Boundedly Rational Decision Making in Continuous-Time

Chandler Lester

Working Draft of Ch.2 of My Dissertation

Abstract

Continuous-time macroeconomic literature has grown remarkably in recent years. As work on continuous-time models becomes, more prevalent macroeconomists need to adapt essential discrete-time methods to continuous-time. This paper modifies adaptive learning techniques to continuous-time. One approach to accomplish this task is shadow-price learning (SP-learning), a framework in which agents forecast their expected shadow prices. To use this framework efficiently, we first need a tractable continuous-time linear-quadratic (LQ) environment. While discrete-time LQ problems are common in the literature, there is very little work on continuous-time LQ problems. Thus, our contributions are two-fold. We build a continuous-time LQ framework for solving Hamilton-Jacobi-Bellman (HJB) equations using iterative methods and implement adaptive learning techniques in this new setting.

1 Introduction

The macroeconomics toolkit has significantly expanded in recent years due to increased access to computational power and interdisciplinary research. One promising modeling framework emerging from this development is stochastic continuous-time modeling. Continuous-time models have existed in economics literature for over thirty years, becoming popular during the period Black and Scholes (1973) was first published. During this time economists published papers using the continuous-time framework including, Brock and Mirman (1972), Merton (1969, 1975), and Mirrlees (1971). However, many of these works could only examine specific aspects of models, such as the steady-state distribution of key parameters, as economists did not have techniques for solving the systems of partial differential equations that represent most continuous-time models. Now, with methods drawn from the field of applied mathematics, it has become feasible to solve more continuous-time macroeconomic models.

Continuous-time macroeconomic models have become increasingly popular for two distinct reasons. First, the field of finance has long favored continuous-time modeling, thus building macroeconomic models in continuous-time allows economists to include financial frictions as in Brunnermeier and Sannikov (2014). Second, as we previously mentioned, solutions to many macroeconomic models can now be easily found—because of better computers and new solution methods—and these solutions often include detailed distributional information. Several works that take advantage of this property are Ahn et al. (2018), Achdou et al. (2020), Kaplan et al. (2018) and Gabaix et al. (2016). As this class of models becomes popular, economists must redevelop traditional macroeconomic modeling techniques to create richer models in this continuous-time framework. This paper modifies adaptive learning techniques for use with continuous-time economies.

Currently, the continuous-time macroeconomic literature consists primarily of models that depend on rational expectations. Rational expectations is a standard modeling technique where agents within economics are assumed to understand theoretical models correctly—the agents know the value of all parameters in the model and understand the distribution of any unobserved processes. It is improbable that individuals in the real world have this level of knowledge about the economy. However, individuals can likely perceive the world around them and gradually adjust their expectations based on their observations—adaptive learning takes this approach.

Allowing for adaptive learning, as opposed to rational expectations, in macroeconomic models avoids allowing agents to have unrealistic amounts of information about the system by instead allowing them to gather information on the economy over time slowly. This technique was developed initially in Bray (1982) and been further refined in more recent work Evans and Honkapohja (2001). Adaptive learning is an attractive modeling tool since rational expectations often make to many strict assumptions about agents' knowledge of parameter values and the distribution of parameters.

Additionally, adaptive learning models often converge to a rational expectations equilibrium over time; however, if a model has two rational expectations equilibria, an adaptive learning model may only converge to one—the equilibria learned by these agents would then be stable under adaptive learning whereas the other equilibria would not. Therefore adaptive learning techniques are beneficial when economists want to examine the stability or particular outcomes.

Despite this, rational expectations is a standard model assumption and the emerging continuous-time literature centers on rational expectations models—some continuous-time asset pricing models use Bayesian methods, for instance, Hansen and Sargent

(2019a) and Hansen and Sargent (2019b). However, these methods require agents' to have prior belief over the distribution of parameters another strong assumption. We instead concentrate on an adaptive learning technique called shadow-price learning, or SP-learning, outlined in Evans and McGough (2018). Under SP-learning agents view their optimization problem as a two-period problem.

During the first period (today), they use a forecast of their shadow-price to form the best possible choices for today, given those choices' impacts on tomorrow (the second period). Hence this learning mechanism focuses on an agent's ability to generate optimal forecasts and the agent's ability to make optimal decisions with the forecasted information, an issue discussed in (Marimon and Sunder, 1993, 1994; Hommes, 2011). In continuous-time, this problem is very similar; however, instead of having today and tomorrow, the agents examine the trade-off between choices using the change in parameters over time—in other words—the continuous-time version of SP-learning examines derivatives of variables with respect to time.

We develop a tractable setting for SP-learning by building a continuous-time linear-quadratic (LQ) framework. The LQ environment aids the study of adaptive learning techniques due to the linearity of first-order conditions, generality, and certainty equivalence in this framework. In economics, the LQ framework is useful for approximations of complex economies since these models can contain lots of information. There is wide-ranging literature on discrete-time economic optimal linear regulator problems that includes several works on optimal policies such as Benigno and Woodford (2004) and Benigno and Woodford (2006), as well as a wealth of papers on techniques and developing the LQ framework in economics, Kendrick (2005), Amman and Kendrick (1999), and Benigno and Woodford (2012). Because of the richness of this framework and the sparse usage of continuous-time LQ problems in economics, further exploration of this technique is necessary.

Although continuous-time LQ problems are not common in economics, some economists have examined this type of modeling framework. Hansen and Sargent (1991) develops a framework for continuous-time LQ problems. Several chapters of this book examine various models and the identification of parameters in this setting. The LQ framework we build in this paper differs from Hansen and Sargent (1991), as it does not use solution methods based on the Lagrangian. Instead, we take a value function approach. Value function methods are conventional in the discrete-time economics literature, and many continuous-time problems in other fields feature similar solution methods.

We build this framework by outlining a basic discrete LQ problem and then describing a similar continuous-time problem, using a value function approach for both settings. We work through both types of problems, so those familiar with only the discrete case can more easily see the parallels between these two settings. After setting up the LQ problems, we look at solution methods for the resulting algebraic Riccati equations (AREs). Though there are many methods for solving AREs, we concentrate on iterative Newtonian methods, as in Kleinman (1968), as this method better complements the adaptive learning environment in later sections. Also, discrete-time LQ systems commonly use iterative methods (Hansen and Sargent, 2013).

After developing a continuous-time LQ framework, we can then examine continuous-time adaptive learning rules. Before reworking discrete-time adaptive learning rules into continuous-time rules, we need to consider several important items. First, does an agent have "continuous" observations of continuous variables, or do they have discrete observations? If these observations are discrete, are they taken at specific points in time or over intervals, and does the spacing of these points or intervals matter?

We take a simplified approach, drawing from empirical economics and finance literature. Bergstrom (1993), a general survey of continuous-time econometric methods, highlights that continuous-time systems can be measured accurately with exact discrete-time equivalents that take time-interval lengths into account, a conclusion initially drawn from Phillips (1959) and discussed further in Bergstrom (1984). In finance, Kellerhals (2001) uses discrete-time data to measure continuous-time financial systems while carefully implementing exact discrete-time models as in the economics literature. Additional work on this topic includes Aït-Sahalia (2010), which examines the maximum likelihood estimation of continuous model parameters using discrete data points. All of these works find that it is possible to measure continuous-time systems with discrete data.

When using learning algorithms to forecast an agent's perception of the model, we implement the exact discrete-time method since—despite the model parameters evolving continuously—as it is most likely that agents observe the data discretely but at fine intervals. The agents observe data as it becomes available, and they observe all data points. Concentrating on this approach for the agent's sampling of the data allows for more direct tie-ins with typical discrete learning methods. Extensions to this work may include observation intervals that vary from the data generating process's time intervals and data that arrive at random intervals.

The contributions of this work are two-fold. First, to create a modeling framework in which we can develop adaptive learning techniques, we construct a novel continuous-time LQ framework. We outline this framework and discuss it in detail in sections 2 and 3. Continuous-time optimal linear regulator problems similar to those outlined in this paper do exist in other disciplines; however, problems outside of economics do not usually include key features such as stochasticity and discounting. Second, we use this new LQ framework to develop continuous-time shadow-price learning in section 4. Also, we demonstrate parallels between the discrete and continuous models and derive a continuous-time version of recursive least squares (RLS).

The bulk of this is done in section 2.2 and section 4.

The paper precedes as follows. Section 2 builds a simple LQ problem without interaction terms or stochasticity. This section also examines iterative solution methods with a univariate test case and convergence of the discrete test case to the continuous one under small time increments. Section 3 studies a more complicated univariate model with stochasticity as well as this model's solutions, the convergence results with the equivalent discrete-time model. Preliminary results for a simple learning algorithm and the convergence of a discrete-time learning rule to the continuous solution are discussed in section 4. We evaluate a simple economic model in section 5; the model used is a simple Robin Crusoe economy as in Evans and McGough (2018). Section 6 concludes.

2 The Optimal Linear Regulator Problem

Before examining a continuous-time LQ problem, we start with a review of a typical deterministic discrete case. A simple deterministic optimal linear-quadratic (LQ) problem can be expressed according to the following equations (Ljungqvist and Sargent, 2012),

$$V(x_0) = \max -\mathbb{E} \sum_{t=0}^{\infty} \beta^t \{x_t' R x_t + u_t' Q u_t\} \tag{1}$$

where x_t evolves according to

$$x_{t+1} = Ax_t + Bu_t. \tag{2}$$

Throughout this work, x_t is an $(n \times 1)$ vector of state parameters, in an economic setting x_t might include variables like capital or technological processes. For our purposes, x_t always includes a constant; however, the constant is not necessary (Hansen and Sargent, 2013). Additionally, u_t is a $(m \times 1)$ vector of control processes, again in an economic setting u_t might include investment. In the system above there no interaction terms between x_t and u_t. Often there are interaction terms in LQ problems; however, we have left them out to make the problem simpler. Using the above equations, we can write the Bellman system as,

$$V(x_t) = \max_u \ \{-x_t' R x_t - u_t' Q u_t + \beta \mathbb{E} V(x_{t+1})\}. \tag{3}$$

To solve the Bellman in the LQ framework we first guess that $V(x_t) = -x_t' P x_t$, where P is a positive semi-definite matrix (Hansen and Sargent, 2013). Thus, $V(x_{t+1}) = -x_{t+1}' P x_{t+1} = -(Ax_t + Bu_t)' P (Ax_t + Bu_t)$. Substituting these guesses for $V(x_t)$ and $V(x_{t+1})$ into (3) yields,

$$-x' P x = \max_u \{-x' R x - u' Q u - \beta (Ax + Bu)' P (Ax + Bu)\}. \tag{4}$$

Now if we look at the first order condition with respect to u, we will get

$$-2u' Q - 2\beta u' B' P B = 2\beta A' x' P B$$

or

$$u = -\beta (Q + \beta B' P B)^{-1} (B' P A) x = -F x \tag{5}$$

we can now eliminate u in the Bellman system and write a recursive solution for P using the Riccati equation,

$$P_{j+1} = R + \beta A' P_j A - \beta^2 A' P_j B (Q + \beta B' P_j B)^{-1} B' P_j A \tag{6}$$

where j denotes the iterations. Using this recursive system, we can find the solution to the discrete-time ARE under certain assumptions. We will focus on the stability conditions for the continuous-time case; for a treatment of the discrete-time case see Hansen and Sargent (2013), Lewis (1986), or Anderson and Moore (2007).

2.1 The Continuous-Time Optimal Linear Regulator

This problem can be similarly approached in continuous-time. We now examine the continuous-time optimal linear regulator problem using a system similar to—but not the same as—the one in the previous section. Additionally, to simplify arithmetic for this problem, we assume that A is symmetric.[1] In the continuous-time setting, our maximization problem becomes,

$$V(x_0) = \max -\mathbb{E}\int_{t=0}^{\infty} e^{-\rho t}\{x_t'Rx_t + u_t'Qu_t\}dt \tag{7}$$

where x_t evolves according to,

$$dx_t = Ax_t dt + Bu_t dt \tag{8}$$

Equation (8) is a standard expression of a continuous-time process, in continuous-time the levels of variables over time do not summarize their evolution—instead the changes in a variable describe how it grows over time Dixit (1992).

The continuous problem relies on the Hamilton-Jacobi-Bellman (HJB) equation as opposed to the discrete-time Bellman system. First, we write down our problem discretely using the power series expansion of $e^{-\rho\Delta}$, $(1-\rho\Delta)$, as a representation of our discount over a period of time (Dixit, 1992). Here Δ represent the increments of

[1] In the appendix there is a version of this problem that does not assume A is symmetric.

the time periods.

$$V(x_t) = \max_u \{-x_t' R x_t \Delta - u_t' Q u_t \Delta + (1-\rho\Delta)\mathbb{E}[V(x_{t+\Delta})]\}. \tag{9}$$

After simplifying the system and taking the limit as $\Delta \to 0$, the HJB simplifies to

$$\rho V(x) = \max_u \Big(- x' R x - u' Q u + V_x(x)\frac{dx_t}{dt}\Big) \tag{10}$$

Now, if we guess that $V(x) = -x'Px$ equation (10) becomes

$$\begin{aligned} -\rho x' P x &= \max_u \Big(- x' R x - u' Q u - 2x' P \frac{dx_t}{dt}\Big) \\ &= \max_u \{-x' R x - u' Q u - 2x' P (Ax + Bu)\} \end{aligned} \tag{11}$$

Taking the first order condition with respect to u yields,

$$-2u'Q - 2x'PB = 0$$

or

$$u = -Q^{-1}B'Px = -\tilde{F}x. \tag{12}$$

This equation is our policy function for u in the continuous-time system. Note that the policy for u is not the same as the discrete case policy. We should expect the policies for the discrete and continuous-time cases to differ, since expectations[2] and discounting in discrete and continuous-time vary.

[2]In discrete-time, $\mathbb{E}[V(x_{t+1}] = x_{t+1}Px_{t+1} = (Ax_t + Bu_t)P(Ax_t + Bu_t)$. While in continuous-time expectations depend on Itô's lemma, $\mathbb{E}[V(x_{t+\Delta})] = V_x(x)\frac{dx_t}{dt} = 2x'P(Ax + Bu)$.

Plugging (12) into (11) produces

$$-\rho x'Px = -x'Rx - x'PBQ^{-1}B'Px - 2x'PAx + 2x'PBQ^{-1}B'Px.$$

Simplifying further yields,

$$R + 2PA - PBQ^{-1}B'P - \rho P = 0. \tag{13}$$

To solve this system we will begin with the Lyapunov equation for our optimal linear regulator problem,

$$2\tilde{A}_i'P_i = -(R + \tilde{F}_i'Q^{-1}\tilde{F}_i).$$

Here, $\tilde{A}_i = A - \frac{1}{2}I\rho - B\tilde{F}_i$, $\tilde{F}_i = Q^{-1}B'P_{i-1}$, and i indexes each iteration. Subtracting, $2\tilde{A}_i'P_{i-1}$ from both sides will yield,

$$2\tilde{A}_i'(P_i - P_{i-1}) = -2\tilde{A}_i'P_{i-1} - \tilde{F}_i'Q^{-1}\tilde{F}_i + R. \tag{14}$$

we can then rewrite this as,

$$P_i = P_{i-1} - (2\tilde{A}_i')^{-1}(2\tilde{A}_i'P_{i-1} - \tilde{F}_i'Q^{-1}\tilde{F}_i + R). \tag{15}$$

Alternatively we can avoid subtracting $\tilde{A}_i'P_{i-1}$ from both sides and get,

$$P_i = -(2\tilde{A}_i')^{-1}(\tilde{F}_i'Q^{-1}\tilde{F}_i + R). \tag{16}$$

To ensure solutions to (15) and (16) are asymptotically stable and exist, several conditions must be met (Lewis, 1986; Anderson and Moore, 2007; Evans and McGough, 2018).

LQ.1 The matrix R is symmetric positive semi-definite and thus can be decomposed in $R = DD'$ by rank-decomposition, and the matrix Q is symmetric positive definite.

LQ.2 The matrix pair (A,B) is *stabilizable*—there exists a matrix $\tilde{F}$ such that $A - B\tilde{F}$ is stable, meaning the eigenvalues of $A - B\tilde{F}$ have modulus less than one.

LQ.3 The pair (A,D) is *detectable*—if y is a non-zero eigenvector of A associated with eigenvector μ then $D'y = 0$ only if $|\mu| < 0$. Detectability implies that the feedback control will plausibly stabilize any unstable trajectories.

The conditions outlined in LQ.1-LQ.3 are standard in optimal linear regulator literature and are necessary for stable solutions in both discrete and continuous time-invariant problems. LQ.1 can be interpreted as a condition on the concavity of the system, making sure that the system in bounded above. Additionally, LQ.2 ensures that the value function $V(x)$ does not become infinitely negative by guaranteeing that it is possible to find a policy F that drives the state x to zero.

Theorem 1. *If the conditions outlined in LQ.1-LQ.3 are true, then the continuous-time ARE has a unique positive semi-definite solution P*

For a proof of theorem 1 see Lewis (1986).

2.2 Convergence of the Discrete Case to the Continuous Case

The discrete and continuous LQ problems outlined in the previous sections had different solutions because these systems have several differences that cause them to evolve differently over time. In this section, we rewrite the discrete problem and show that the discrete solution converges to the continuous one under certain variable transformations as the discrete-time intervals become increasingly small.

Theorem 2. *The discrete-time system outlined in* (1) *and* (2) *can be transformed so that its' solutions converge to the solutions from the continuous-time system outlined in* (7) *and* (8).

Proof. To begin, we start with the typical continuous-time system given by equations (7) and (8). To discretize this system, we rewrite (7) as a summation over time periods that increment over integers and an integral over individual time increments, Δ.

$$-\mathbb{E}\sum_{k=0}^{\infty}\int_{t=\Delta k}^{\Delta(k+1)}\{e^{-\rho t}(x_t'Rx_t+u_t'Qu_t)\}dt=-\mathbb{E}\sum_{k=0}^{\infty}\int_{\Delta k}^{\Delta(k+1)}\{e^{-\rho t}f(x_t,u_t,t)\}dt \quad (17)$$

For convenience the boundaries on the integral will be changed from $(\Delta k,\Delta(k+1))$ to $(0,\Delta)$, thus $f(x_t,u_t,t)$ must be transformed to $f(x_{t+s},u_{t+s},t+s)$ and integrated over ds. Using a Taylor approximation, the function becomes,

$$\begin{aligned}f(x_{\Delta k+s},u_{\Delta k+s},\Delta k+s)=&x_{\Delta k}'Rx_{\Delta k}+u_{\Delta k}'Qu_{\Delta k}+2x_{\Delta k}'R(x_{\Delta k+s}-x_{\Delta k})+2u_{\Delta k}'Q(u_{\Delta k+s}-u_{\Delta k})\\&+R(x_{\Delta k+s}-x_{\Delta k})^2+Q(u_{\Delta k+s}-u_{\Delta k})^2.\end{aligned}$$

This function can be further simplified since $x_{\Delta k+s}-x_s=(Ax_{\Delta k}+Bu_{\Delta k})s$ and $u_{\Delta k+s}-u_t=\dot{u}s$ where $\dot{u}$ is a smooth function that summarizes that change in u over an increment of time. Using these substitutions only a few terms in the function will remain—as $s^2\approx 0$ in the continuous-time limit,

$$f(x_{\Delta k+s},u_{\Delta k+s},\Delta k+s)=x_{\Delta k}'Rx_{\Delta k}+u_{\Delta k}'Qu_{\Delta k}+2x_{\Delta k}'R(Ax_{\Delta k}+Bu_{\Delta k})s+2u_{\Delta k}'Q\dot{u}s$$

plugging this into (17) yields,

$$-\mathbb{E}\sum_{k=0}^{\infty}\int_{s=0}^{\Delta} e^{-\rho(\Delta k+s)}\{x'_{\Delta k}Rx_{\Delta k}+u'_{\Delta k}Qu_{\Delta k}+2x'_{\Delta k}R(Ax_{\Delta k}+Bu_{\Delta k})s+2u'_{\Delta k}Q\dot{u}s\}ds$$

Focusing on the inter integral,

$$\int_{s=0}^{\Delta} e^{-\rho(\Delta k+s)}\{x'_{\Delta k}Rx_{\Delta k}+u'_{\Delta k}Qu_{\Delta k}+2x_{\Delta k}R(Ax_{\Delta k}+Bu_{\Delta k})s+2u'_{\Delta k}Q\dot{u}s\}ds$$
$$=-\frac{1}{\rho}e^{-\rho\Delta k}[e^{-\rho\Delta}-1](x'_{\Delta k}Rx_{\Delta k}+u'_{\Delta k}Qu_{\Delta k}).$$

Plugging this result[3] back into the main summation term and replacing k with t while setting $\hat{x}_t = x_\Delta$, $\hat{u}_t = u_\Delta$, and $\hat{\rho} = \rho\Delta$ yields,

$$-\mathbb{E}\sum_{t=0}^{\infty}\frac{1}{\hat{\rho}}e^{-\hat{\rho}t}[1-e^{-\hat{\rho}}](\hat{x}'_t R\hat{x}_t+\hat{u}'_t Q\hat{u}_t)\Delta \tag{18}$$

to get this into the typical discrete LQ format, as in (1), β, R, and Q must be appropriately transformed. The discount factor β becomes $\beta(\Delta) = e^{-\hat{\rho}}$, R is now $R(\Delta) = \frac{1}{\rho}(1-e^{-\hat{\rho}})R$, and $Q(\Delta) = \frac{1}{\rho}(1-e^{-\hat{\rho}})Q$.

Lastly, the equation for the evolution of the state variables must be transformed by applying the Euler-Maruyama method to equation (8) yielding,

$$x_{\Delta(t+1)} = (I + A\Delta)x_\Delta + B\Delta u_\Delta \tag{19}$$

where I is an identity matrix. Thus the transformed coefficients are $A(\Delta) = (I + A\Delta)$ and $B(\Delta) = B\Delta$. $\square$

[3]The term $\int_0^\Delta e^{-\rho(\Delta k+s)}\{2x_{\Delta k}R(Ax_{\Delta k}+Bu_{\Delta k})s+2u'_{\Delta k}Q\dot{u}s\}ds$ goes to zero after implementing integration by parts and then using the power series expansion of $e^{-\rho(\Delta)}$, $(1-\rho\Delta)$.

2.2.1 A Numerical Illustration

Now that we have shown all of the necessary variable transformations, we can examine the convergence of the transformed discrete-time system to the continuous-time system. As shown in figure 1 after decreasing Δ from 1.0 to 0.001 the transformed discrete-time system converges to the same solution as the continuous-time system.

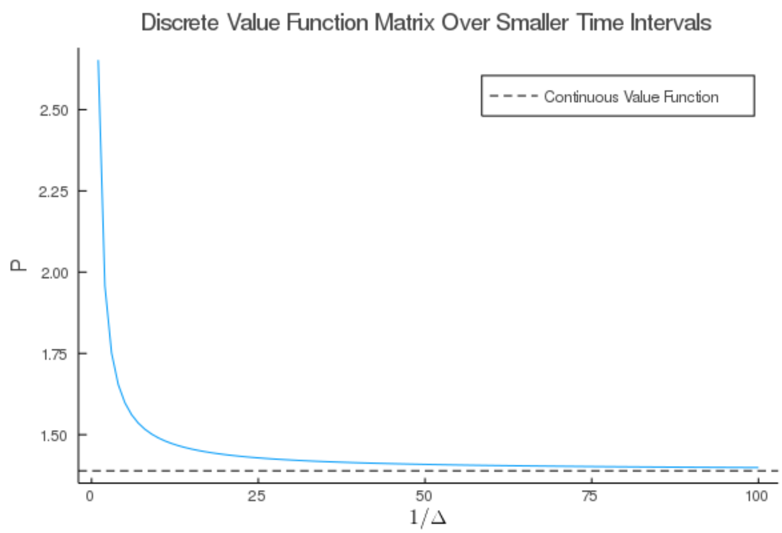


Figure 1

Figure 1 displays the unique tie between the discrete-time LQ solutions and the continuous-time version. Now we expand our analysis to a case with stochasticity.

3 A More Complex Model

In this section, we outline a more complicated system with stochasticity and interaction terms between x and u. This optimal linear regulator problem takes the following form,

$$-\mathbb{E}\int_{t=0}^{\infty} e^{-\rho t}\{x_t'Rx_t + u_t'Qu_t + 2x_t'Wu_t\}dt. \tag{20}$$

Where the state of the system, x_t, evolves according to,

$$dx_t = Ax_t dt + Bu_t dt + CdZ_t \tag{21}$$

here dZ_t is the increment of the Wiener process[4] and A is assumed to be symmetric. As before x_t is a $(N \times 1)$ vector of state variables and u_t is a $(M \times 1)$ vector of control variables.

The HJB for this problem can be found similarly to (11). For this system, the HJB is,

$$\rho V(x) = \max_u \; - x'Rx - u'Qu - 2x'Wu + \frac{1}{dt}\mathbb{E}\left(V_x(x)dx_t + \frac{1}{2}V_{xx}(x)(dx_t)^2 \right). \tag{22}$$

Note that unlike the HJB in (11), this HJB equation has an additional term which comes from applying Itô's lemma to the stochastic process for dx_t (Dixit, 1992). This additional term changes the proposed $V(x)$ (Hansen and Sargent, 2013). In this setting the value function takes the form,

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

where ξ does not depend on our state or control variables. Plugging the proposed value function into (22) yields,

$$\rho x'Px + \rho\xi = \max_u \{x'Rx + u'Qu + 2x'Wu + 2x'P(Ax + Bu) + P(CC')\}. \tag{23}$$

[4]The increment of the Wiener process can be approximated as $dZ_t = \varepsilon_t\sqrt{dt}$ where $\varepsilon_t \sim \mathrm{N}(0, 1)$. Thus, $\mathbb{E}[dZ_t] = 0$ and $\mathbb{E}[(dZ_t)^2] = dt$

This yields the following policy for u,

$$u = -(Q')^{-1}(W + PB)'x = -Fx. \tag{24}$$

Now, plugging this policy into (23) and rewriting the result in a general form produces,

$$\rho P = R + F'QF - 2WF + 2A'P - 2PBF \tag{25}$$

$$\rho\xi = PCC'. \tag{26}$$

Note that is equation is similar to the discrete stochastic case discussed in Hansen and Sargent (2013) in that the matrix C that multiplies the Wiener process dZ_t does not impact P, instead it affects ρ. Thus the matrix P is independent of the stochasticity in this problem. Steady-state solutions for this type of system can be found similarly to the system in section 2.1 using the following recursive scheme

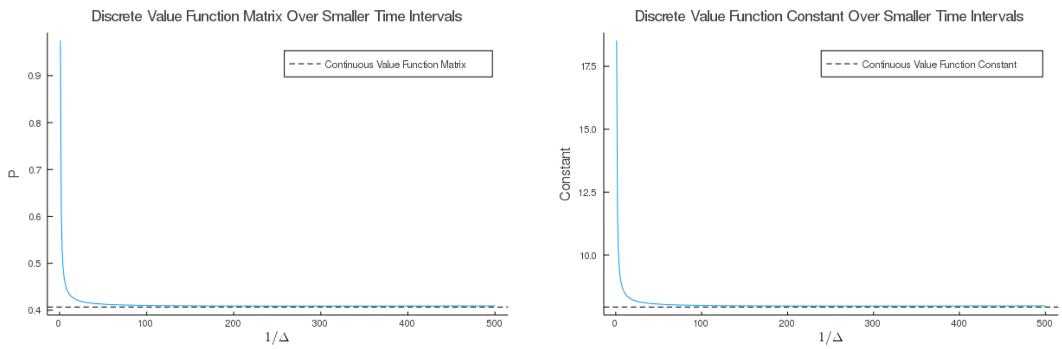
$$P_i = -(2\tilde{A}_i')^{-1}(\tilde{F}_i'Q^{-1}\tilde{F}_i + R - 2W\tilde{F}_i) \tag{27}$$

$$\xi_i = \rho^{-1}\text{trace}(P_{i-1}CC'), \tag{28}$$

where $\tilde{A}_i = (A - B\tilde{F}_i - .5\rho)$ and $\tilde{F}_i = (Q')^{-1}(W + P_{i-1}B)'$.

3.1 Convergence in the Complex Case

Before moving on, it is worth noting that under transformations similar to those in section 2.2 a discrete version of this system converges to the continuous model we described in the previous section. The necessary transformations are β becomes $\beta(\Delta) = e^{-\hat{\rho}}$, R is now $R(\Delta) = \frac{1}{\rho}(1 - e^{-\hat{\rho}})R$, $Q(\Delta) = \frac{1}{\rho}(1 - e^{-\hat{\rho}})Q$, $W(\Delta) = \frac{1}{\rho}(1 - e^{-\hat{\rho}})W$, $A(\Delta) = (I + A\Delta)$ and $B(\Delta) = B\Delta$, and $C(\Delta) = C\sqrt{\Delta}$ where $\hat{\rho} = \rho\Delta$. To test convergence for this model, we used the same univariate case as in section 2 with $W = 1.0$ and $C = 1.0$. The rate of convergence for the matrix P in the complex case is similar to the rate of convergence in the simple case considered earlier.



(a) (b)

Figure 2

In figure 2a the transformed discrete system's value function, or P matrix, converges to the continuous system's value function, and in figure 2b the value function's constant term ξ converges to the continuous-time system's constant. Figure 2 demonstrates that even in the more complicated model, the discrete-time system's solutions can limit to the continuous-time solutions.

4 Learning Dynamics

Now, we have enough information to capture an agent's behavior under bounded rationality, assuming that LQ.1-LQ.3 are satisfied. Modeling the agent's behavior can be done using a continuous-time analog to recursive least squares (Lewis et al., 2007). Recursive least squares is an essential discrete-time learning technique implemented in adaptive learning literature (Evans and Honkapohja, 2001). A method with continuous observations can be derived using generalized recursive methods and the link between the continuous and discrete-time Kalman filters.

4.1 Continuous-Time Recursive Least Squares

Recursive algorithms are used to estimate parameters and states in a wide variety of models. However, as provocatively stated in Ljung and Söderström (1983), "There is only one recursive identification method. It contains some design variables to be chosen by the user." While this statement is not true for all models, we can use the same general algorithm for a wide variety of linear regression and state-space models. This relationship between recursive algorithms has been often noted for the Kalman filter and LQ problems as in Ljungqvist and Sargent (2012); however we explored this relationship with two other standard recursive algorithms in economics—recursive least squares (RLS) and the Kalman filter.

Parallels between the Kalman filter and RLS are well understood in economics research and have been cited in Branch and Evans (2006) and Sargent (1999). First, we study the recursive least squares algorithm. Then we examine the Kalman filter and transform this recursive algorithm into the RLS algorithm to better understand their linkage before examining both these systems in continuous-time. The relationship between these discrete and continuous-time recursive algorithms has been noted in Ljung (1977) and Lewis et al. (2007). The recursive least squares approach begins with the following difference equation model,

$$y_t = \theta' x_t + e_t$$

where $e_t \sim \mathrm{N}(0, 1)$. Here we can estimate the model by choosing an estimate the minimizes the errors of the model. We select a least-squares method,

$$V_N(\theta) = \frac{1}{N} \sum_{t=1}^{N} \alpha_t [y_t - \theta' x_t]^2 \tag{29}$$

where N is the number observations in the data and α_t is a weighting vector that depends on time. The weighting vector α_t is indirectly related to the gain sequence in adaptive learning literature, for the derivation of RLS it is often set to 1. Implementing this least-squares method we can derive RLS,

$$\begin{aligned}\hat{\theta}_t &= \hat{\theta}_{t-1} + \frac{1}{t}\mathcal{R}_t^{-1}x_t\alpha_t[y_t - \hat{\theta}'_{t-1}x_t],\\ \mathcal{R}_t &= \mathcal{R}_{t-1} + \frac{1}{t}[\alpha_t x_t x_t' - \mathcal{R}_{t-1}]\end{aligned}$$

This recursive algorithm estimates coefficients based on observations and an estimate of the second moment $\mathcal{R}_t$. The avoid the matrix inversion in the system above we can instead use $\mathcal{P}_t = \bar{\mathcal{R}}_t^{-1}$.

$$\begin{aligned}\mathcal{P}_t &= [\mathcal{P}_{t-1}^{-1} + x_t\alpha_t x_t']^{-1}\\ &= \mathcal{P}_{t-1} - \frac{\mathcal{P}_{t-1}x_t x_t'\mathcal{P}_{t-1}}{1/\alpha_t + x_t'\mathcal{P}_{t-1}x_t}.\end{aligned}$$

Thus our system will become,

$$\hat{\theta}_t = \hat{\theta}_{t-1} + L_t[y_t - \hat{\theta}'_t x_t], \tag{30}$$

$$L_t = \frac{\mathcal{P}_{t-1}x_t}{1/\alpha_t + x_t'\mathcal{P}_{t-1}x_t}, \tag{31}$$

$$\mathcal{P}_t = \mathcal{P}_{t-1} - \frac{\mathcal{P}_{t-1}x_t x_t'\mathcal{P}_{t-1}}{1/\alpha_t + x_t'\mathcal{P}_{t-1}x_t}. \tag{32}$$

Now, we will look at the Kalman filter and how it relates to RLS. Suppose we have the following state-space model,

$$\text{Transition Equation: } z_{t+1} = F_t z_t + w_t, \tag{33}$$

$$\text{Measurement Equation: } y_t = H_t z_t + e_t \tag{34}$$

Where $\{w_t\} \sim N(0, R_t)$ and $\{e_t\} \sim N(0, r_t)$, r_t and R_t may be defined as constants. The Kalman Filter for this system can be described by the following equations,

$$z_{t+1} = F_t x_t + K_t[y_t - H_t x_t], \tag{35}$$

$$K_t = \frac{F_t \mathcal{P}_t H_t'}{r_t + H_t \mathcal{P}_t H_t'}, \tag{36}$$

$$\mathcal{P}_{t+1} = F_t \mathcal{P}_t F_t' + R_t - F_t \mathcal{P}_t H_t'[r_t + H_t \mathcal{P}_t H_t']^{-1} H_t \mathcal{P}_t F_t'. \tag{37}$$

Note the parallels between this and the system in (32). We can, in fact, imagine these as the same two algorithm. If we re-imagine the state-space model we used to derive the recursive least squares algorithm as,

$$\text{Transition Equation: } \theta_{t+1} = \theta_t + \nu_t \tag{38}$$

$$\text{Measurement Equation: } y_t = \theta_t' x_t + e_t \tag{39}$$

where $\nu_t \sim N(0, R_t)$ and $e_t \sim N(0, r_t)$, this comes from (29), the Kalman filter will become our RLS system when $R_t = 0$ and $r_t = 1/\alpha_t$. This particular RLS system will have a decreasing gain where the gain parameter $\gamma_t = 1/t$. The transition equation in (38) is now the transition equation for model parameters θ_t instead of data x_t, as shown in (38) the parameters in this setting are constant over time. The measurement equation in (39) is essentially the same as the measurement equation in (34); however, now there is uncertainty about the parameters θ_t as apposed to the data x_t.

Re-writing the Kalman filter for the system described in (38) and (39) yields,

$$\hat{\theta}_{t+1} = \hat{\theta}_t + K_t[y_t - x_t'\theta_t], \tag{40}$$

$$K_t = \frac{\mathcal{P}_t x_t}{1/\alpha_t + x_t' \mathcal{P}_t x_t}, \tag{41}$$

$$\mathcal{P}_{t+1} = \mathcal{P}_t - \mathcal{P}_t x_t [1/\alpha_t + x_t' \mathcal{P}_t x_t]^{-1} x_t' \mathcal{P}_t. \tag{42}$$

As we can see this is equivalent to the system in (30)-(32) with $K_t = L_t$, and some modified timing conventions. The Kalman filter and RLS are equivalent algorithms under these assumptions, which is useful since the continuous-time Kalman filter has a more intuitive and well-documented derivation than continuous-time RLS.

The continuous-time Kalman filter is used when measurements are continuous functions of time. In this section, we derive the continuous-time Kalman filter (Lewis et al., 2007). First, we modify (33)-(34) to depend on increments of time (Δ),

$$z_{k+1} = (I + F_k \Delta) z_k + w_k$$
$$y_k = H_k x_k + e_k$$

here the covariance matrix for $\{w_k\}$ is $R_k \Delta$ and the covariance matrix for $\{e_k\}$ is $r_k/(\Delta)$. First, we examine what happens to the Kalman gain in (36) as $\Delta \to 0$. Our Kalman gain becomes,

$$K_t = \frac{(I + F_t \Delta) \mathcal{P}_t H_t'}{(r_t/\Delta) + H_t P_t H_t'}$$

re-arranging this we will get

$$\frac{1}{\Delta} K_t = \frac{(I + F_t \Delta) \mathcal{P}_t H_t'}{r_t + H_t \mathcal{P}_t H_t' \Delta}$$

and taking the limit of this will yield,

$$\lim_{\Delta \to 0} \frac{1}{\Delta} K_t = \mathcal{P}_t H_t' r_t^{-1} \tag{43}$$

this is our continuous-time Kalman gain. Now, we can examine (37),

$$\mathcal{P}_{t+1} = (I + F_t\Delta)\mathcal{P}_t(I + F_t\Delta)' + R_t\Delta - (I + F_t\Delta)\mathcal{P}_t H_t'[(r_t/\Delta) + H_t\mathcal{P}_t H_t']^{-1} H_t\mathcal{P}_t(I + F_t\Delta)'.$$

Eliminating and terms and dividing by Δ yields,

$$\frac{1}{\Delta}\mathcal{P}_{t+1} = \frac{1}{\Delta}\mathcal{P}_t + F_t\mathcal{P}_t + \mathcal{P}_t F_t' + R_t - (I + F_t\Delta)\mathcal{P}_t H_t'[r_t + H_t\mathcal{P}_t H_t'\Delta]^{-1} H_t\mathcal{P}_t(I + F_t\Delta)'.$$

Then, taking the limit as $\Delta \to 0$,

$$\lim_{\Delta\to 0} \frac{1}{\Delta}\big(\mathcal{P}_{t+1} - \mathcal{P}_t\big)\dot{\mathcal{P}}_t = F_t\mathcal{P}_t + \mathcal{P}_t F_t' + R_t - \mathcal{P}_t H_t'[r_t]^{-1} H_t\mathcal{P}_t$$

this equation is our continuous-time covariance updating equation.

Last, we derive the estimate updating equation. In this setting (35) will become,

$$\hat{z}_{t+1} = (I + F_t\Delta)\hat{z}_t + K_t[y_t - H_t\hat{z}_t]$$

diving this by Δ will give us,

$$\frac{1}{\Delta}(\hat{z}_{t+1} - \hat{z}_t) = F_t\hat{z}_t + \frac{K_t}{\Delta}[y_t - H_t\hat{z}_t].$$

Now, we can take the limit as $\Delta \to 0$ and use equation (43),

$$\dot{\hat{z}}_t = F_t\hat{z}_t + \mathcal{P}_t H_t' r_t^{-1}[y_t - H_t\hat{z}_t]$$

this will be our systems estimate updating equation.

Thus our continuous-time Kalman filter for this system can be described by the following equations. Our system can be rewritten as,

$$\dot{z} = Fz + w \tag{44}$$

$$y = Hz + v \tag{45}$$

Here w and v are error terms and $w \sim N(0, R)$ and $v \sim N(0, r)$. Our Kalman filter for this system is,

$$\dot{\mathcal{P}} = F\mathcal{P} + \mathcal{P}F' + R - \mathcal{P}H'r^{-1}H\mathcal{P} \tag{46}$$

$$K = \mathcal{P}H'r^{-1} \tag{47}$$

$$\dot{\hat{z}} = F\hat{z} + K[y - H\hat{z}] \tag{48}$$

now that we have established how to derive the continuous-time Kalman filter we can use this to get a continuous version of RLS.

We can rewrite the system in (38)-(39) as,

$$\dot{\theta}_t = \nu_t \tag{49}$$

$$y_t = \theta_t' x_t + e_t \tag{50}$$

Now, $\nu_t \sim N(0, R)$ where $R = 0$ and variance for e_t is $r_t = 1/\alpha_t$, our RLS system will be

$$\dot{\mathcal{P}} = -\alpha_t \mathcal{P}x'x\mathcal{P} \tag{51}$$

$$K = \alpha_t \mathcal{P}x' \tag{52}$$

$$\dot{\hat{\theta}}_t = K[y_t - \hat{\theta}_t' x_t]. \tag{53}$$

We can also derive this more rigorously starting from a discretized version of the model. The discretized version of our model with an undetermined time step Δ is,

$$
\begin{aligned}
\theta_{k+1} &= \theta_k \\
y_k &= \theta_k' x_k + e_k
\end{aligned}
$$

Where, the covariance matrix for $e_k \sim N(0, \frac{1}{\alpha_k \Delta})$ as in Lewis et al. (2007). First, we can examine the gain term in (31). Writing (31) in this setting we'll have,

$$
\begin{aligned}
L_t &= \mathcal{P}_{t-1} x_t [(1/\alpha_t \Delta) + x_t \mathcal{P}_{t-1} x_t']^{-1} \\
&= \mathcal{P}_{t-1} x_t \Delta [1/\alpha_t + x_t \mathcal{P}_{t-1} x_t' \Delta]^{-1}.
\end{aligned}
$$

Dividing through by Δ and then taking the limit as $\Delta \to 0$ we get,

$$
K = \lim_{\Delta \to 0} \frac{1}{\Delta} L_t = \alpha_t \mathcal{P}_{t-1} x_t \tag{54}
$$

Next, if we look at (32) we can rewrite this equation as,

$$
\begin{aligned}
\mathcal{P}_t - \mathcal{P}_{t-1} &= -\mathcal{P}_{t-1} x_t x_t' \mathcal{P}_{t-1} [(1/\alpha_t \Delta) + x_t \mathcal{P}_{t-1} x_t']^{-1} \\
&= -\mathcal{P}_{t-1} x_t x_t' \mathcal{P}_{t-1} \Delta [1/\alpha_t + x_t \mathcal{P}_{t-1} x_t' \Delta]^{-1}.
\end{aligned}
$$

Dividing through by Δ and taking the limit as $\Delta \to 0$,

$$
\dot{\mathcal{P}}_t = -\mathcal{P}_{t-1} x_t x_t' \mathcal{P}_{t-1} = -K x_t' \mathcal{P}_{t-1}. \tag{55}
$$

Last, we can derive the continuous-time estimate updating equation (30). Rewriting this equation and diving through by Δ yields,

$$\frac{1}{\Delta}\left(\hat{\theta}_t - \hat{\theta}_{t-1}\right) = \frac{1}{\Delta} L_t[y_t - \hat{\theta}'_{t-1} x_t].$$

Limiting this as $\Delta \to 0$ we get,

$$\dot{\hat{\theta}}_t = K[y_t - \hat{\theta}'_{t-1} x_t]. \tag{56}$$

These equations we have just derived are the same as the Kalman filter equations in (51)-(53). Again we have shown that the RLS and Kalman Filter equations are the same under particular assumptions. The equivalence of these two is intuitive even at glance, if our models are both set up in a set-space setting. The Kalman filter and RLS are similar algorithms; however, the Kalman filter estimates the states of our model and RLS estimates our regression coefficients. From a Bayesian perspective the Kalman filter use's Bayes' Law to update the estimate of our state as we observe more data about the system. As shown under certain conditions RLS is equivalent to a Kalman filter applied to a system with random walk coefficients (Sargent, 1999).

4.2 Adaptive Learning Rules in Continuous-Time

Using the approach outlined in the previous section, we can now write an adaptive learning algorithm for the system outlined in (20) and (21). In this algorithm agents will gain information about a data process for x_t and use this information to update their predictions of parameters and shadow-prices in turn their decisions will impact the states that they observe.

The agent will update their estimates of the systems transition matrix, A, and the shadow price parameters which we will denote as H using the continuous analog of recursive least squares. Estimated values of A and H will then impact the agent's policy decision and the shadow prices they observe next period. Recall from previous sections that the process for the evolution of our states follows,

$$dx_t = Ax_t dt + Bu_t dt + CdZ_t$$

and our policy function is

$$u = -(Q')^{-1}(W + PB)'x = -Fx. \tag{57}$$

Below the learning algorithm is outlined. Here $\mathcal{P}_t$ is the second moment matrix for x_t using all data between the initial observation and the current observation and γ_t is the gain sequence that measures the response of estimates to forecast errors. For simplicity, we assume that the gain is constant—$\nu = 0$ and $\kappa = 0.01$. Additionally, $F^{SP}(H_t, B)$ is the policy under shadow price learning and $T^{SP}(H_t, A_t, B)$ is the T-map—a link between agent's perception and the actual system, we will describe both functions as well as the link between H and P in the following section.

$$\begin{aligned}
dx_t &= Ax_t dt + Bu_t dt + CdZ_t \\
d\mathcal{P}_t &= -\gamma_t \mathcal{P}_t x_t x_t' \mathcal{P}_t dt \\
dH_t' &= \gamma_t \mathcal{P}_t x_t (\lambda_t - H_t x_t)' dt \\
dA_t' &= \gamma_t \mathcal{P}_t x_t (dx_t - Bu_t dt - A_t x_t dt)' \\
u_t &= -F^{SP}(H_t, B)x_t = -\frac{1}{2}(Q')^{-1}(2W - H'B)x_t \\
\lambda_t &= T^{SP}(H_t, A_t, B)x_t \\
\gamma_t &= \kappa(t + N)^{-\nu}.
\end{aligned} \tag{58}$$

We have formatted the learning algorithm in terms of changes in levels as opposed to time derivatives to more closely fit the formatting of stochastic processes in macroeconomic literature. Analytically, this will give us the same results as the time derivative format.

4.2.1 Continuous-time Policies and the T-map

The optimal linear regulator solution methods outlined in the previous sections were recursive, meaning that given an approximation to the solution $V_k(x)$ a new approximation $V_{k+1}(x)$ can be obtained. Note that here k is not a measure of time. This approach conveniently lends itself to learning algorithms as the first approximation $V_k(x)$ can be viewed as the perceived value function, using $V_k(x)$ one can then compute the induced value function $V_{k+1}(x)$. Utilizing the notation from the previous section and allowing $V^P(x)$ to represent the induced value function and P to the perceived value function we can write $V^P(x)$ as

$$\rho V^P(x) = \max_u \{-x'Rx - u'Qu - 2x'Wu - 2x'P(Ax + Bu) - P(CC')\} \tag{59}$$

Here the agents choose u to solve this problem. Equation (59) can be used to describe the agent's control decisions and the induced value function. The unique optimal control decision for perceptions P is given by,

$$u = -F(P)x = -(Q')^{-1}(W + PB)'x.$$

When examining induced value function for perceptions P, we first examine the deterministic case where $C = 0$. In this case the induced value function is given by

$V^P(x) = -x'T(P)x$ where

$$T(P) = (2\tilde{A}')^{-1}(F'Q^{-1}F + R - 2WF) \tag{60}$$

here $\tilde{A} = A - \frac{1}{2}I\rho - BF$. The right-hand-side of $T(P)$ is similar to the Riccati equation (27). Knowing this, we can conclude that the fixed point of this T-map identifies the solution to the agent's optimal control problem. In the stochastic case where $C \neq 0$ our T-map is given by,

$$T^{\varepsilon}(\tilde{P}) = \tilde{P} - \rho^{-1}\text{trace}(\tilde{P}CC')$$

where $T(\tilde{P}) = \tilde{P}$. Optimally decision making in this setting is determined by the fixed pint of $T^{\varepsilon}(P^*_{\varepsilon})$, P^*_{ε}. The fixed point of the stochastic system is directly related to the solution for the deterministic case, P^*, by the following equation

$$P^*_{\varepsilon} = P^* - \rho^{-1}\text{trace}(P^*CC').$$

Thus, the solution to the deterministic problem yields the solution to the stochastic problem.

4.3 Shadow Price Learning

The learning dynamics outlined thus far have made strong assumptions about an agent's knowledge of the value function. In the problem outlined in (59), an agent understands that the value function is quadratic in x, knows how to solve for the matrix P by iterating on the Riccati equation, and knows parameters A and B. In the following section, we modify these assumptions. As opposed to assuming the agent knows A and B, we assume that the agent does know B, indicating they understand how their control decisions impact the state. However, the agent is not assumed to know the parameters of the state-contingent transition dynamics. Meaning they must estimate A. Additionally, the agent in the following problem is not assumed to know how to solve the programming problem. Instead, they use a simple forecasting model to estimate the value of the state tomorrow—the shadow price of the state. The agent then uses this estimate along with an estimate of the transition equation to determine the best control response for today.

We now outline a learning framework in which the agent forms expectations of future shadow prices. The boundedly optimal behavior modeled in this section is shadowing price learning or SP-learning (Evans and McGough, 2018). Under SP-learning, the agent believes that the shadow price, λ, is linear in x. Thus they can forecast the shadow price as,

$$\lambda_t = Hx_t + \mu_t \tag{61}$$

where μ_t is some error term. Using this perceived law of motion (PLM), we can create a T-map for the agent's perceptions using equation (11). we first estimate that,

$$\mathbb{E}[V_x(x)] = \lambda^e = Hx$$

where λ^e is the updated estimate of λ. Plugging this into (11) we get,

$$\rho V(x) = \max_u \{-x'Rx - u'Qu - 2x'Wu + (Hx)'(Ax + Bu) + \frac{1}{2}(H'CC')\}.$$

Our first order condition with respect to u will yield,

$$u = -\frac{1}{2}(Q^{-1})'(2W - H'B)x = -F^{SP}(H, B)x \tag{62}$$

our policy function. Then to get the mapping from the PLM to the actual law of motion (ALM) we use the envelope condition,

$$\rho V_x(x) = \rho\lambda^e = -2x'R - 2u'W + 2x'A'H + u'B'H. \tag{63}$$

we can rewrite (63) as,

$$\lambda^e = \rho^{-1}\{-2x'R - 2u'W + 2x'A'H + u'B'H\}$$

or

$$\begin{aligned} \lambda^e &= T^{SP}(H, A, B)x \\ &= \rho^{-1}\big(-2R + 2H'A - (H'B - 2W)F^{SP}(H, B)\big)x. \end{aligned} \tag{64}$$

This is the T-map that will be used to model the agent's boundedly rational behavior. The fixed points of this mapping correspond to equilibrium values of P since agents' perceptions are in line with the correct data-generating process expectations are formed optimally.

4.3.1 Stability of shadow-price learning dynamics

The stability of the T-map is essential to learning dynamics.

Conjecture 1. *Assuming that LQ.1-LQ.3 hold, there exists an $n \times n$ solution P^* to the Riccati equation given any symmetric positive definite initial matrix P_0 (Evans and McGough, 2018). Therefore $T^m(P_0) \to P^*$ as $m \to \infty$ and*

1. *$T(P^*) = P^*$—the solution P^* is a fixed point of the T-map.*

2. *$DT_v(\mathrm{vec}(P^*))$ is stable—has eigenvalues less than one.*

3. *P^* is the unique fixed point of T among the class of $n \times n$, symmetric positive semi-definite matrices.*

Thus, if the Riccati equation has asymptotically stable solutions, the T-map for the system is stable. Conjecture 1 is proved to be true in the discrete-time setting in Evans and McGough (2018), based on numerical and analytical results it conjectured to hold true in the continuous-time setting as well.

Next, we will examine the solutions and stability of the learning system using $A = 0.0$, $R = Q = B = 1.0$, $W = C = 0$, and $\rho = 0.05$. With these values, our T-map (64) can be rewritten as a function of H. This function $T(H)$ has two fixed points. One at $\tilde{H} \approx 2.880$ and a second solution at $H^* \approx -2.778$. This second solution is consistent with the solutions for P from both the continuous iterative scheme and the icare function, since $H = -2P$. Directly comparing the solution for P from the iterative schemes and $-\frac{1}{2}H^*$ there is a difference of 2.220×10^{-16}.

The solution H^* is stable, based on stability conditions for the ARE and the T-map. For the continuous-time ARE to be stable, $A + BF^{SP}(H^*, B)$ must have eigenvalues with real parts less than one, and our T-map must satisfy the condition that $DT^{SP}(H^*, A_t, B)$ has eigenvalues with real parts less than one. H^* meets these stability conditions as,

$$A + BF^{SP}(H^*, B) = -0.975, \quad DT^{SP}(H^*, A, B) = -39.012.$$

However, the unstable solution $\tilde{H}$ does not meet these criteria as

$$A + BF^{SP}(\tilde{H}, B) = 1.025, \quad DT^{SP}(\tilde{H}, A, B) = 41.012.$$

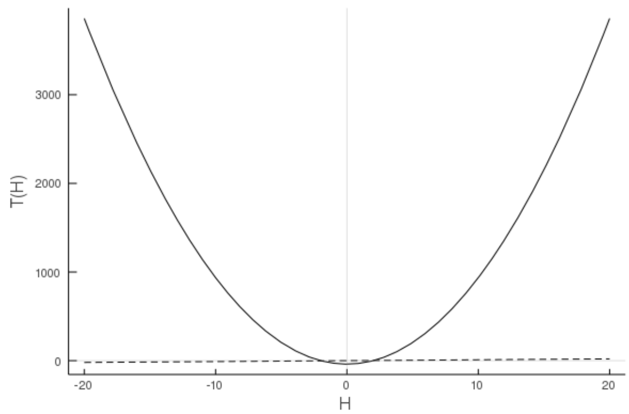


Figure 3: T-map

Now that we have examined stability conditions and derived a continuous-time version of RLS, we can examine the convergence of the learning algorithm outlined in (58). In the next section, we apply adaptive learning techniques to a univariate test case.

4.4 Continuous-Time Learning Results

In this system of equations, the agent's behavior is modeled by a system of equations that measures the change in each parameter that occurs over time. These changes are then used to update parameter estimates. As shown below in figure 4, when using an approximation of the length of the time increment ($dt \approx 0.001$) and a constant gain ($\gamma = 0.01$) the method outlined in (58) will be yield convergence results.

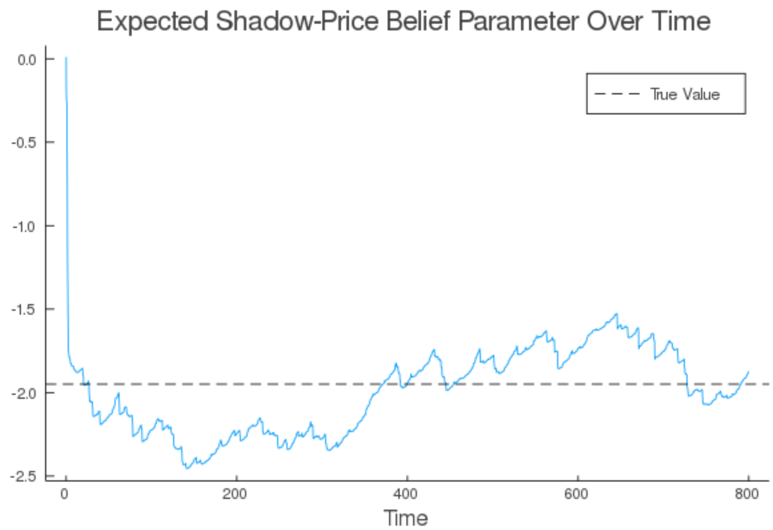


Figure 4

We can compare this to a discrete-time system where an agent's bounded rational behavior can be modeled by (Evans and McGough, 2018),

$$
\begin{aligned}
x_t &= Ax_{t-1} + Bu_{t-1}dt + C\varepsilon_t \\
\mathcal{R}_t &= \mathcal{R}_{t-1} + \gamma_t(x_t x_t' - \mathcal{R}_{t-1}) \\
H_t' &= H_{t-1} + \gamma_t \mathcal{R}_{t-1}^{-1} x_{t-1}(\lambda_{t-1} - H_{t-1}x_{t-1})' \\
A_t' &= A_{t-1} + \gamma_t \mathcal{R}_{t-1}^{-1} x_{t-1}(x_t - Bu_{t-1} - A_{t-1}x_{t-1})' \\
u_t &= -F^{SPD}(H_t, A_t, B)x_t \\
&= (2Q - \beta B'HB)^{-1}(\beta B'HA_t - 2W')x_t \\
\lambda_t &= T^{SPD}(H_t, A_t, B)x_t \\
&= \Big(-2R - 2WF^{SPD}(H_t, A_t, B) + \beta A_t'H\big(A_t + BF^{SPD}(H_t, A_t, B)\big)\Big)x_t \\
\gamma_t &= \kappa(t+N)^{-\nu}.
\end{aligned}
\tag{65}
$$

Recall, $\mathcal{R}_t$ is the inverse of $\mathcal{P}_t$. Using the same parameter values as in the previous example this system has comparable convergence results,

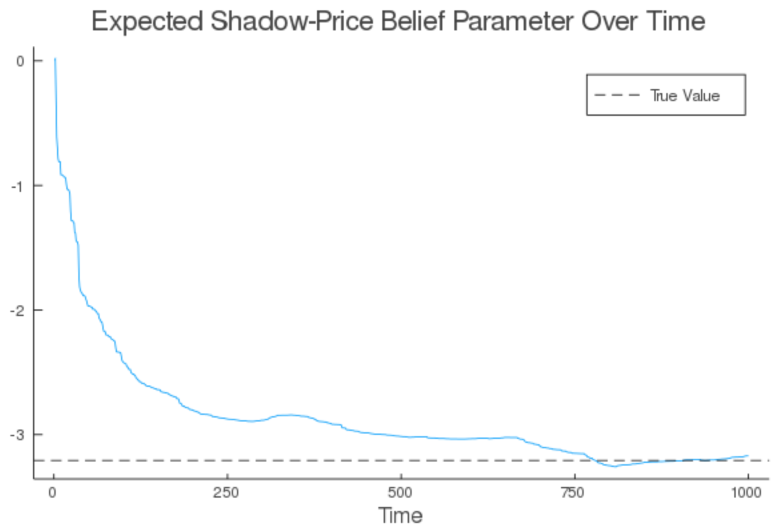


Figure 5

it is important to note that convergence occurs in these models after a different number of iterations. Since $dt \approx 0.001$, when the length of time is ten in the continuous-time case this corresponds to $10,000$ learning iterations. The rates of convergence will differ between these two cases as the discrete version will take in a single observation each period and weight the agent's response to the observation by the gain parameter $\gamma_t = 0.01$. In the continuous-time version the agent is observing $1,000$ data points in the time a discrete learner would observe a single value, thus the agent will weight their response to the forecast error by the gain parameter γ and then update their estimate after each observation.

4.5 Convergence in the Context of Learning

In section 3, we showed that under certain transformations our discrete-time system's value function matrix P can converge to the continuous-time solution. Similarly, the discrete learning rule outlined in equation (65) with $\gamma_t = (0.01)\Delta$ will converge to the continuous-time expected shadow price parameter when Δ is sufficiently small.

Figure 6 shows how the discrete learning rule responds under the transformations in section 3 with select values of Δ.[5]

Figure 6

As shown in figure 6 the modified discrete learning rule gradually gets closer to the continuous-time rational expectations solutions as Δ gets increasingly small.

5 A Robinson Crusoe Economy

Now that we have developed the modeling framework for continuous-time LQ problems and examined learning rules in this setting, we can examine a more involved model.

We begin with a simple Robinson Crusoe economy as in, Evans and McGough (2018). The representative agent in this model maximizes the following.

$$\max -\mathbb{E} \int_{t=0}^{\infty} e^{-\rho t} \left((c_t - b_t)^2 + \phi l_t^2 \right) \tag{66}$$

[5]The learning iterations in figure 6 have been re-scaled for easier representation. Each iteration is equivalent to a discrete time period $t = 1, 2, \ldots 10,000$ that contains Δ^{-1} observations. Meaning that for $\Delta = 1/4$ this graph is displaying the results from $40,000$ iterations

subject to the following

$$
\begin{aligned}
y_t &= A_1 s_t \\
ds_t &= (y_t - c_t - s_t)dt + dZ_t \\
s_t &= l_t \\
b_t &= b^*
\end{aligned}
\tag{67}
$$

where dZ_t is the increment of the Wiener process.

The model we have outlined in (66) and (67) is a simplified version of the Robinson Crusoe (RC) model used in Evans and McGough (2018). In this model, the agent has only one consumable good, fruit, and only one means of production, growing trees from seeds of the fruit. Thus, income y_t can be thought of as fruit and consumption c_t as the consumption of that fruit along with its seeds. The production of the fruit comes from planting seeds, s_t. In the system above, the change in the number of seeds over time depends on growing conditions—represented by the increment of the Wiener process dZ_t—and leftovers from consumption. In this model, work is burdensome and causes disutility for the worker ($\phi > 0$). Lastly, b_t is a bliss point represented by the constant b^*.

We have simplified this model, to maintain similarities between a continuous and discrete case. For instance, we do not have a possible time lag in production—in this model, young trees and old trees produce the same amount. Additionally, the bliss point is non-stochastic, and there are no productivity shocks. Instead, production only depends on the availability of seeds.

we need to transform this system to get into the same format as (20) and (21) to find a steady-state using the methods described in section 3. Setting the state vector as $x_t = (1, s_t)'$ and the control as $u_t = c_t$ the states will evolve according to,

$$dx_t = Ax_t dt + Bu_t dt + CdZ_t$$

where

$$A = \begin{bmatrix} 0 & 0 \\ 0 & A_1 - 1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \quad C = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Now the R, Q, and W matrices from equation (20) must be found. For this system, these are defined as follows,

$$R = \begin{bmatrix} b^{*2} & 0 \\ 0 & \phi \end{bmatrix}, \quad Q = 1.0, \quad W = \begin{bmatrix} -b^* \\ 0 \end{bmatrix}.$$

Using these matrices and parameter values we can now calculate the rational expectations equilibrium for this system.

5.1 Learning in the Continuous RC Model

In this setting, it is likely that our agent does not know the parameters of the production function, or the value of an additional tree tomorrow. However, the agent can use the system outlined in (58) to forecast these unknown values. As the agent gains more information they can update their parameter estimates using (58); the matrices B, C, R, Q, and W; and initial values for A_t, H_t, $\mathcal{R}_t$, and λ_t.

Under the learning rules described in (58), the agent learns parameters for the matrix H, and the matrix A (in this case both are a 2×2 matrix). To generate data for this model, an approximation for dt was necessary. For the following results, we used $dt \approx \Delta = 1/25$. Additionally, we used a constant gain term where $\kappa = 0.01$, and $\nu = 0$.

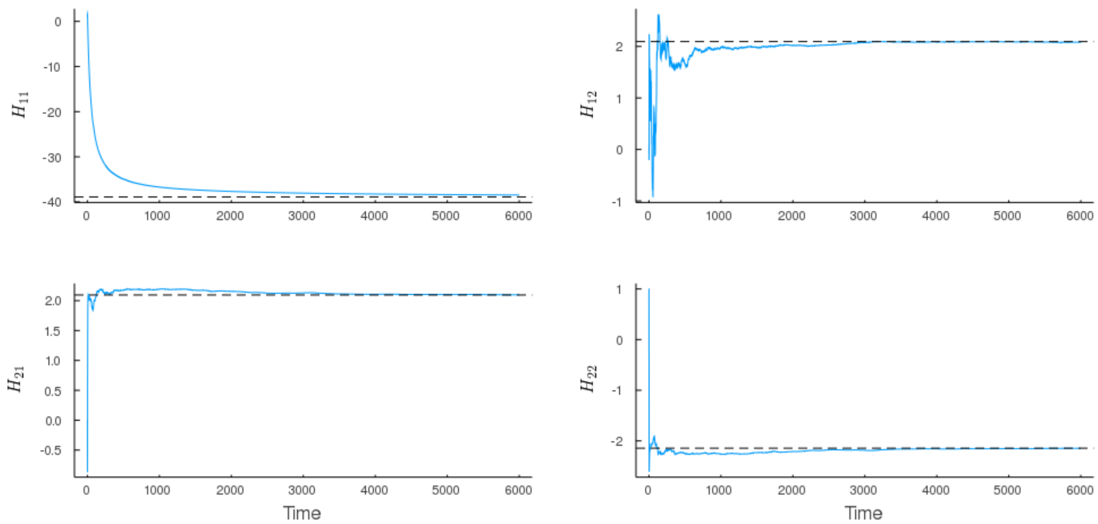


Figure 7: Expected Shadow-Price Parameters

As shown in figure 7 an agent with boundedly rational behavior modeled by (58) will be able to generate close estimate of the steady state shadow price parameters. In figure 7 we have plotted 150,000 learning iterations, which is equivalent to approximately 6,000 discrete time-periods. Fewer time periods are necessary for the convergence of H_{12}, H_{21} and H_{22}.

5.2 Learning in the Discrete RC Model

A discrete version of this model with, as outlined in Evans and McGough (2018), converges similarly with the same constant gain parameter. Below we have plotted 20,000 discrete periods to make it easier to compare the convergence of this system to the continuous system in section 5.1.

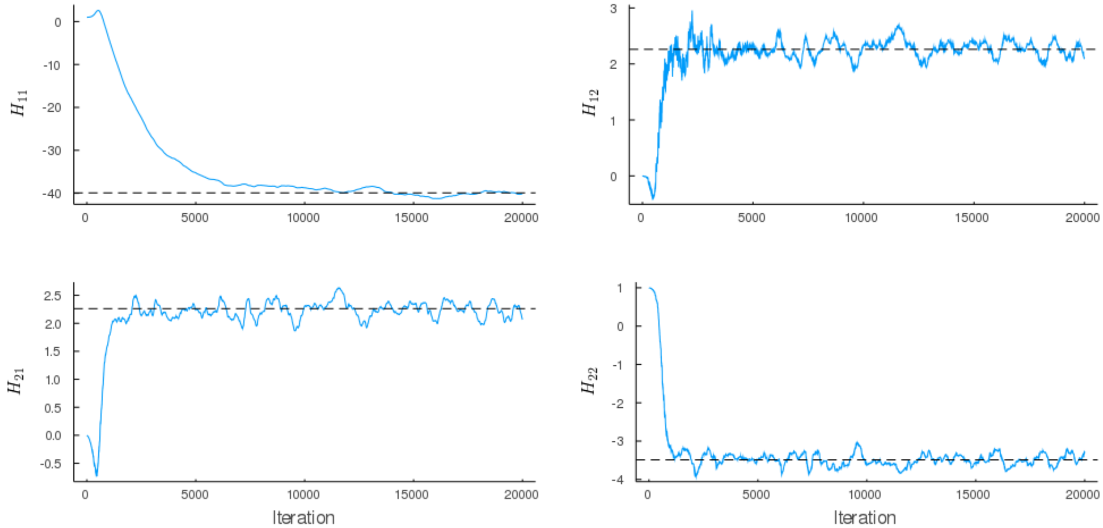


Figure 8: Expected Shadow-Price Parameters

6 Conclusion

As continuous-time macroeconomic literature expands, it is necessary to modify and re-evaluate discrete modeling techniques in this framework. Adaptive learning mechanisms are particularly essential to modify as they relax the strong assumption of rational expectations—the belief that agents forecast optimally. The shadow-price learning technique outlined in the previous sections goes beyond easing rational expectations, as it also examines the optimality of an agent's decisions as they optimize according to their forecasts. Since agents in this setting use available information to forecast their shadow-prices and then make control decisions based on their forecasts (Evans and McGough, 2018).

It was beneficial to develop a continuous-time linear-quadratic framework for macroeconomic models to efficiently implement shadow price learning in a continuous-time environment. Other disciplines, such as engineering, frequently use continuous-time linear quadratic methods (Vrabie et al., 2009; Lewis, 1986). However, very few examples of economic models in this framework exist (Hansen and Sargent, 1991). After building this general framework, we were able to examine convergence results and equilibrium stability in this class of models.

With this continuous-time LQ framework, we were able to implement a continuous analog to recursive least squares and analyze a continuous-time T-map. This system yielded results that suggest an agent can learn to optimize decisions in both simple univariate cases and with more sophisticated models. Obvious extensions to this work include expanding the analysis to more complex economies and examining continuous-time learning methods outside of the linear-quadratic framework.

A Algebraic Riccati Equation Solutions

To verify the convergence of (6), (15), and (16) a simple univariate system was tested. In this test case, $A = 0$, $B = 1$, $R = 2$, $Q = 1$, $\beta = .95$, and $\rho = -\ln\beta$ (for consistency between the continuous and discrete discount rates). Below, is a table comparing the results of the iterative methods to output from MATLAB's built-in functions for solving AREs, icare for continuous systems and idare for discrete ones.

Solution Comparisons

Iterative Scheme	**Iterative Solution**	**MATLAB Solution**	**Difference**
Equation (6)	2.0000	2.0004	4.1670e-04
Equation (15)	1.3887	1.3894	6.3507e-04
Equation (16)	1.3887	1.3894	6.3507e-04

Table 1: Iterative Scheme Results

As table 1 shows the results from the iterative schemes are fairly close to the standard MATLAB solutions.[6] Additionally, (15) and (16) output identical solutions in our simple case and should be able to be used interchangeably.

B OLRP with Fewer Symmetry Assumptions

Here we outline a continuous-time optimal linear regulator problem without symmetry assumptions. In this section we revisit the continuous-time problem in section 3 and relax the assumption that the matrix A is symmetric. In this setting the an agent

[6]The iterative solutions were found using julia not MATLAB. This may contribute to the difference between the iterative solutions and MATLAB functions as julia and MATLAB round differently.

faces the following optimization problem,

$$V(x_0) = \max -\mathbb{E}\int_{t=0}^{\infty} e^{-\rho t}\{x_t' R x_t + u_t' Q u_t + 2x_t' W u_t\}dt. \tag{68}$$

Where the state of the system, x_t, evolves according to,

$$dx_t = Ax_t dt + Bu_t dt + CdW_t \tag{69}$$

here dW_t is the increment of the Wiener process. The HJB for this problem can be found similarly to (11). For this system, the HJB will be,

$$\rho V(x) = \max_u \; -x'Rx - u'Qu - 2x'Wu + \mathbb{E}\bigg(V_x(x)\dot{x} + \frac{1}{2}V_{xx}(x)\dot{x}^2\bigg). \tag{70}$$

In this setting the value function takes the form (Hansen and Sargent, 2013),

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

where ξ does not depend on the state or control variables. Plugging the proposed value function into (70) yields,

$$\rho x'Px + \rho\xi = x'Rx + u'Qu + 2x'Wu + x'P(Ax + Bu)(Ax + Bu)'Px + P(CC'). \tag{71}$$

This yields the following policy for u,

$$u = -(Q')^{-1}(W + PB)'x = -Fx. \tag{72}$$

Now, plugging this policy into (71) and rewriting the result in a general form produces,

$$\rho P = R + F'QF - 2WF + PA + A'P - PBF - F'B'P \tag{73}$$

$$\rho\xi = PCC'. \tag{74}$$

This is similar to the discrete stochastic case discussed in Hansen and Sargent (2013). The steady-state solution for this system can be found similarly to the system in section 2.1 using the following iterative scheme

$$P_i = -(I_n \otimes \tilde{A}' + \tilde{A}' \otimes I_n)^{-1}\text{vec}(\tilde{F}_i'Q^{-1}\tilde{F}_i + R - 2W\tilde{F}_i)$$

$$\xi_i = \rho^{-1}trace(P_{i-1}CC'),$$

where $\tilde{A}_i = (A - B\tilde{F}_i - .5\rho)$ and $\tilde{F}_i = (Q')^{-1}(W + P_{i-1}B)'$.

References

Achdou, Y., Han, J., Lasry, J., Lions, P., and Moll, B. (2020). Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach. *Review of Economic Studies*.

Ahn, S., Kaplan, G., Moll, B., Winberry, T., and Wolf, C. (2018). When Inequality Matters for Macro and Macro Matters for Inequality. *NBER Macroeconomics Annual*, 32(1).

Aït-Sahalia, Y. (2010). Estimating Continuous-Time models with discretely sampled data. *Advances in Economics and Econometrics, Theory and Applications, Ninth World Congress*.

Amman, H. M. and Kendrick, D. A. (1999). Should Macroeconomic Policy Makers Consider Parameter Covariances? *Computational Economics*, 14(3):263–267.

Anderson, B. D. and Moore, J. B. (2007). *Optimal Control Linear Quadratic Methods*. Dover.

Benigno, P. and Woodford, M. (2004). Optimal Monetary and Fiscal Policy: A Linear-Quadratic Approach. In *NBER Macroeconomics Annual 2003, Volume 18*, NBER Chapters, pages 271–364. National Bureau of Economic Research, Inc.

Benigno, P. and Woodford, M. (2006). Optimal Taxation in an RBC Model: A Linear-Quadratic Approach. *Journal of Economic Dynamics and Control*, 30(9-10):1445–1489.

Benigno, P. and Woodford, M. (2012). Linear-quadratic Approximation of Optimal Policy Problems. *Journal of Economic Theory*, 147(1):1–42.

Bergstrom, A. (1984). Continuous Time Stochastic Models and Issues of Aggregation over Time. In Griliches, Z. and Intriligator, M., editors, *Handbook of Economics*, volume 2, chapter 2, pages 1145–1212. Amsterdam: Elsveier.

Bergstrom, A. (1993). Survey of Continuous Time Econometrics. In Barnett, William A. and Giancarlo, Gandolfo and Hillinger, Clauda, editor, *Dynamic Disequilibrium Modeling: Theory and Applications*, chapter 1, pages 3–25. Cambridge University Press.

Black, F. and Scholes, M. (1973). Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, 81(3):637–654.

Branch, W. A. and Evans, G. W. (2006). A simple recursive forecasting model.(report). *Economics Letters*, 91(2):158.

Bray, M. (1982). Learning, Estimation, and the Stability of Rational Expectations. *Journal of Economic Theory*, 26:318–339.

Brock, W. and Mirman, L. (1972). Optimal Economic Growth and Uncertainty: The Discounted Case. *Journal of Economic Theory*, 4(3):479–513.

Brunnermeier, M. K. and Sannikov, Y. (2014). A Macroeconomic Model with a Financial Sector. *American Economic Review*, 104(2):379–421.

Dixit, A. (1992). The Art of Smooth Pasting. STICERD - Theoretical Economics Paper Series, Suntory and Toyota International Centres for Economics and Related Disciplines, LSE.

Evans, G. and McGough, B. (2018). Learning to Optimize. *Draft*.

Evans, G. W. and Honkapohja, S. (2001). *Learning and Expectations in Macroeconomics*. Princeton University Press.

Gabaix, X., Lasry, Jand Lions, P., Moll, B., and Qu, Z. (2016). The Dynamics of Inequality. *Econometrica*, 84(6):2071–2111.

Hansen, L. and Sargent, T. (1991). *Rational Expectations Econometrics*. Underground classics in economics. Westview Press. Includes bibliographical references (pages 283-293).

Hansen, L. and Sargent, T. (2013). *Recursive Models of Dynamic Linear Economies*. Princeton University Press.

Hansen, L. and Sargent, T. (2019a). Macroeconomic Uncertainty Prices when Beliefs are Tenuous. *NBER working paper*.

Hansen, L. and Sargent, T. (2019b). Structured Uncertainty and Model Misspecification. *University of Chicago, Becker Friedman Institute for Economics Working Paper No. 2018-77*.

Hommes, C. (2011). The Heterogeneous Expectations Hypothesis: Some Evidence from the Lab. *Journal of Economic Dynamics and Control*, 35(1):1–24.

Kaplan, G., Moll, B., and Violante, G. (2018). Monetary Policy According to HANK. *American Economic Review*, 108(3):697–743.

Kellerhals, P. B. (2001). *Financial Pricing Models in Continuous Time and Kalman Filtering*, volume 506. Springer.

Kendrick, D. (2005). Stochastic Control for Economic Models: Past, Present and the Paths Ahead. *Journal of Economic Dynamics and Control*, 29(1-2):3–30.

Kleinman, D. L. (1968). On an Iterative Technique for Riccati Equation Computations. *IEE Transactions on Automatic Control*, 13(1):114–115.

Lewis, F., Xie, L., and Popa, D. (2007). *Optimal and Robust Estimation: With an Introduction to Stochastic Control Theory*. CRC Press, Second edition.

Lewis, F. L. (1986). *Optimal Control*. Wiley-Interscience.

Ljung, L. (1977). Analysis of recursive stochastic algorithms. *IEEE transactions on automatic control*, 22(4):551–575.

Ljung, L. and Söderström, T. (1983). *Theory and Practice of Recursive Identification*. MIT Press.

Ljungqvist, L. and Sargent, T. J. (2012). *Recursive Macroeconomic Theory*. Mit Press.

Marimon, R. and Sunder, S. (1993). Indeterminacy of Equilibria in a Hyperinflationary World: Experimental Evidence. *Econometrica*, 61(5):1073–1107.

Marimon, R. and Sunder, S. (1994). Expectations and Learning under Alternative Monetary Regimes: An Experimental Approach. *Economic Theory*, 4(1):131–162.

Merton, R. (1969). Lifetime Portfolio Selection Under Uncertainty - Continuous-Time Case. *Review of Economics and Statistics*, 51(3):247–257.

Merton, R. (1975). An Asymptotic Theory of Growth Under Uncertainty. *The Review of Economic Studies*, 42(3):375–393.

Mirrlees, J. (1971). *Optimum Accumulation Under Uncertainty: the Case of Stationary Returns to Investment*, pages 36–50. Palgrave Macmillan UK, London.

Phillips, A. W. (1959). The Estimation of Parameters in Systems of Stochastic Differential Equations. *Biometrika*, 46(1-2):67–76.

Sargent, T. J. (1999). *The Conquest of American Inflation*. Princeton University Press.

Vrabie, D., Pastravanu, O. C., Abu-Khalaf, M., and Lewis, F. L. (2009). Adaptive Optimal Control for Continuous-Time Linear Systems Based on Policy Iteration. *Automatica*, 45:477–484.