is estimated but B is assumed known, which would be typical.

- For real-time learning results we need an additional assumption, LQ.RTL: the state dynamics are well-behaved under optimal decision-making, i.e. are stationary and have a non-singular second-moment matrix.

Theorem 4 (Asymptotic optimality of SP learning in LQ model). *If LQ.1 - LQ.3 and LQ.RTL are satisfied then, locally,* (H_t, A_t) *converges to* (H^*, A) *almost surely when the recursive algorithm is augmented with a suitable projection facility, and* $F^{SP}(H_t, A_t, B)$ *converges to* $-F^*$*.*

The proof of Theorem 4 combines known properties of LQ problems with the stochastic approximation tools from the adaptive learning literature.

Extension: We show it is unnecessary for agents to estimate and forecast shadow prices for exogenous states. This is convenient for applications.

Theorem 4 is a striking result:

- SP learning converges asymptotically to fully rational forecasts and fully optimal decisions.

- By including perceived shadow prices, we have converted an infinite-horizon problem into a two-period optimization problem.

- The agent is learning over its lifetime based on a single ‘realization’ of its decisions and the resulting states.

Remark 1: The “projection facility” in many applications is rarely needed.

Remark 2: like adaptive learning of expectations, the system is self-referential. Here this comes from the impact of perceived shadow prices on actual decisions.

Alternative implementation: value-function learning

- In SP learning agents estimate the SP vector λ for state x. An alternative implementation is to estimate $V(x)$

$$V(x) = -x'Px$$

  and make decisions on this basis.

- They use P and the rhs of Bellman's equation to obtain revised $\hat{V}_t$.

- Estimates of P are updated over time using a regression of $\hat{V}_t$ on linear and quadratic terms in the state x_t.

- Theorem 5 provides a corresponding result for value-function learning.

Alternative implementation: Euler-equation learning

- Another alternative implementation of bounded optimality is EE learning.

- One-step-ahead Euler equations exist in special cases (Appendix B), e.g. if x_{t+1} does not depend on endogenous states x_t. Then λ_t can be eliminated from the FOC to give the Euler equation

$$Qu_t + W'x_t + \beta B'E_t(Rx_{t+1} + Wu_{t+1}) = 0.$$

- Under EE-learning the agent forecasts its own future decision u_{t+1} using

$$u_t = -Fx_t.$$

Theorem 6 provides a corresponding result for EE learning.

Example: SP Learning in the Investment Model

We give results here for the single agent problem in which the firms treats p_t as an exogenous process. We rewrite the problem in terms of installed capital

$$z_t(\omega) = (1-\delta)k_{t-1}(\omega) + (1-\mu)I_{t-1}(\omega),$$

where $k_t(\omega) = z_t(\omega) + \mu I_t(\omega)$. The firm problem is then

$$\max_{I_t(\omega)} \hat{E}(\omega) \sum_{t \geq 0} \beta^t \left(p_t f(z_t(\omega) + \mu I_t(\omega)) - (J + q_t(\omega)) I_t(\omega) - \frac{\gamma}{2} I_t(\omega)^2 \right)$$

$$\begin{aligned} z_{t+1}(\omega) &= (1-\delta) z_t(\omega) + (1-\delta\mu) I_t(\omega) \\ v_{t+1} &= \rho v_t + \varepsilon^p_{t+1}, \; q_{t+1} = \bar{q} + \varepsilon^q_{t+1} \\ p_{t+1} &= a_0 + a_1 p_t + a_2 v_t + \varepsilon^p_{t+1} \end{aligned}$$

Exogenous and endogenous states are: $x_{1t} = (1, v_t, q_t(\omega), p_t)$, $x_{2t} = z_t(\omega)$, and the control is $u_t = I_t(\omega)$.

The FOC for I_t gives the key behavioral equation

$$J + q_t(\omega) + \gamma I_t(\omega) = \mu p_t f'(z_t(\omega) + \mu I_t(\omega)) + \beta(1 - \delta\mu)\hat{E}_t \lambda^z_{t+1}(\omega).$$

Decision-making requires also $\hat{E}_t \lambda^z_{t+1}(\omega)$, using forecasts of the state at $t+1$, including $\hat{E}_t p_{t+1}$ based on estimates of a_0, a_1, a_2, and estimates of

$$\lambda^z_t(\omega) = H_0 + H_1 v_t + H_2 q_t + H_3 p_t + H_4 z_t(\omega).$$

The FOC for z_t

$$\lambda^z_t(\omega) = p_t f'(z_t(\omega) + \mu I_t(\omega)) + \beta(1 - \delta)\hat{E}_t \lambda^z_{t+1}(\omega)$$

gives a new data point for $\lambda^z_t(\omega)$ for updating estimates H.

We give the numerical results for the LQ case $f(k) = k^\alpha$ with $\alpha = 1$, and for a non-LQ case with $\alpha = 0.3$.

For the price process we use the REE price process from market demand $p_t = \alpha_0 - \alpha_1 y_t + v_t$ with $\alpha_0 = 10$, $\alpha_1 = -1.1$. The learning gain parameter is 0.2 and the other parameters were set at

$$\beta = .95,\ \mu = .95,\ \delta = .1,\ J = 2,\ \bar{q} = 0,\ \gamma = 2,\ \rho = .7.$$

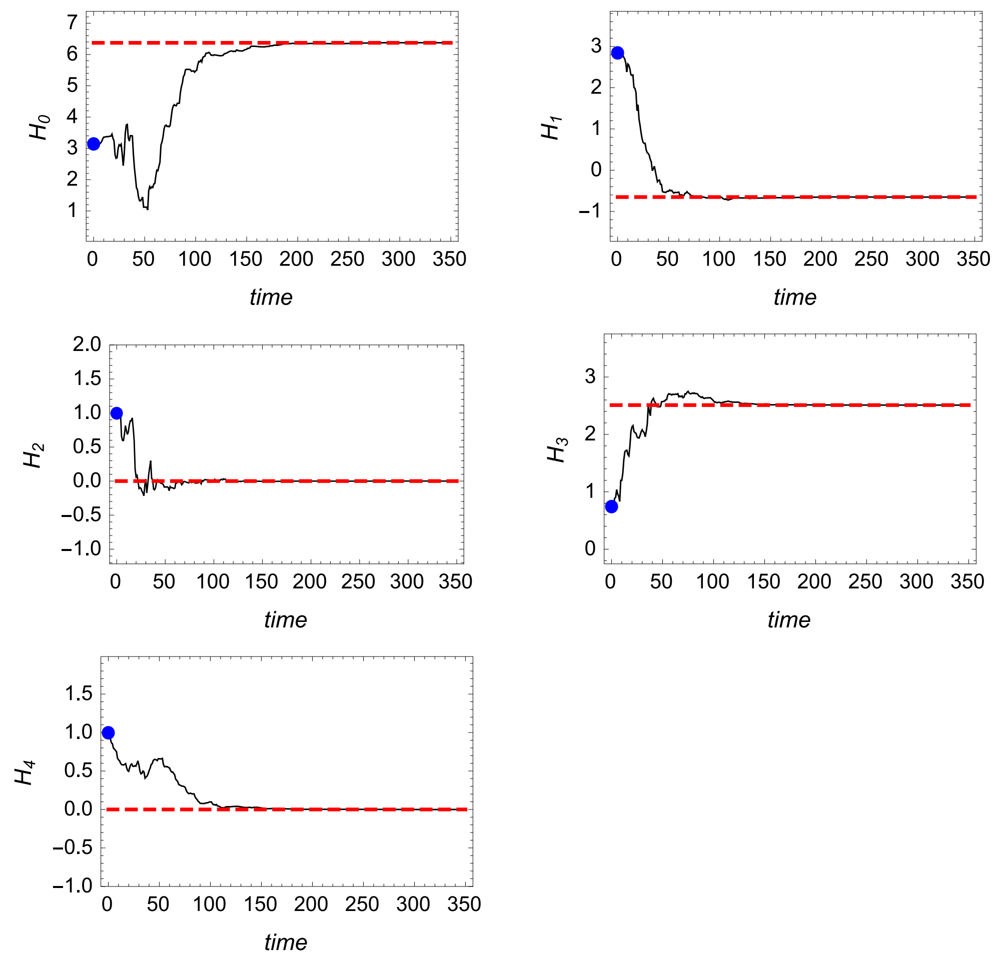
The price processes are given by

$$\begin{aligned} \text{LQ case:} \qquad & p_t = 0.43 + 0.70 p_{t-1} - 0.12 v_{t-1} + \varepsilon_t^p \\ \text{Non-LQ case:} \qquad & p_t = 1.31 + 0.85 p_{t-1} - 0.15 v_{t-1} + \hat{\varepsilon}_t^p \end{aligned}$$

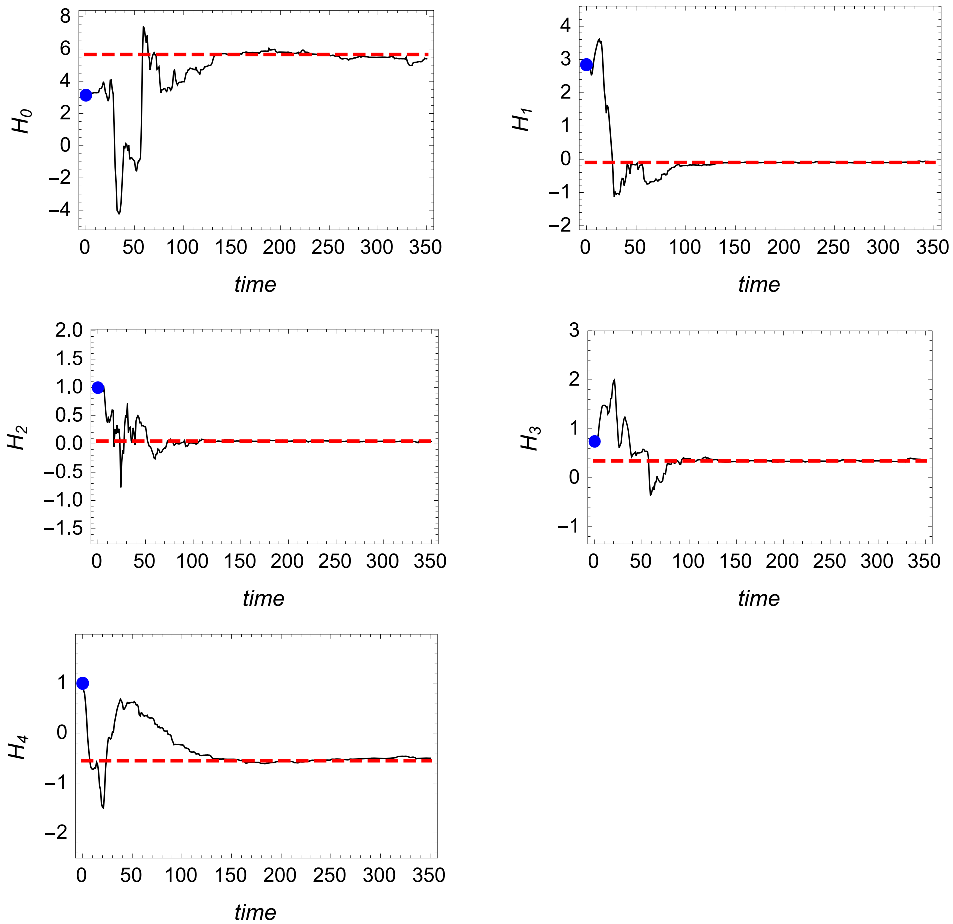
Optimal beliefs H^* are given by

$$\begin{aligned} \text{LQ case} \quad &: \quad H^* = (6.37, -0.65, 0, 2.51, 0) \\ \text{Non-LQ case} \quad &: \quad H^* = (5.66, -0.097, 0.05, 0.34, -0.55) \end{aligned}$$

The following figures show real-time plots of estimated SP parameters H. These indicate convergence to optimal decision-making.



LQ case $\alpha = 1$



Non-LQ case $\alpha = 0.3$

Example: SP Learning in a Crusoe economy

$$\max -E \sum_{t \geq 0} \beta^t \left((c_t - b^*)^2 + \phi s_{t-1}^2 \right)$$
$$\text{s.t.} \quad s_{t+1} = A_1 s_t + A_2 s_{t-1} - c_t + \mu_{t+1}$$

Output is fruit/sprouting trees. Young trees need weeding. Under SP learning Bob estimates the SPs of young and old trees:

$$\lambda_{it} = a_i + b_i s_t + d_i s_{t-1}, \text{ for } i = 1, 2, \text{ and thus}$$

$$\hat{E}_t \lambda_{it+1} = a_i + b_i (A_{1t} s_t + A_{2t} s_{t-1} - c_t) + d_i s_t, \text{ for } i = 1, 2.$$

These plus the FOC for the control

$$c_t = b^* - \frac{\beta}{2} \hat{E}_t \lambda_{1t+1}.$$

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

The FOCs for the states give updated estimates of SPs

$$\begin{aligned}\lambda_{1t} &= -2\phi s_t + \beta A_{1t}\hat{E}_t\lambda_{1t+1} + \beta\hat{E}_t\lambda_{2t+1} \\ \lambda_{2t} &= \beta A_{2t}\hat{E}_t\lambda_{1t+1},\end{aligned}$$

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

Proposition: *Provided LQ.RTL holds, Robinson Crusoe learns to optimally consume fruit.*

Note: LQ.RTL necessarily holds if $A_2 \geq 0$ is not too large and shocks have small support.

EE learning is also possible using a second-order Euler equation. See paper.

Example: SP Learning in a Ramsey Model

The stochastic Ramsey model illustrates a general equilibrium model in a non-LQ setting.

Representative household ω has one unit of labor and maximizes

$$\begin{aligned} \max \quad & E \sum \beta^t U(c_t(\omega)) \\ & s_{t+1}(\omega) = (1 + r_t) s_t(\omega) + w_t - c_t(\omega). \end{aligned}$$

Competitive firms use CRTS technology and $y = zf(k)$ where y, k are output, capital per unit of labor and $\log z_t$ is AR(1) stationary with $Ez_t = 1$. In equilibrium factors are paid their marginal products and $k_t = s_t = s_t(\omega)$.

Households will choose

$$U'(c_t(\omega)) = \beta \hat{E}_t \lambda_{t+1}(\omega),$$

where $\lambda_t(\omega) = (1+r_t)U'(c_t(\omega))$ is the shadow value of an extra unit of $s_t(\omega)$. The full state is $x_t' = (1, s_t(\omega), w_t, r_t, z_t)$, but only $s_t(\omega)$ is endogenous to the agent.

To avoid multicollinearity issues we assume agents estimate

$$\lambda_t(\omega) = H_\gamma(\omega) + H_s(\omega)s_t(\omega) + H_z(\omega)z_t$$

and use this and their flow budget constraint to forecast $\hat{E}_t \lambda_{t+1}(\omega)$, and thus to solve for

$$c_t(\omega) = c(s_t(\omega), w_t, r_t, z_t, H(\omega))$$

Assume homogeneity of beliefs. Then $s_t(\omega) = s_t = k_t$ and under SP learning the recursive system is

$$\begin{aligned}
\mathcal{R}_t &= \mathcal{R}_{t-1} + \gamma_t(\tilde{x}_{t-1}\tilde{x}'_{t-1} - \mathcal{R}_{t-1}) \text{ for } \tilde{x}'_t = (1, k_t, z_t) \\
H_t &= H_{t-1} + \gamma_t \mathcal{R}_t^{-1}\tilde{x}_{t-1}(\lambda_{t-1} - H'_{t-1}\tilde{x}_{t-1}) \\
z_t &= \varepsilon_t z_{t-1}^{\rho} \\
w_t &= z_t(f(k_t) - f'(k_t)k_t) \\
r_t &= z_t f'(k_t) - \delta \\
c_t &= c\,(k_t, r_t, w_t, z_t, H_t) \\
k_{t+1} &= (1 + r_t)k_t + w_t - c_t \\
\lambda_t &= (1 + r_t)U'(c_t)
\end{aligned}$$

Illustration

For log utility, Cobb-Douglas production, and $\delta = 1$, we can obtain the explicit RE solution and analytical REE shadow price λ_t^* function.

The red line is initial beliefs. Under learning there is convergence to the black dashed line. The dashed blue line is the $\lambda^*(k)$ in the REE.

Long-run beliefs appear to coincide to first order with $\lambda^*(k)$ of the REE.

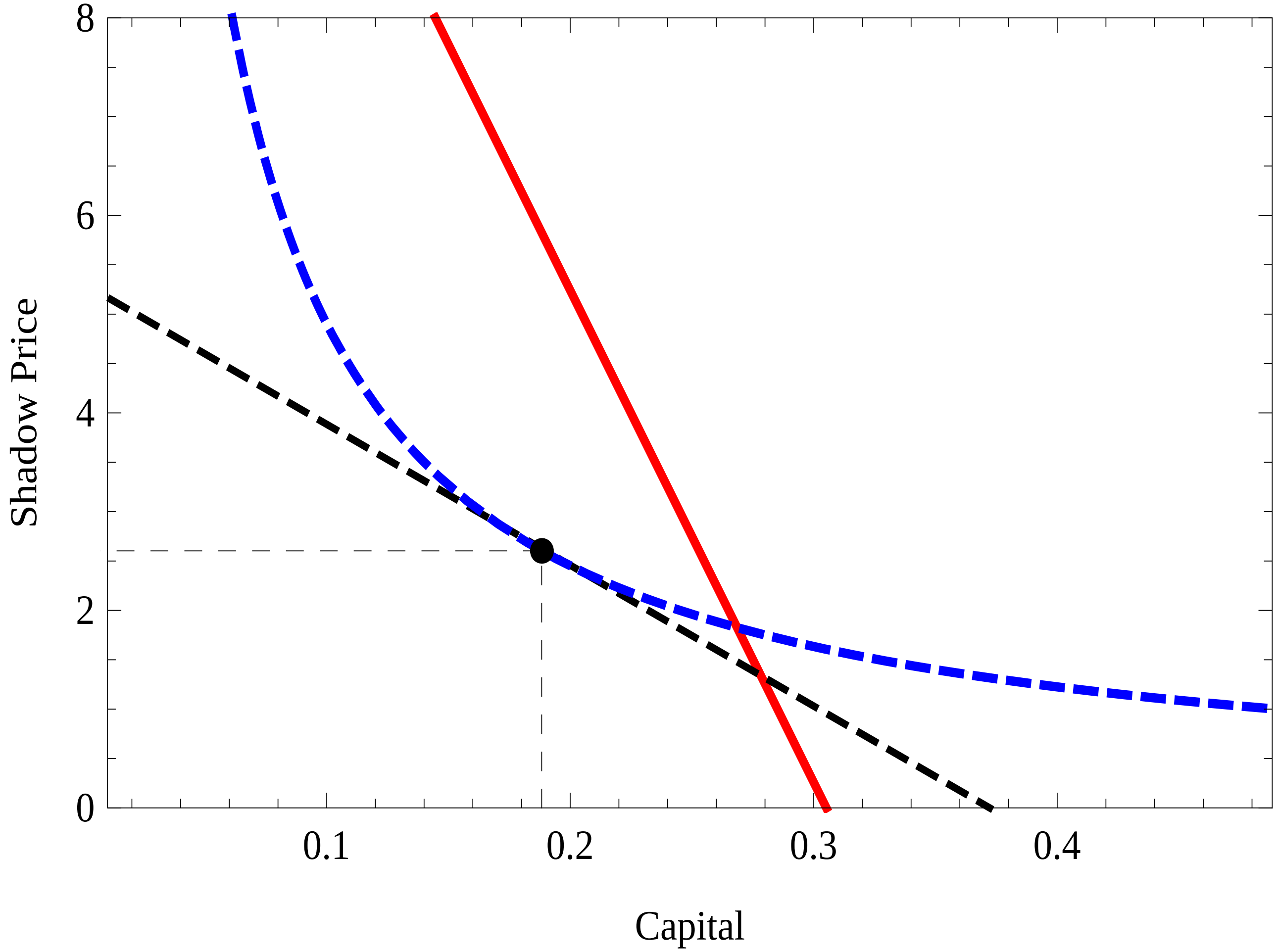


Figure 3: Red: initial beliefs. Black: final beliefs. Blue: true $\lambda(k)$ in REE.

Example: SP Learning in a NK Model

- A final example: a NK model (without capital) with a Rotemberg price-adjustment friction. The utility flow for household-firm ω is

$$\frac{1}{1-\sigma}\left(c_t(\omega)^{1-\sigma}-1\right)+\log\left(\frac{m_{t-1}(\omega)}{\pi_t(\omega)}\right)-\frac{h_t(\omega)^{1+\chi}}{1+\chi}$$
$$-\frac{\gamma}{2}\left(\frac{p_t(\omega)}{p_{t-1}(\omega)}-1\right)^2+\mathcal{L}(b_{t-1}(\omega)).$$

  We consider both homogeneous and heterogeneous expectations cases.

- ω treats $z_t=(1,R_{t-1},\pi_t,p_t,y_t,\mathcal{T}_t,g_t,\varphi_t)$ as exogenous. g_t is an AR(1) govt. spending shock and φ_t is an AR(1) interest-rate shock.

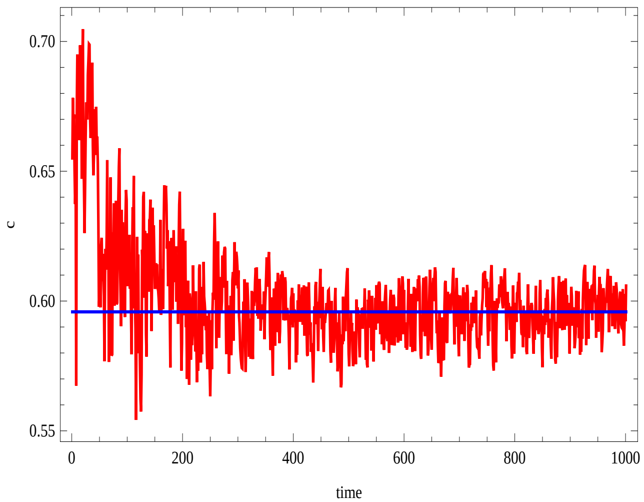
- The endogenous states and the controls of ω are

$$
\begin{aligned}
s_t(\omega) &= (m_{t-1}(\omega), b_{t-1}(\omega), p_{t-1}(\omega))' \\
u_t(\omega) &= (c_t(\omega), m_t(\omega), p_t(\omega))'.
\end{aligned}
$$

- Agent ω requires shadow prices for $\lambda_t(\omega) = (\lambda_t^b(\omega), \lambda_t^p(\omega))'$. We assume PLM $\lambda_t = \mathbf{\Psi}(\omega) x_t(\omega)$, for regressors $x_t(\omega)$.

- We set $x_t(\omega) = (1, g_t, \varphi_t, R_{t-1})$ in the homogeneous case and $x_t(\omega) = (1, g_t, \varphi_t, R_{t-1}, b_{t-1}(\omega))$ in the heterogeneous case.

With homogeneous expectations and small constant gain learning ($\hat{\gamma} = 0.015$) we find (Figure 4) convergence of beliefs to a distribution centered on the SP learning restricted perceptions equilibrium (RPE).



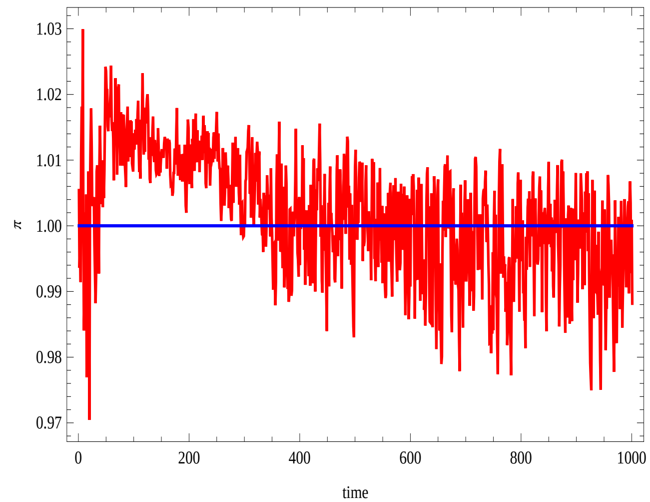


Figure 4: c_t, π_t in NK model, homogeneous agents.

- Now consider case of heterogeneous agents. Agents can borrow from or lend to each other. We introduce a disutility of debt above a certain level. Suppose two agents types.

- Different beliefs leads to different savings rates, bond and money holdings and prices $p_{t-1}(\omega)$.

- Note: simulations under SP learning are not greatly more difficult than for the homogeneous case.

- Figure 5 shows simulations with zero initial debts but initial heterogeneous beliefs.

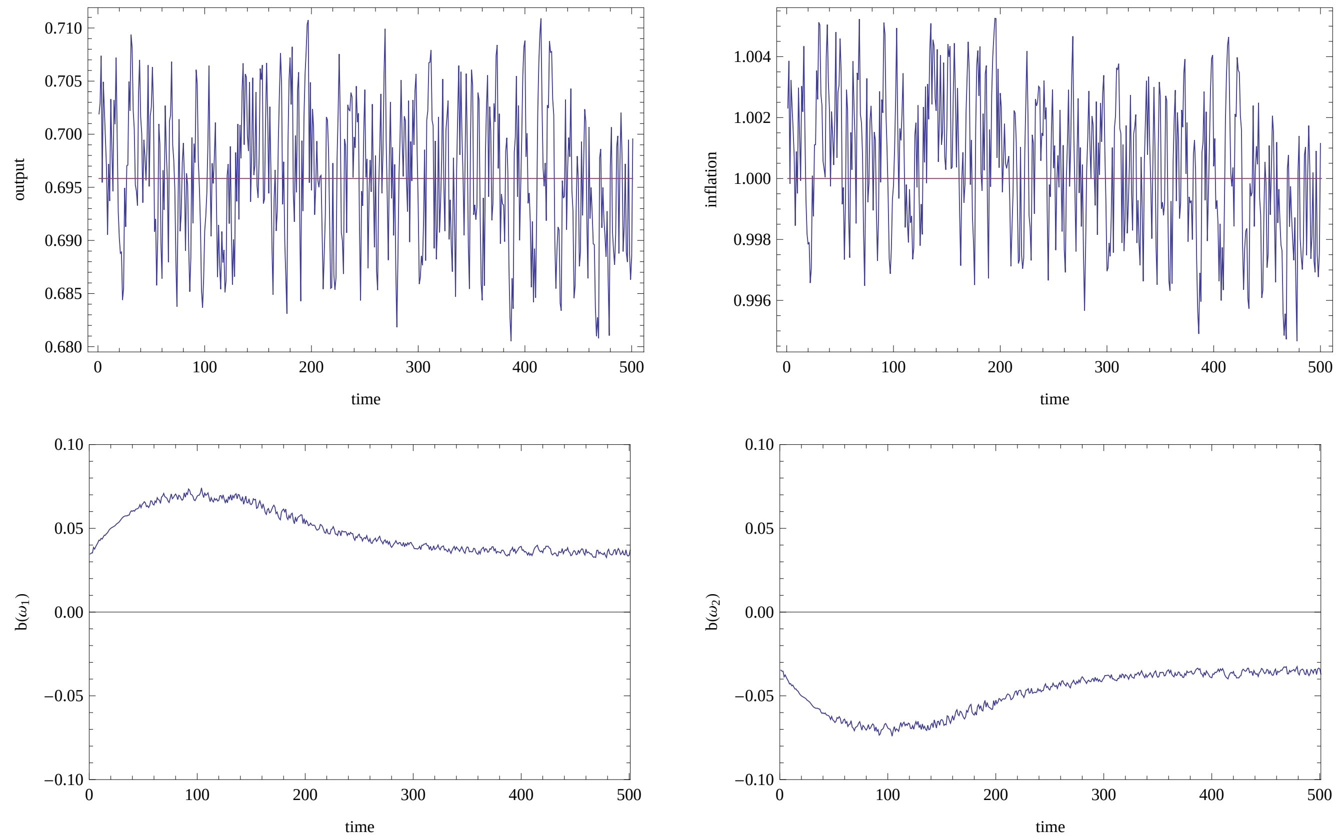


Figure 5: Paths under SP learning with heterogeneous initial beliefs for the two types of agents. Dislike of debt above a specified threshold.

Conclusions

- SP learning can be applied in general dynamic stochastic optimization problems and within general equilibrium models.

- The approach is formulated at the agent level and allows for heterogeneity in general equilibrium settings.

- It is tractable because agents need only solve 2-period optimization problems using one-step ahead forecasts of states and shadow prices.

- SP learning is boundedly optimal but converges to optimal decisions.

- Current work – Applications:

  – SP learning in general equilibrium models with heterogeneous agents.

  – Develop a general procedure for implementing SP-learning in this setting.

  Current work – Extensions:

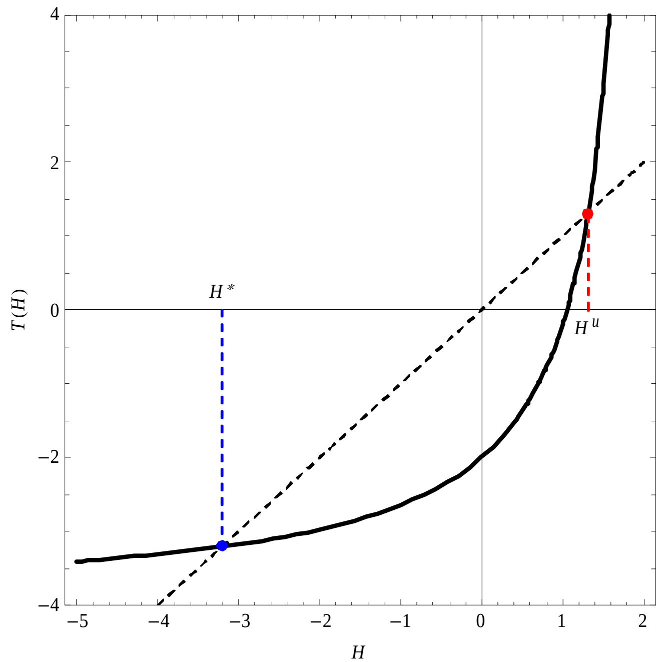
  – SP learning with inequality constraints (e.g. borrowing constraints).

  – Value function learning in qualitative choice models.

  – Study implications of persistent deviations from full optimization due to misspecified shadow price models or "escape paths" under constant gain learning.

ADDENDA

Illustration of need for Projection Facility (PF): Without PF an unusual sequence of shocks can lead to perceptions H that impart explosive dynamics.



Univariate T-map.

Optimal decision-making in Investment model

The key SP learning behavioral equation is

$$J + q_t(\omega) + \gamma I_t(\omega) = \mu p_t f'(z_t(\omega) + \mu I_t(\omega)) + \beta(1 - \delta\mu)\hat{E}_t \lambda^z_{t+1}(\omega).$$

This requires a shadow price forecast made based on:

$$\lambda^z_t(\omega) = H_0 + H_1 v_t + H_2 q_t(\omega) + H_3 p_t + H_4 z_t(\omega).$$

$\hat{E}_t \lambda^z_{t+1}(\omega)$ is obtained based on estimated H and forecasts $\hat{E}_t v_{t+1} = \rho v_t$, $\hat{E}_t q_{t+1}(\omega) = \bar{q}$, $\hat{E}_t p_{t+1}$ from the estimated p_t equation and $\hat{E}_t z_{t+1}(\omega)$ from the known transition equation

$$\hat{E}_t z_{t+1}(\omega) = (1 - \delta) z_t(\omega) + (1 - \delta\mu) I_t(\omega).$$

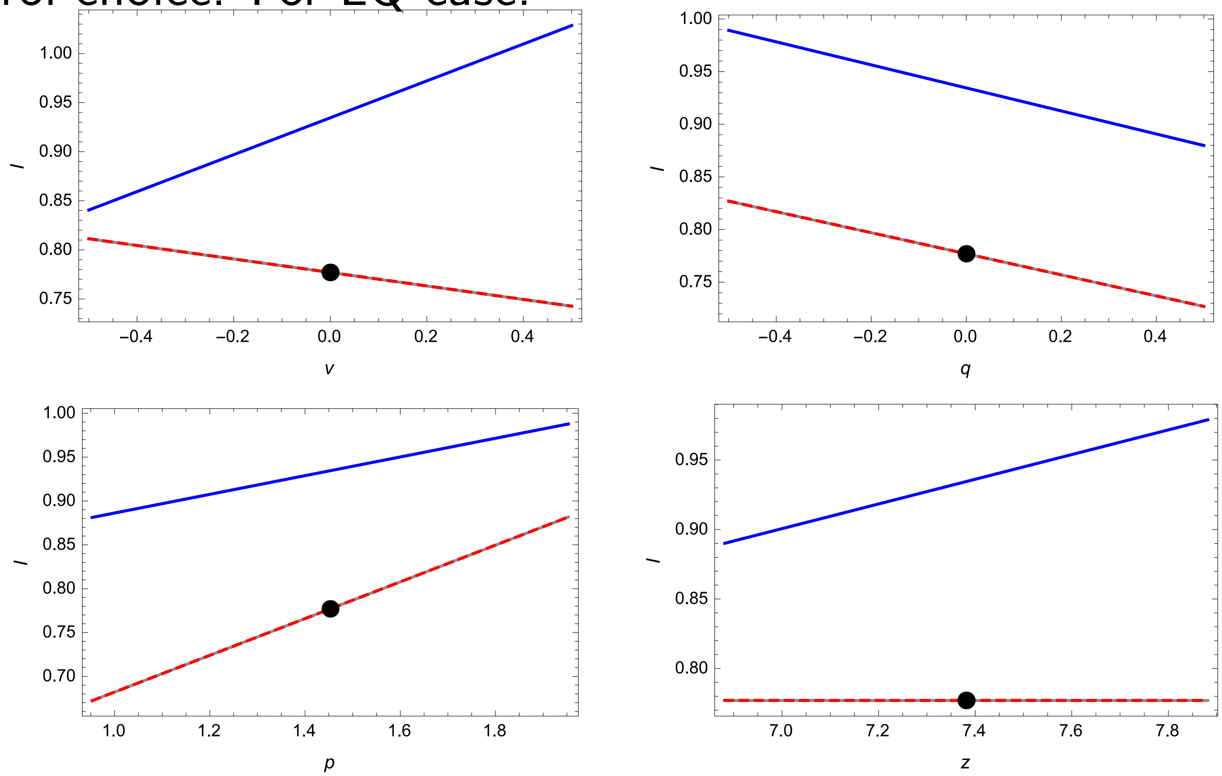
Given $(z_t(\omega), q_t(\omega), p_t)$ the control $I_t(\omega)$ and forecasts $\hat{E}_t \lambda^z_{t+1}(\omega)$ can be solved simultaneously.

The “envelope” equation

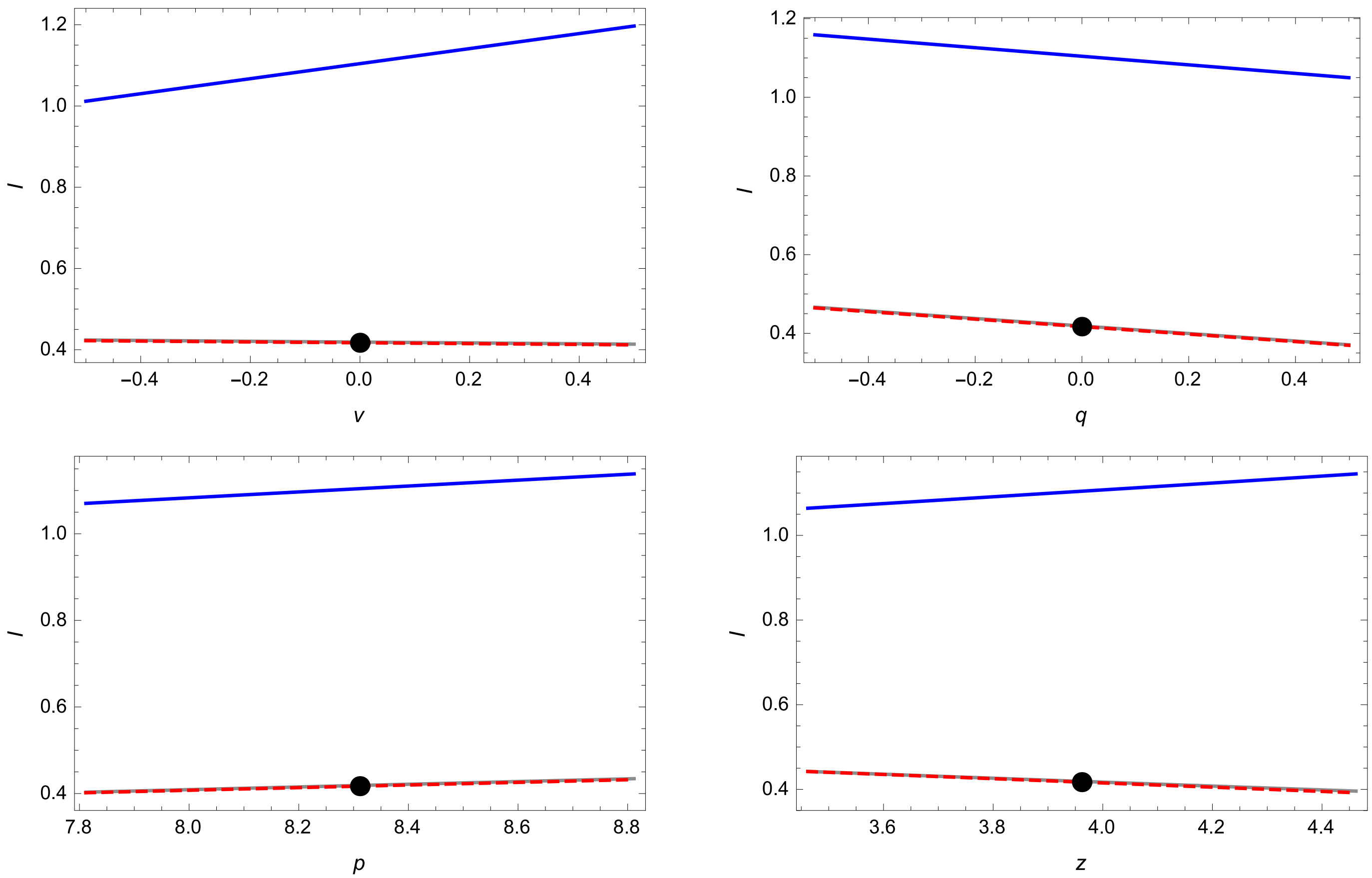
$$\lambda_t^z(\omega) = p_t f'(z_t(\omega) + \mu I_t(\omega)) + \beta(1 - \delta)\hat{E}_t \lambda_{t+1}^z(\omega)$$

is then used to obtain a new data point for $\lambda_t^z(\omega)$. This is used for updating estimates H.

Investment model, control choices: The first figure shows control choice against the indicated state (other states at steady state values), with blue = control choice for initial beliefs, gray = control choice after 350 periods and red = optimal control choice. For LQ case:



Control choices in non-LQ case



Euler-equation learning in Crusoe model

EE learning is also possible using a second-order Euler equation:

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

To implement use forecasts of $\hat{E}_t c_{t+i}$ from estimates of

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

SP learning and EE-learning are not identical, but both are asymptotically optimal. This can be seen from a numerical calculation of their largest eigenvalue, shown in Figure 1.

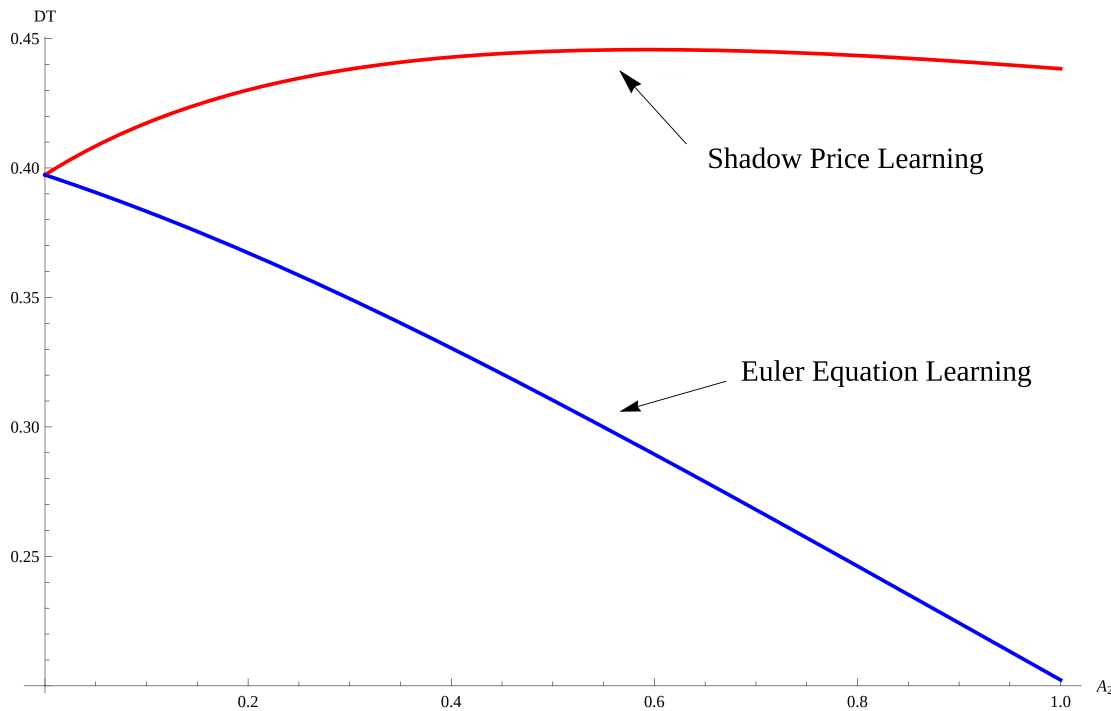


Figure 2: Largest eigenvalue of DT^{SP} and DT^{EL} under SP and EE learning.

Why are EE-learning and SP learning different?

Here $\text{dim}(u) = 1$ and $\text{dim}(x) = 2$. The PLMs are

$$\text{SP PLM: } \lambda_t = Hx_t \text{ vs EE PLM: } u_t = Fx_t$$

so SP learning estimates 4 parameters whereas EE learning estimates 2 parameters.

The SP PLM requires less structural information than the EE PLM. For the SP PLM to be equivalent to the EE PLM, agents would need to understand the structural relation between λ_1 and λ_2 and to impose this restriction in estimation.

NK decision rules

$$c_t(\omega)^{-\sigma} = \beta \hat{E}_t \lambda^b_{t+1}(\omega)$$

$$\begin{aligned} & \beta(\nu - 1)\left(\frac{p_t(\omega)}{p_t}\right)^{1-\nu}\left(\frac{y_t}{p_t}\right)\hat{E}_t \lambda^b_{t+1}(\omega) \\ = \quad & -\frac{\gamma(\pi_t(\omega) - 1)}{p_{t-1}(\omega)} - h_t(\omega)^{\chi} h_{p(\omega)}(t) + \beta \hat{E}_t \lambda^p_{t+1}(\omega) \end{aligned}$$

$$m_t(\omega) = \beta R_t (R_t - 1)^{-1} c_t(\omega)^{\sigma}$$

NK aggregation

$$c_t(\omega) = \left(\int c_t(\omega, \varpi)^{\frac{\nu-1}{\nu}} d\varpi \right)^{\frac{\nu}{1-\nu}}$$

$$\begin{aligned} y_t(\varpi) &= \int c_t(\omega, \varpi) d\omega + g_t(\varpi) \\ y_t &= \int c_t(\omega) d\omega + g_t \\ m_t^s &= \int m_t(\omega) d\omega \\ 0 &= \int b_t(\omega) d\omega. \end{aligned}$$

NK model disutility of debt

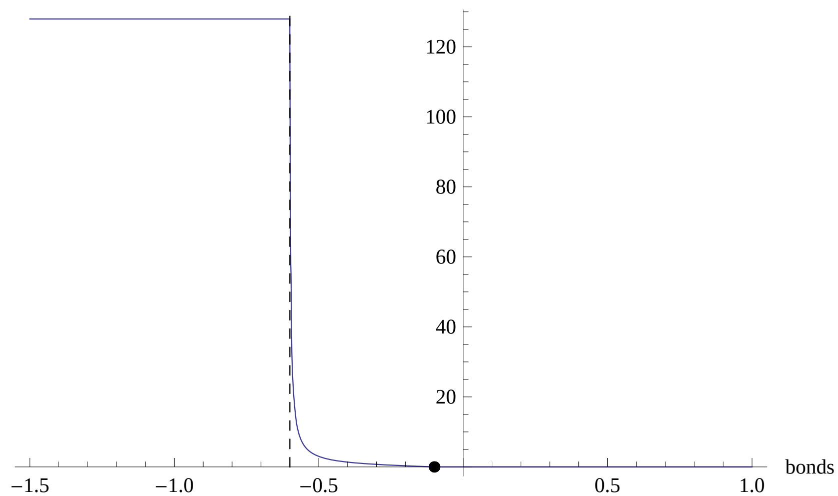


Figure 4: Marginal disutility of debt in NK model.

In Figure 6 we start with homogeneous beliefs but different debt levels. We also assume a stronger dslike of debt, i.e. $\hat{b} = 0$. In this case lending between agents goes to zero asymptotically.

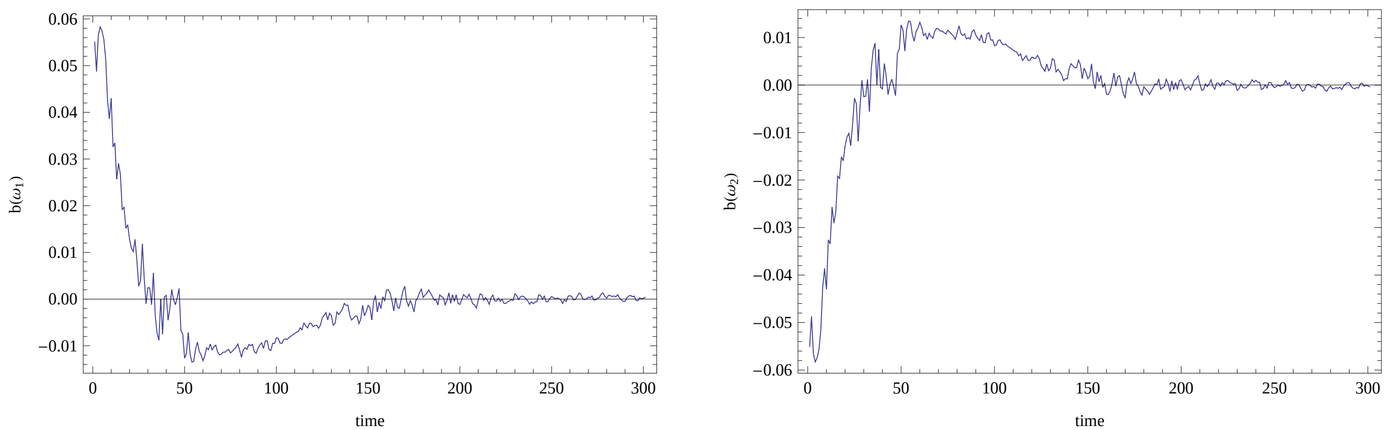


Figure 6: Heterogeneous initial bonds but homogeneous beliefs.