Agent-level Adaptive Learning

George W. Evans

University of Oregon and University of St Andrews

Bruce McGough

University of Oregon

March 30th, 2020

Abstract

Adaptive learning is a boundedly-rational alternative to rational expectations that has been increasingly used in macroeconomics, monetary economics and financial economics. The key ideas and techniques are presented in our companion Encyclopedia article "Adaptive Learning in Macroeconomics," which focused primarily on a reduced-form approach. In the current article we show that an agent-level approach to adaptive learning can be used to provide microfoundations for adaptive learning.

Two central issues of bounded rationality are simultaneously addressed at the agent level: replacing fully-rational expectations of key variables with econometric forecasts and boundedly optimal decisions-making based on those forecast. Alternative implementations are developed, using the Real Business Cycle (RBC) model as a laboratory, and we show explicitly how to aggregate agent decision-making to obtain the path of the economy under agent-level adaptive learning. The specific implementations developed include shadow-price learning (and its anticipated-utility counterpart, iterated shadow-price learning), Euler-equation learning, and long-horizon learning.

Within the context of a linearized RBC and using simulations of time paths following a structural change – specifically a permanent increase in government spending – we illustrate the major differences of each implementation from the rational expectations paths, as well as the similarities and differences among themselves, and with reduced-form learning.

The different implementations of agent-level adaptive learning have differing advantages. A major advantage of shadow-price learning is its ease of implementation within the nonlinear RBC model. Compared to reduced-form learning, which has been the benchmark adaptive learning approach, in part because of its ease of application, agent-level learning both provides microfoundations which ensure robustness to the Lucas critique and provides the natural framework for applications of adaptive learning in heterogeneous-agent models.

1 Introduction

The benchmark implementation of adaptive learning in linear or linearized macroeconomic models is the reduced-form (RF) approach. Under this approach, the model's reduced-form expec-

tational difference equations are modified by replacing the rational expectations operator with a boundedly rational alternative. Via this method of modeling adaptive learning, stability of rational expectations equilibria (REE) can be assessed, complex learning dynamics can be introduced, and empirical and policy analysis can be conducted. The reduced-form implementation of learning, which is covered in detail in our companion Encyclopedia article, Evans and McGough (2020) (EM2020), provides a natural and efficient mechanism through which rational expectations can be replaced by adaptive learning counterparts. However, some care must be taken when introducing bounded rationality and adaptive learning into micro-founded DSGE models, especially when the invariance of the micro-foundations is central, e.g. when conducting policy analysis. After all, reduced-form learning incorporates bounded rationality after agent-level decisions, market clearing and aggregation are imposed, yet surely boundedly rational agents make boundedly rational decisions that are reflected in market-clearing and aggregation outcomes. Are these outcomes the same as, or even related to the outcomes obtained under reduced-form learning? And how, exactly, do boundedly rational agents make boundedly rational decisions? These are the questions addressed by agent-level learning.

The introduction of adaptive learning at the agent level requires the specification of the learning implementation's behavioral primitives, equilibrium notions and stability concepts. The *cognitive consistency principle* – that the sophistication of a model's agents should be comparable to the sophistication of the modeler – guides our behavioral assumptions. These assumptions identify how boundedly rational agents make forecasts *and* how they make decisions based on these forecasts. The equilibrium notions clarify the details of how coordination of per-period decisions is achieved and, when possible, specify the appropriate notions of boundedly optimal beliefs and behaviors. The stability concepts characterize whether the per-period, *temporary equilibrium* outcomes are achievable, and how and whether the induced economic dynamics result in the convergence of beliefs and behaviors to their optimal counterparts. In this paper we direct careful focus on the discussion and development of the behavioral primitives and equilibrium notions, leaving the stability concepts to passively present themselves. See Evans and McGough (2018b) for a more nuanced discussion of stability.[1]

We take, as our platform for the development of agent-level learning, a benchmark Real Business Cycle (RBC) model, with public spending used to induce policy dynamics. Rational agents in the RBC model make optimal forecasts based on their beliefs and make optimal decisions based on their forecasts; the associated equilibrium concept – specifically, a *restricted perceptions equilibrium* (RPE) – guarantees that their decisions result in aggregates that evolve under stochastic laws consistent with their beliefs. This characterization of agent behavior emphasizes that, when introducing bounded rationality at the agent level, two closely-related stands must now be taken by the modeler: we must specify how boundedly-rational agents form forecasts *and* how they make decisions given these forecasts. This latter stand, sometimes referred as *boundedly-optimal decision-making*, discriminates among implementations of agent-level learning.

There is a variety of behavioral primitives characterizing bounded optimality in the literature,

---

[1]For an application of boundedly optimal decision making based on value function learning, when agents face a discrete choice decision, see Evans, Evans, and McGough (2019).

with almost all of them exclusively applied to linearized models. Prominent implementations include the Euler-equation approach and long-horizon learning. Briefly, under the Euler-equation approach agents make one-period-ahead forecasts and form per-period decision schedules based on their corresponding first-order conditions (FOC) and flow constraints; under long-horizon learning, agents make forecasts at all forward leads and form per-period schedules based on their corresponding lifetime budget constraint. In each case these schedules are then coordinated each period in temporary equilibrium to generate the economy's aggregate outcomes.[2]

In this Encyclopedia article we develop the *shadow-price (SP) approach*, which posits that agents make per-period marginal decisions using estimates of the shadow price of savings to capture the requisite inter-temporal trade offs. We develop shadow-price learning within the context of the non-linear model: agents make forecasts using estimated linear models, but the attendant control decisions and market clearing involve no approximations. The model is in an RPE if agents' forecasting models are optimal among linear models with the same regressors. We show through simulation that the shadow-price approach provides tractable access to equilibrium dynamics, all the while allowing for retention of the RBC model's intrinsic non-linearities.

We then consider the linearized RBC model. After examining the model's REE and simulating the reduced-form learning dynamics, we study the shadow-price approach, and find that it is equivalent, in this simple environment, to Euler-equation learning. In contrast to this, we find that the long-horizon approach yields results that differ from those obtained under SP-learning. This difference, which reflects the close association of long-horizon learning and the anticipated-utility approach, is discussed in detail; further, an iterated version of SP-learning nests Euler-equation learning and anticipated utility.

We organize this paper as as follows: Section 2 covers the non-linear RBC model. Section 2.1 begins with a general modeling environment, rich enough to allow the easy incorporation of a variety of behavioral models. Section 2.2 then adopts the rational expectations hypothesis and characterizes the model's REE as a fixed point to a *T-map* that acts on plausible transition dynamics. We begin our investigations of agent-level learning in section 2.3 by considering the anticipated-utility approach within the context of the non-linear model. Here agents perceive the model's transition dynamics to be linear, and they make decisions fully optimally against these perceptions by solving the corresponding Bellman systems. The model's RPE is characterized as a fixed point of a T-map that takes linear perceptions to linear projections.

Section 2.4 uses the anticipated-utility model to motivate and develop the shadow-price approach. The shadow price of the state is naturally interpreted as the derivative of Bellman's value function, and this interpretation provides the link to anticipated utility. Instead of assuming agents uncover the value function implied by their linear perceptions, the SP approach posits that agents estimate dependence of the derivative of the value function – the shadow price – on observables, and use these estimates to forecast shadow prices and take decisions. Again, the model's RPE is

---

[2]Euler-equation learning, which has been implicitly assumed in many adaptive learning models with long-lived agents, is discussed in in the context of a linearized New-Keynesian model in Evans and Honkapohja (2006). Long-horizon learning was introduced in New-Keynesian models in Preston (2005) and employed, for example, in Eusepi and Preston (2010) and Eusepi and Preston (2012).

characterized as a fixed point of a T-map that takes linear perceptions to linear projections, but this time the projections include the forecast model for the shadow price. Section 2.5 then demonstrates the computational tractability of the shadow-price approach by conducting a policy experiment in the non-linear modeling environment. Finally, section 2.6 briefly addresses a number of issues related to agent-level implementations of adaptive learning.

It is common in both the RE and adaptive learning literature to focus attention on linearized versions of DSGE model, and so in Section 3 the linearized RBC model is analyzed. Section 3.1 develops the linearized model's REE. In Section 3.2 we conduct the standard reduced-form learning analysis, and then various implementations of agent-level learning, including shadow-price, Euler-equation and long-horizon learning, are developed in Sections 3.3 – 3.6, and the outcomes are compared with those obtained under reduced form learning and in the non-linear model. The lessons from these analyses complement both the work in Section 2.4 as well as in our companion article: the SP-approach allows for a tractable implementation of agent-level learning in non-linear environments; at the same time, reduced-form and agent-level analyses conducted in linear or linearized models do provide substantive insights. Section 4 concludes.

2 The non-linear real business cycle model

In this section we present the shadow-price implementation of agent-level learning within a canonical RBC-type modeling environment. We characterize rational behavior and the model's rational expectations equilibrium; however, because our interest involves agent-level bounded rationality and its aggregate implications, we will begin by developing our economic environment in a manner that is flexible enough to accommodate boundedly rational forecasting and decision making.

2.1 The modeling environment

There are many identical firms, each producing a homogeneous good that can be used for consumption or investment. The representative firm owns constant-returns-to-scale technology $y = zf(k,n)$, where k is the quantity of installed capital, n is the quantity of labor hired, and z is total-factor productivity, which is assumed to follow

$$z_t - \bar{z} = \rho(z_{t-1} - \bar{z}) + v_t. \tag{1}$$

Here $|\rho| < 1$ and $v_t \overset{iid}{\sim} U(-\varepsilon,\varepsilon)$, with $\varepsilon > 0$ small enough to guarantee $z_t > 0$. Firms are profit maximizers who rent capital, hire labor and sell goods in competitive markets. There are no installation costs, so the representative firm's problem is static: in this simple framework, firm behavior is independent of the assumptions governing whether agents are rational. Solving the representative firm's period t decision problem yields the following demand schedules for capital and labor:

$$
\begin{aligned}
k_t &= k^d(R_t, w_t, z_t) \\
n_t &= n^d(R_t, w_t, z_t),
\end{aligned}
$$

where R_t is the real rental rate of capital and w_t is the real wage. These schedules characterize the period t behavior of the representative firm.

There is a unit mass of households (agents) indexed as $i \in \mathscr{I}$. Even though we will assume agent homogeneity, the distinction between individual and aggregate variables is paramount when considering agent-level behaviors; therefore, we will track agents' indices even while acknowledging the attendant introduction of tedious notation.

Agent $i \in \mathscr{I}$ owns capital and is endowed with labor, and faces per-period consumption/savings and labor/leisure decisions. The agent's flow utility is given by

$$u(c_t(i)) - \chi(n_t(i)), \tag{2}$$

where $c_t(i)$ and $n_t(i)$ is the quantity of goods consumed and the quantity of labor supplied by agent i in period t. Here, as usual, $u', \chi', \chi'' > 0$ and $u'' < 0$. The flow constraint of agent i is given by

$$a_t(i) = (1 + r_t)a_{t-1}(i) + w_t n_t(i) - c_t(i) - \tau, \tag{3}$$

where $a_{t-1}(i)$ are the real assets held by agent i at the beginning of period t, τ is government spending financed by lump-sum taxes, which is taken for simplicity as constant, and which will we use for comparative dynamics experiments.[3] Also, $r_t = R_t - \delta_t$ is the real interest rate, and $\delta_t = \delta + \iota_t$ is the stochastic depreciation rate. Here it is assumed that $\iota_t \overset{iid}{\sim} U(-\bar{\iota}, \bar{\iota})$, with $0 < \bar{\iota} < \delta$.

We now turn to a general description of agent behavior, and this description begins with a discussion of forecasts. The flow utility (2) captures how an agent "thinks about" today; and, by determining $a_t(i)$, the flow constraint (3) predicts, in part, how an agent's decisions today affect his future. There are, however, additional considerations: an agent's future will also depends on the evolution of variables that are exogenous to him, e.g. future prices, as well as on how his future self will behave, e.g, in response to these future prices. A behavioral model that involves the consideration of tradeoffs between today and the future must, therefore, provide mechanisms by which an agent forecasts the future, including the behavior of his own future self.

Within the context of rational decision making, forecasting future behavior is conceptually straightforward: an agent simply assumes that his future self will make decisions optimally. In fact, it is precisely this assumption that is leveraged by the principal of optimality to transform complex sequential decision problems into the more tractable Bellman formulations. When considerations of boundedly rational decision making are in play, alternative assumptions about future behaviors are required, and become of central importance.

With our discussion of forecasts in mind, we may proceed to provide a general description of agent behavior. We assume that agent i makes per-period decisions to maximize the weighted sum of the flow utility and some measure of the discounted value of future flow utility. Period t decisions are made subject to the flow constraint, as well as with respect to whatever forecasting mechanism is adopted. We use $\psi_{t-1}(i)$ to denote this forecasting mechanism, which implicitly

---

[3]Careful modeling of heterogeneity requires the inclusions of a borrow-constraint. Because we will focus on the homogeneous-agent case, we suppress this constraint.

includes the information set available to the agent when making forecasts.[4] It may be helpful to think of $\psi_{t-1}(i)$ as representing the forecasting models used by agent i in period t, together with the data on which these models condition; alternatively, in a rational expectations equilibrium, $\psi_{t-1}(i)$ may be taken as the model-consistent conditional distribution of all variables.

Consider now the decision problem of agent i in period t. He "wakes up in the morning" holding assets $a_{t-1}(i)$ and facing exogenous spending τ. He then contemplates consumption/savings and labor/leisure decisions given his flow utility, his flow constraint, his measure of the future (still neither defined nor notated), and his forecasting mechanism $\psi_{t-1}(i)$. These contemplations yield the following demand and supply schedules for goods, capital and labor:

$$
\begin{aligned}
c_t(i) &= c^d\left(r_t, w_t, a_{t-1}(i), \psi_{t-1}(i)\right) \\
a_t(i) &= a^s\left(r_t, w_t, a_{t-1}(i), \psi_{t-1}(i)\right) \\
n_t(i) &= n^s\left(r_t, w_t, a_{t-1}(i), \psi_{t-1}(i)\right),
\end{aligned}
\tag{4}
$$

where we note that, due to the timing assumptions of the RBC model, the capital quantity $a_t(i)$ is supplied inelastically in period $t+1$. These schedules characterize the period t behavior of agent i.

The behaviors of the representative firm and the agents are coordinated through the competitive markets for goods, capital and labor. This per-period *temporary equilibrium*, a concept originally put forward by Hicks (1946), determines contemporary prices as functions of exogenous shocks and predetermined agent-level variables. Capital and labor market clearing may be written

$$
\begin{aligned}
k^d\left(r_t + \delta_t, w_t, z_t\right) &= \int_{\mathscr{I}} a_{t-1}(i)\,di \\
n^d\left(r_t + \delta_t, w_t, z_t\right) &= \int_{\mathscr{I}} n^s\left(r_t, w_t, a_{t-1}(i), \psi_{t-1}(i)\right)di.
\end{aligned}
$$

Solving for r_t and w_t provides the temporary equilibrium maps

$$
\begin{aligned}
r_t &= \mathscr{TE}_r\left(z_t, \delta_t, \{a_{t-1}(i), \psi_{t-1}(i)\}_{i \in \mathscr{I}}\right) \\
w_t &= \mathscr{TE}_w\left(z_t, \delta_t, \{a_{t-1}(i), \psi_{t-1}(i)\}_{i \in \mathscr{I}}\right).
\end{aligned}
\tag{5}
$$

These maps, together with the agents' schedules, can then be used to determine the period t values of all other variables. In particular, letting $\mathscr{TE}_r(t)$ and $\mathscr{TE}_w(t)$ be the period t temporary equilibrium maps for r and w, with the attendant arguments suppressed for notational ease, we have

$$
\begin{aligned}
c_t &= \int_{\mathscr{I}} c^d\left(\mathscr{TE}_r(t), \mathscr{TE}_w(t), a_{t-1}(i), \psi_{t-1}(i)\right)di \\
k_{t+1} &= \int_{\mathscr{I}} a^d\left(\mathscr{TE}_r(t), \mathscr{TE}_w(t), a_{t-1}(i), \psi_{t-1}(i)\right)di.
\end{aligned}
\tag{6}
$$

It is worth observing that these maps suggest a dependence of equilibrium outcomes on the entire distribution of assets and beliefs rather than only on aggregates or other sufficient statistics.

---

[4]The reason for the timing convention used here will be made clear below.

The flexible modeling environment just developed provides a conceptual relationship between the behaviors of individual agents and the implied, per period equilibrium outcomes. To make further progress and thereby produce a useful model, a stand must be taken on the precise manner in which agents form forecasts and take decisions, and we turn to this task now. We begin with the benchmark behavioral model of rationality; we then back off the assumption of rationality and discuss Kreps' model of anticipated utility before turning to our preferred implementation, the shadow-price approach.

2.2 Rational expectations

We assume that all agents have the same information set, which, in period t, includes all variables dated t and earlier. To aid notation, let $x_t = (k_t, z_t, \delta_t)$.[5] The common beliefs of agents are summarized by a map ψ, which acts as

$$(x_t, v_{t+1}, \iota_{t+1}) \xrightarrow{\psi} (r_{t+1}, w_{t+1}, x_{t+1}).$$

Thus $\psi : \mathbb{R}^3 \oplus \mathbb{R} \oplus \mathbb{R} \to \mathbb{R} \oplus \mathbb{R} \oplus \mathbb{R}^3$ captures agents' perception of the evolution of the relevant economic aggregates.

To make his time t decisions regarding consumption, savings, labor and leisure, agent $i \in \mathscr{I}$ solves the following program:

$$
\begin{aligned}
V\left(a_{t-1}(i), r_t, w_t, x_t\right) &= \max u(c_t(i)) - \chi(n_t(i)) + \beta E_t^{\psi} V\left(a_t(i), r_{t+1}, w_{t+1}, x_{t+1}\right) \\
a_t(i) &= (1+r_t)a_{t-1}(i) + w_t n_t(i) - c_t(i) - \tau \\
(r_{t+1}, w_{t+1}, x_{t+1}) &= \psi\left(x_t, v_{t+1}, \iota_{t+1}\right)
\end{aligned}
\tag{7}
$$

Here, as usual, $\beta \in (0,1)$ is the discount factor and V measures the continuation value of the state. The operator E_t^{ψ} emphasizes that expectations condition on the perceived transition dynamic ψ. The objective $(u - \chi) + \beta E V$ captures the *weighted sum of flow utility and some measure of discounted value of the future flow utility* mentioned in the previous section. With $\psi_{t-1}(i) = \psi$, the solution to this program gives policy functions corresponding to the schedules (4).

It is worth noting a redundancy in the state vector that will become more evident when homogeneity is imposed: in an REE contemporaneous prices (r_t and w_t) are determined by x_t, and therefore they could be omitted from the agent's state. However, this determination is an equilibrium outcome that results from the coordination of agent behaviors through market clearing. Before market clearing is imposed prices should be taken as independent variables against which agents form schedules. It is precisely this independence, i.e. flexibility of prices, that allows these schedules to be coordinated in temporary equilibrium.

---

[5]Throughout the paper, all vectors should be interpreted as columns, even when expressed horizontally (which we do to save space). Thus x_t is a 3×1 vector.

By imposing homogeneity we may identify $a_{t-1}(i)$ with k_t; and, remembering that x_t includes k_t, we may write market clearing as

$$k^d(r_t + \delta_t, w_t, z_t) = k_t$$
$$n^d(r_t + \delta_t, w_t, z_t) = n^s(r_t, w_t, x_t, \psi). \tag{8}$$

Solving these equations for prices and combining with agents' schedules yields the temporary equilibrium maps $\star_t = \mathscr{TE}_\star(x_t, \psi)$ for $\star \in \{r, w, a, c, n\}$. Since $a_t = k_{t+1}$, the map $\mathscr{TE}_a$, together with the exogenous processes z_t and δ_t, completely determine the evolution of the economy for fixed beliefs ψ.

For the economy to be in a rational expectations equilibrium, agents beliefs ψ must align with the implied data generating process as determined by the temporary-equilibrium coordination of agents' actions given their beliefs. To make this explicit, first note that given beliefs ψ, we may write

$$x_{t+1} = T_x(\psi)(x_t, v_{t+1}, \iota_{t+1}) \equiv \begin{pmatrix} \mathscr{TE}_a(x_t, \psi) \\ (1-\rho)\bar{z} + \rho z_t + v_{t+1} \\ \delta + \iota_{t+1} \end{pmatrix}. \tag{9}$$

Thus the realized data-generating process (DGP) is given by the map $T(\psi) : \mathbb{R}^3 \oplus \mathbb{R} \oplus \mathbb{R} \to \mathbb{R} \oplus \mathbb{R} \oplus \mathbb{R}^3$, where

$$(r_{t+1}, w_{t+1}, x_{t+1}) = T(\psi)(x_t, v_{t+1}, \iota_{t+1}) \equiv \begin{pmatrix} \mathscr{TE}_r(T_x(\psi)(x_t, v_{t+1}, \iota_{t+1}), \psi) \\ \mathscr{TE}_w(T_x(\psi)(x_t, v_{t+1}, \iota_{t+1}), \psi) \\ T_x(\psi)(x_t, v_{t+1}, \iota_{t+1}) \end{pmatrix}. \tag{10}$$

A beliefs function ψ that satisfies $T(\psi) = \psi$ characterizes a rational expectations equilibrium: given beliefs ψ, agents are making optimal forecasts and decisions which themselves induce a data generating process that confirms the agents beliefs.

Before turning to anticipated utility, it is worthwhile pausing here to make a connection between the above development of rational expectation behavior and equilibrium and related concerns when bounded rationality is assumed. Let $\mathscr{F}$ be an "appropriate" collection of continuously differentiable functions acting on $\mathbb{R}^5$, where the vague notation of "appropriateness" used here is meant to restrict attention to plausible perceived transition dynamics ψ. Then $T : \mathscr{F} \to \mathscr{F}$, and may be interpreted as taking the agents' *perceived law of motion* (PLM) to the corresponding *actual law of motion* (ALM). A fixed point of this T-map aligns perceived and realized dynamics and thus identifies an equilibrium. Exactly the same notion of a T-map is used when analyzing models of bounded rationality, and in fact comes to play a central role both in the establishment of equilibrium existence, and, through the *E-stability* principle, the assessment of equilibrium stability. See EM2020 for a thorough discussion of the E-stability principle.

2.3 Anticipated utility

The behavioral assumption of rational expectations requires that agents have a full understanding of the stochastic dynamics of the model's endogenous and exogenous variables; and even in

this simple, homogeneous-agent framework, these stochastic dynamics are characterized by a non-linear transition function in five variables. We may look to the cognitive consistency principle, discussed in the introduction, for guidance, and it suggests a simpler alternative: assume agents approximate the transition dynamics using linear models and then make optimal decisions conditional on these linear beliefs. In these types of behavioral models, the agents' linear models are commonly assumed to be updated as new data become available; however, importantly, in each period, agents take decisions *as if* their estimated forecasting models are, and will continue to be correct. This is the *anticipated-utility* approach developed in Kreps (1998).

To operationalize this approach within the context of the current model, let $x_t = (1, k_t, z_t, \delta_t)$ and assume that current and lagged x, as well as lags of prices and agent-i-specific variables, are in agent i's information set. For the purposes for forecasting aggregates, we assume that agent i holds the following forecasting models, which summarize his beliefs:

$$k_{t+1} = \psi_k'(i) \cdot x_t, \ r_{t+1} = \psi_r'(i) \cdot x_{t+1}, \ \text{and } w_{t+1} = \psi_w'(i) \cdot x_{t+1}. \tag{11}$$

Note that, in contrast to the rational case, beliefs here are captured by a finite list of parameters. We further assume, without loss of generality, that all agents know the processes governing the productivity shock and the depreciation shock.[6] Coupling this with (11) gives the PLM

$$(r_{t+1}, w_{t+1}, x_{t+1}) = \psi(i) \cdot (x_t, \nu_{t+1}, \iota_{t+1}), \tag{12}$$

where $\psi(i) = (\psi_k(i), \psi_r(i), \psi_w(i))$. The expression (12) gives the perceived linear transition dynamics for agent i, and is analogous to the perceived transition function ψ in the previous section.

Taking the perceived transition dynamics (12) as given, under the anticipated-utility approach, agent i makes time t decisions by solving following program:

$$V(a_{t-1}(i), r_t, w_t, x_t) = \max u(c_t(i)) - \chi(n_t(i)) + \beta E_t^{\psi(i)} V(a_t(i), r_{t+1}, w_{t+1}, x_{t+1})$$
$$\text{subject to } a_t(i) = (1 + r_t)a_{t-1}(i) + w_t n_t(i) - c_t(i) - \tau, \text{ and the transition (12).} \tag{13}$$

Now the expectations operator $E_t^{\psi(i)}$ emphasizes conditioning on the perceived transition dynamic $\psi(i)$. For given beliefs $\psi(i)$, the solution to the program (13) gives policy functions corresponding to the schedules (4).

As in the rational case, we may now impose homogeneity, so that $\psi(i) = \psi$ and $a_{t-1}(i) = k_t$, and then use capital and labor market clearing to obtain the temporary equilibrium maps $\star_t = \mathscr{TE}_\star(x_t, \psi)$ for $\star \in \{r, w, a, c, n\}$. Finally, as in the rational case, since $a_t = k_{t+1}$ the map $\mathscr{TE}_a$ together with the exogenous processes z_t and δ_t completely determine the evolution of the economy for fixed beliefs ψ.

It remains to discipline the allowable beliefs of agents. A rational expectations equilibrium requires that agents form beliefs optimally in the sense that their perceived transition dynamic

---

[6]Since these processes are observed and exogenous, agents can use standard procedures to estimate them without having material impact on the model's aggregate dynamics.

corresponds to the transition dynamic implied by their perceptions. A similar equilibrium concept can be applied here and in other models of bounded rationality. We say that the economy is in a restricted perceptions equilibrium if agents forecasting models are optimal among those under consideration: see EM2020 for further discussion of RPEs. For the case at hand, we require that the forecast models (12) are optimal among all linear models conditioning on the same regressors.

We can make this explicit. Just as in the rational case, the temporary equilibrium maps determine the DGP:

$$(r_{t+1}, w_{t+1}, k_{t+1}) = T(\psi)(x_t, v_{t+1}, \iota_{t+1}).$$

We may now define the vectors $T_\star(\psi) \in \mathbb{R}^4$, for $\star \in \{r, w, k\}$, as follows:

$$0 = E^{T(\psi)} \left(r_{t+1} - T_r(\psi)' \cdot x_{t+1} \right) \cdot x_{t+1}$$
$$0 = E^{T(\psi)} \left(w_{t+1} - T_w(\psi)' \cdot x_{t+1} \right) \cdot x_{t+1}$$
$$0 = E^{T(\psi)} \left(k_{t+1} - T_k(\psi)' \cdot x_t \right) \cdot x_t,$$

where $E^{T(\psi)}$ is the expectation is taken against the stationary distribution of the DGP implied by beliefs ψ. Thus $T_\star(\psi)$ is the orthogonal projection of the variable $\star$ on the relevant regressors.

Now, admittedly abusing notation, let $T(\psi)$ be the matrix with columns $T_k(\psi), T_r(\psi)$, and $T_w(\psi)$ respectively. Thus, depending on context, $T(\psi)$ may be the DGP implied by beliefs ψ or the projection of this process onto a collection of regressors. In the former case the map T can be thought of as taking the perceived transition dynamic to the implied transition dynamic, just as in the rational model. In the latter case, it takes the perceived forecast-model coefficients to the realized, or actual coefficients implied by the beliefs.

Because agents' forecast models are necessarily misspecified – after all they have linear beliefs in a non-linear world – it is the latter interpretation of the T-map that provides the needed concept for equilibrium characterization and assessment; and thus we fix this interpretation going forward. Recall that an REE is characterized by the transition dynamic determined as a fixed point of that model's T-map. The same is true here: a fixed point of this model's T-map determines the linear forecasting models, and thus determines the perceived linear transition dynamics, corresponding to and subsequently characterizing an RPE of the model. In an REE agents use the optimal non-linear forecasting model; in an RPE agents use the optimal linear forecasting model conditional on the allowed set of regressors.

Among the concerns directed against reliance on REE as the natural equilibrium notion are the unrealistic level of sophistication required of, and unreasonable amount of knowledge held by, rational agents. The latter concern speaks to the assumed understanding of the transition dynamic: how are agents suppose to have gained this understanding? The anticipated-utility approach, when coupled with the assumption of linear forecasting models, can be augmented with adaptive learning algorithms and thereby address this concern. In this augmented model, agents use recursive least squares to estimate their forecasting models based on past data; they then behave *as if* their estimated models are correct and thus solve their programming problems; markets clear and new aggregate data become available which are then used by agents to update their forecasting models

in real time. If, over time, agents beliefs converge to RPE beliefs then the question of how agents might obtain optimal beliefs is answered.

The former concern, that the sophistication needed to make optimal decisions in a dynamic stochastic environment is unrealistically high, remains unresolved in the approach just described, and indeed the only manner in which the decision problem under the anticipated-utility approach is simpler than under rational expectations is that some of the transition dynamics that are non-linear under RE are linear under anticipated utility. The agent is still required to solve a sophisticated dynamic program that is analytically intractable, and even in this very simple model already has six states and is therefore computationally burdensome.

Thus the anticipated-utility approach, coupled with linear forecasts, addresses the knowledge problem but not the sophistication problem. Furthermore, when coupled with adaptive learning, the anticipated-utility approach raises a new concern about decision making with misspecified models: it is not obvious that solving for the fully optimal policy functions is advisable if the agents' forecast models are inaccurate. For these reasons, we prefer an approach that at once simplifies the agents' decision problem and avoids the potential pitfalls of indefinitely extrapolating poorly estimated models, all the while maintaining reliance on linear forecasting rules and preserving RPE as the equilibrium concept. We call our behavioral model *the shadow-price approach*, and we turn to its development now.

2.4   The shadow-price approach

Under the anticipated-utility approach, agent *i* makes per period control decisions by balancing their impacts on his flow utility and on the continuation value implied by his beliefs. The computation of this continuation value in practice presents as an analytically intractable solution to a functional equation, a technical feat that should not be taken as necessarily achievable by economic agents. We may again look to the cognitive consistency principle for guidance, and it suggests a simpler alternative: we assume that agents estimate their continuation values and then make decisions based on these estimates.

A behavioral model along these lines, anchored to estimated value functions, is quite feasible, and in some cases quite natural: see section 2.6.3 below. However, for many models, including the model at hand, an estimate of the derivative of the value function, i.e. the shadow price, more naturally facilitates per period decision making; therefore we pursue a shadow-price implementation of agent-level learning here.

The narrative underlying our behavioral primitives is quite simple: each period each agent makes *marginal* decisions, that is, they weigh perceived marginal impacts on today and tomorrow of variations in their controls decisions. What distinguishes the shadow-price approach is that the perceived change in the value of tomorrow is measured by perceived prices, specifically, by the product of the perceived shadow price with the forecasted change in the state.

To operationalize this approach we continue to let $x_t = (1, k_t, z_t, \delta_t)$ and we adopt the same

11

assumptions about information sets and forecasting models, i.e. agent i holds the following beliefs:

$$k_{t+1} = \psi'_k(i) \cdot x_t, \ r_{t+1} = \psi'_r(i) \cdot x_{t+1}, \text{ and } w_{t+1} = \psi'_w(i) \cdot x_{t+1}, \tag{14}$$

where we continue to assume that agents know the processes governing z_t and δ_t. The first-order conditions (FOCs) for the problem (13) motivate our behavioral assumptions. Let $\lambda_t(i)$ be the period t shadow price of $a_{t-1}(i)$. The FOCs may be written

$$
\begin{array}{rl}
\text{Euler:} & u'(c_t(i)) = \beta \lambda^e_{t+1}(i) \\
\text{Labor:} & \chi'(n_t(i)) = u'(c_t(i)) \cdot w_t \\
\text{Envelope:} & \lambda_t(i) = (1+r_t)u'(c_t(i)) \\
\text{Budget:} & a_t(i) = (1+r_t)a_{t-1}(i) + w_t \cdot n_t(i) - c_t(i) - \tau,
\end{array}
\tag{15}
$$

where $\lambda^e_{t+1}(i)$ is the perception made by agent i in period t of the value of an additional unit of savings in period $t+1$.[7]

The above equations determine period t control decisions conditional $\lambda^e_{t+1}(i)$. Under rational expectations $\lambda^e_{t+1}(i) = E_t(1+r_{t+1})u'(c_{t+1}(i))$ where the expectation is taken against the equilibrium distribution of future aggregates as well as against the distribution of the agent's own future decisions, which themselves optimally condition on future aggregates. Under anticipated utility $\lambda^e_{t+1}(i) = E^{\psi(i)}_t(1+r_{t+1})u'(c_{t+1}(i))$ where the expectation is taken against the perceived distribution of future aggregates, as summarized by $\psi(i)$, as well as against the distribution of the agent's own future decisions, which themselves condition on future aggregates in a manner that is optimal given the agents perceptions. Under the shadow-price approach, agents use a linear model to estimate the dependence of the shadow price on state variables, and then use this estimated model to form $\lambda^e_{t+1}(i)$.

The shadow price of a state variable is the partial derivative of the value function against that state variable. Thus the shadow price in principle depends on the entire state vector, and so it would seem natural for the linear model used by the agent to estimate the shadow price to include all state variables. However, as noted above, in equilibrium there are redundancies in the state vector, and these redundancies will result in multi-collinearity if all state variables are included as regressors.[8] With this in mind, we assume agents think that the shadow price of savings depends on their own savings stock as well as prices; thus we propose the following PLM for shadow prices:

$$
\begin{align}
\lambda_t(i) &= \psi'_\lambda(i) \cdot (1, a_{t-1}(i), r_t, w_t) \tag{16} \\
\lambda^e_{t+1}(i) &= \psi'_\lambda(i) \cdot \left(1, a_t(i), r^e_{t+1}, w^e_{t+1}\right) = \Psi_\lambda\left(\psi(i)\right) \cdot (1, a_t(i), x_t), \tag{17}
\end{align}
$$

where equation (17), which is equation (16) stepped forward, emphasizes a linear dependence of $\lambda^e_{t+1}(i)$ on $a_t(i)$. Also, this equation further emphasizes that, given beliefs $\psi(i)$, which now include

---

[7]As is demonstrated in Evans and McGough (2018b), only state variables that are endogenous to the agent require associated estimates for shadow prices.

[8]Issues of multicolinearity are greatly mitigated in heterogeneous agent models: see Section 2.6 for discussion.

$\psi_\lambda(i)$, and given the PLMs (14), we may write $\lambda_{t+1}^e(i)$ as a linear function $\Psi_\lambda(\psi(i))$ of the time t control $a_t(i)$ and the time t states k_t, z_t and δ_t.

Given beliefs, the first-order conditions (15) and the forecast model (17) comprise the behavioral model for the shadow-price approach: the supply and demand schedules for agent i corresponding to (4), as well the agent's realized shadow price, are computed in terms of beliefs, prices and states by solving (15) and (17) for $c_t(i), n_t(i), a_{t+1}(i), \lambda_t(i)$. As above, we may now impose homogeneity, so that $\psi(i) = \psi$ and $a_{t-1}(i) = k_t$, and then use capital and labor market clearing to obtain the temporary equilibrium maps

$$\star_t = \mathscr{TE}_\star(x_t, \psi) \text{ for } \star \in \{r, w, a, c, n, \lambda\}.$$

Since $a_t = k_{t+1}$, the map $\mathscr{TE}_a$, together with the exogenous processes z_t and δ_t, completely determine the evolution of the economy for fixed beliefs ψ. Additionally, the equilibrium notation appropriate for the shadow-price approach is the same as under anticipated utility, with the added requirement that the beliefs ψ_λ are optimal among similarly conditioned linear models.

We can make this explicit. As in Section 2.3, the temporary equilibrium maps determine the DGP $T(\psi)$ implied by beliefs ψ:

$$(r_{t+1}, w_{t+1}, k_{t+1}) = T(\psi)(x_t, v_{t+1}, \iota_{t+1}).$$

Letting $\hat{x}_t = (1, r_t, w_t, k_t)$, we may further define the vectors $T_\star(\psi)$, for $\star \in \{r, w, k, \lambda\}$, as follows:

$$0 = E^{T(\psi)}(r_{t+1} - T_r(\psi)' \cdot x_{t+1}) \cdot x_{t+1}$$
$$0 = E^{T(\psi)}(w_{t+1} - T_w(\psi)' \cdot x_{t+1}) \cdot x_{t+1}$$
$$0 = E^{T(\psi)}(k_{t+1} - T_k(\psi)' \cdot x_t) \cdot x_t,$$
$$0 = E^{T(\psi)}(\lambda_t - T_\lambda(\psi)' \cdot \hat{x}_t) \cdot \hat{x}_t,$$

where the expectation is taken against the stationary distribution of the DGP implied by beliefs ψ. Thus, just as before, $T_\star(\psi)$ is the orthogonal projection of the variable $\star$ on the relevant regressors.

Again abusing notation, let $T(\psi)$ be the matrix with columns $T_k(\psi), T_r(\psi), T_w(\psi)$ and $T_\lambda(\psi)$ respectively. The map T takes linear perceptions ψ to the implied linear projections $T(\psi)$. A fixed point of this map is an RPE of the model.

2.5 Shadow-price learning

The rational expectations model and the anticipated-utility approach have associated with them significant technical impediments for both the agents in the models and for the modelers trying to analyze them. The shadow-price approach, on the other hand, is quite tractable, particularly when augmented with adaptive learning, and we pursue this exercise here.

We assume agents use recursive least squares (RLS) to estimate their linear forecast models: see EM2020 for details on RLS and real-time adaptive learning. For the model at hand this means

that agents use least-squares to estimate the parameters $\psi_\star$ for $\star \in \{k, r, w, \lambda\}$ of their PLMs (14) and (16). The recursive formulation of this estimation procedure may be combined with the temporary equilibrium map and other aspects of the model's dynamics to fully specify the economy's time path.

The state of the economy at the beginning of period t may be taken as

$$(k_t, z_{t-1}, \psi_{k,t-1}, \psi_{r,t-1}, \psi_{w,t-1}, \psi_{\lambda,t-1}, R_{\hat{x},t-1}, R_{x,t-1}),$$

where $R_\star$ is the estimated second moment matrix of the relevant regressors, which is required for recursive updating of least squares. The economy's dynamics are then given by

$$
\begin{aligned}
z_t &= \bar{z}(1-\rho) + \rho z_{t-1} + v_t \\
\delta_t &= \delta + \varepsilon_t \\
x_t &= (1, k_t, z_t, \delta_t)' \\
(r_t, w_t, k_{t+1}, c_t, n_t, \lambda_t) &= \mathscr{TE}(x_t, \psi_{t-1}) \\
\hat{x}_t &= (1, k_t, r_t, w_t)' \\
R_{x,t} &= R_{x,t-1} + \gamma(x_t \otimes x_t' - R_{x,t-1}) \\
R_{\hat{x},t} &= R_{\hat{x},t-1} + \gamma(\hat{x}_t \otimes \hat{x}_t' - R_{\hat{x},t-1}) \\
\psi_{\star,t} &= \psi_{\star,t-1} + \gamma R_{x,t}^{-1} x_t (\star_t - \psi_{\star,t-1}' x_t) \text{ for } \star \in \{k, r, w\} \\
\psi_{\lambda,t} &= \psi_{\lambda,t-1} + \gamma R_{\hat{x},t}^{-1} x_t (\lambda_t - \psi_{\lambda,t-1}' \hat{x}_t)
\end{aligned}
\tag{18}
$$

Here, $\gamma \in (0,1)$ represents the RLS *gain*, which measures the responsiveness of estimates to new data: see EM2020 for more details on alternative gain sequences.[9] Given initial beliefs, capital stock and productivity shock, the system (18) completely determines the evolution of the economy, and may be easily simulated.

As an application, we consider the following policy experiment: spending τ is unexpectedly and permanently raised. We conduct this experiment as a calibration exercise based on the following parametric forms:

$$
\begin{aligned}
\text{Flow utility: } u(c) - \chi(n) &= \log(c) + \xi \log(1-n) \\
\text{Production: } f(k,n) &= k^\alpha n^{1-\alpha}
\end{aligned}
\tag{19}
$$

The model is calibrated as $\beta = .985, \xi = 4.0, \rho = 0.9, \alpha = 1/3, \delta = .025, \bar{\tau} = 0.2, \bar{z} = 1.359, \varepsilon = .005, \gamma = 0.04$. Unless otherwise stated, this calibration is used for all of the numerical work in this paper. The economy is initialized with beliefs corresponding to the RPE of the economy. Spending is then raised by 5% and the model is simulated $50,000$ times. Figure 1 presents the cross-sectional means of the corresponding consumption and capital time-paths, expressed as proportional deviation from the pre-policy change steady-state, with the horizontal dashed lines corresponding to the post-policy change steady-state. Intuition for the figure is postponed until Section 3.3.

---

[9]It is worth observing that the appropriate choice of the gain reflects a tradeoff between filtering and tracking. Smaller gains reduce the impact on estimates of noise in the data, whereas larger gains facilitate quicker adaptation to structural change.
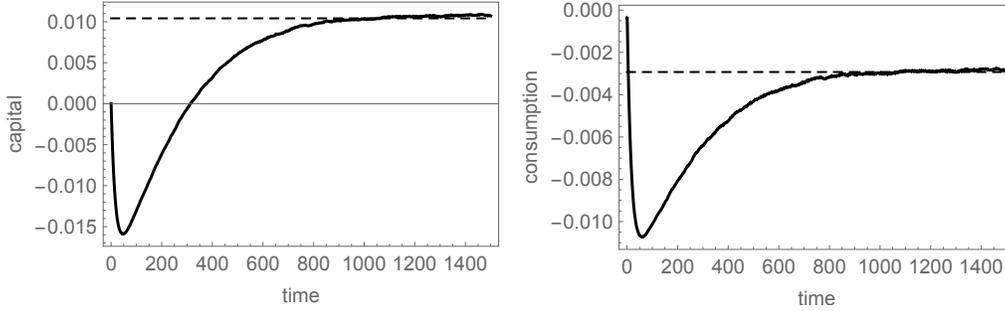
Figure 1: shadow-price learning in a non-linear model

2.6 Discussion

Before turning to the linearized models that provide more familiar environments for the assessment of bounded rationality and adaptive learning, there are some additional considerations that merit discussion. A comprehensive treatment of any of these matters is outside the scope of this paper, and so we content ourselves with brief introductions and references to the literature.

2.6.1 Restricted perceptions equilibria

If each agent is using a forecasting model that is invariant over time and optimal among the models considered, then we say the economy is in a restricted perception equilibrium; and if the models under consideration are linear in parameters then the RPE can be characterized as a solution to a system of orthogonality conditions. Focusing on this latter case, two points are worth noting. First, even though agents' forecasts in an RPE correspond to optimal linear projections, there is no a priori relationship between the RPE and the REE of the linearized model. While in many models an RPE will closely align with the linear approximation of the REE, there exist non-linear models with a unique REE that have multiple RPE associated with the same class of forecasting models, and some of these RPE may be nowhere near the REE (as measured by an appropriate metric).

Second, there are limited existence results for RPE in non-linear environments, and no general results for models with lags. The technical impediment involves the computation of the orthogonality conditions, which requires taking expectations against the asymptotic stationary distributions associated with a given collection of beliefs. These asymptotic distributions are difficult to characterize and control outside of highly stylized dynamic environments. Progress has been made in forward-looking models. See EM2020 further discussion and links to the literature.

2.6.2 Misspecified forecast models

In both the anticipated-utility and shadow-price implementations of agent-level learning discussed above, the forecasting models used are *linear* and thus necessarily misspecified. The agents' linear forecast-model misspecification then led us to the concept of a restricted perceptions equilibrium, which, as just discussed, is itself defined in terms of linear projections. The important point to emphasize here is the considerable flexibility afforded by the requirements that agents' forecast models are linear and that the attendant equilibrium is characterized in terms of this linearity. The *only* requirement is that the forecast models be *linear in coefficients*: agents are welcome to regress on whatever non-linear functions and combinations of observables they like. Thus, for example, agents acting in a regime switching environment might have regime dependent forecast models; or if the economy's intrinsic non-linearity relative to shock size is strong enough, agents and modelers may find it reasonable to condition on higher-order terms to account for curvature. In fact, through the specification of the forecast model, agents' sophistication can be quite finely tuned, allowing them to be modeled anywhere from naive to fully rational.

2.6.3 Value-function learning

The approach developed in section 2.4 above is predicated upon marginal decision making, with the future's margin captured by the estimated shadow price of the state. A natural alternative is to have agents estimate the continuation value itself, rather than its derivative, the shadow price, and then make decisions based on the estimated value function. This implementation of boundedly optimal decision making is termed *value function learning* (VF-learning). In Evans and McGough (2018b) we develop the complete theory of VF-learning within the linear-quadratic environment, and compare it to shadow-price learning and other implementations.

While in many environments VF-learning is qualitatively equivalent to SP learning, its implementation often introduces additional complexity: after all, in most models decisions are characterized in terms of the value function's derivative, which is why we emphasize SP learning in the first place. On the other hand, certain environments – particularly those involving discrete choice considerations – are particularly well-suited for VF-learning. For example, Branch and McGough (2016) use a version of VF-learning to assess the impact of bounded rationality on trade efficiencies in a money-search model; and Evans, Evans, and McGough (2019) examine the McCall model of labor search by having agents learn the optimal reservation wage.

2.6.4 Heterogeneous agent models

For simplicity we have focused on the homogeneous agent version of the real business cycle model; however, our original modeling environment was developed allowing for agent-level heterogeneity. Indeed, as indicated by (5) and (6), given the collections of beliefs, the economy's path can be recursively simulated. These same equations emphasize the computational challenge of assessing the REE even numerically: the time path of aggregates depends on the entire distribution of agents'

savings, as well as on any other idiosyncratic shocks that are incorporated into the heterogeneous agent version of the model (e.g. income shocks). These distributions become infinite-dimensional states whose evolution must be correctly tracked by both agent and modeler, and the computation of the model consistent transition dynamic governing the evolution of these distributions is a serious technical impediment.

The shadow-price approach offers an alternative that is computationally attractive. In particular, neither the modeler nor the agents need be assumed to know the model-consistent transition dynamics – they need only be able to learn about them. Furthermore, there is no requirement that agents track the distributions required to perfectly forecast prices; instead, and much more naturally, the agents can be provided simple forecasting models – perhaps vector auto-regressors in some collection of aggregate observables – which they update over time.[10] Thus to simulate a heterogeneous agent model, simply specify the forecast models agents will use, and provide them with an associated recursive estimation procedure. The only limitation is the number of agents your computer can reasonable handle, and on modern clusters simulations with millions of agents is feasible.

Heterogeneity also mitigates multicollinearity concerns that commonly arise in homogeneous models with learning agents. Mechanically, it can be the case that the natural set of regressors used by boundedly rational agents when estimating forecast models exhibits multicollinearity. This issue arises as an artifact of the homogeneity assumption: the representative agent's individual-specific variables, e.g. savings, correspond also to aggregate variables, e.g. capital stock. With agent-level heterogeneity, these types of correspondences, and hence the multicollinearity issues they induce, are mitigated.

2.6.5 Equilibrium stability

Within the context of agent-level learning, stability concerns remain paramount but also become more nuanced, as three closely related but distinct stability principles organically emerge: market stability, behavioral stability, and expectational stability. These principles reflect the view that only self-reinforcing patterns will be evidenced in actual economic outcomes. Behavioral stability requires that agents update their decision rules in such a way that, provided the environment is stationary, their behavior is asymptotically approximately optimal, where the notion of approximate is context dependent. Thus, asymptotically, agents would not have an incentive to change their decision rule. Market stability concerns how the temporary equilibrium is achieved each period. When expectations are predetermined, this concept is well-established and well-understood, e.g. Hahn (1982). If instead expectations are determined simultaneously with market outcomes, then additional related stability conditions may be required: see Evans and McGough (2018a). Finally, expectational stability governs whether and how the adaptive learning process leads the implied temporary equilibrium path to converge to the restricted perceptions equilibrium.

---

[10]It's worth noting that the associated RPE is essentially the approximation to the REE used by Krusell and Smith (1998) and others. See Giusto (2014) for adaptive learning results in the Krusell-Smith model.

3 The linearized real business cycle model

DSGE models are commonly analyzed by computing linear approximations to the model's behavioral equations and then solving the resulting linearized model. Moreover, almost all of the research on agent-level learning already in the literature has been conducted based on these linearized models. In this section we show how the shadow-price approach can be developed in linearized environments, and how it relates to other implementations of agent-level learning. A conceptual point is worth emphasizing here: in this section we treat the linearized RBC model as governing the dynamics of the economy, and not as an approximation to these dynamics.

3.1 The linearized model

When developing the agent-level approach in the non-linear model we started quite abstractly before discussing the particulars of different behavioral assumptions. To develop the linearized model we must be specific about the nature of the linearization, including the point about which we linearize. For this reason it is most convenient to, in effect, reverse the order of presentation, and instead begin with the rational model under the assumption of homogeneity and attend to its linearization before turning to associated models of boundedly rational decision making.

The system of non-linear expectational difference equations, which, together with boundary conditions, characterize the REE of our RBC model, can be written

$$
\begin{aligned}
\text{Euler:} \quad & u'(c_t) = \beta E_t (1 + r_{t+1}) u'(c_{t+1}) \\
\text{Labor:} \quad & \chi'(n_t) = u'(c_t) \cdot w_t \\
\text{Capital accumulation:} \quad & k_{t+1} = z_t f(k_t, n_t) + (1 - \delta_t) k_t - c_t - \tau \\
\text{Real rate:} \quad & r_t = z_t f_k(k_t, n_t) - \delta_t \\
\text{Real wage:} \quad & w_t = z_t f_n(k_t, n_t) \\
\text{Productivity:} \quad & z_t = \bar{z} + \rho(z_{t-1} - \bar{z}) + v_t \\
\text{Depreciation:} \quad & \delta_t = \delta + \iota_t.
\end{aligned}
\tag{20}
$$

Using $d\star$ to denote the deviation of the variable $\star$ from its non-stochastic steady state, the first-order approximation of these equations can be written

$$
\begin{aligned}
\text{Euler:} \quad & dc_t = E_t dc_{t+1} - \beta c \sigma^{-1} E_t dr_{t+1} \\
\text{Labor:} \quad & \chi'' dn_t = u' dw_t + w u' dc_t \\
\text{Capital accumulation:} \quad & dk_{t+1} = \bar{z} f_k dk_t + \bar{z} f_n dn_t + f dz_t + (1 - \delta) dk_t + k \iota_t - dc_t - d\tau \\
\text{Real rate:} \quad & dr_t = \bar{z} f_{kk} dk_t + \bar{z} f_{kn} dn_t + f_k dz_t - \iota_t \\
\text{Real wage:} \quad & dw_t = \bar{z} f_{nk} dk_t + \bar{z} f_{nn} dn_t + f_n dz_t \\
\text{Productivity:} \quad & dz_t = \rho dz_{t-1} + v_t,
\end{aligned}
$$

where all derivatives are evaluated at the non-stochastic steady state, c here denotes steady-state consumption, and $\sigma = -cu''(c)/u'(c)$ is the relative risk aversion. It is straightforward to find conformable matrices F and G, and reduced-form parameters $(\theta_c, \theta_k, \theta_z, \theta_\delta)$, so that the equilibrium of the linearized model satisfies the following reduced-form system of linear expectational difference equations:

$$dc_t = F \cdot E_t \begin{pmatrix} dc_{t+1} \\ dk_{t+1} \end{pmatrix} + G \cdot dz_t \tag{21}$$

$$dk_{t+1} = \theta_c \cdot dc_t + \theta_k \cdot dk_t + \theta_z \cdot dz_t + \theta_\delta \cdot \iota_t - d\tau \tag{22}$$

$$dz_{t+1} = \rho dz_t + v_t. \tag{23}$$

Equation (22) captures that capital is predetermined. Using standard techniques to identify the unique, appropriately bounded solution to (21) - (23), equilibrium consumption dc_t can be expressed as linear in dk_t, dz_t and ι_t. Equations (22), (23) and the government spending policy then determine the equilibrium time paths of the linearized economy.

Figure 2 reproduces the policy experiment of Figure 1 in the linear model, and for the same calibration. The dark bands provide cross-sectional quartiles and the lighter bands identify outer deciles, giving indication of the cross-sectional variation induced by productivity and depreciation innovations.[11]
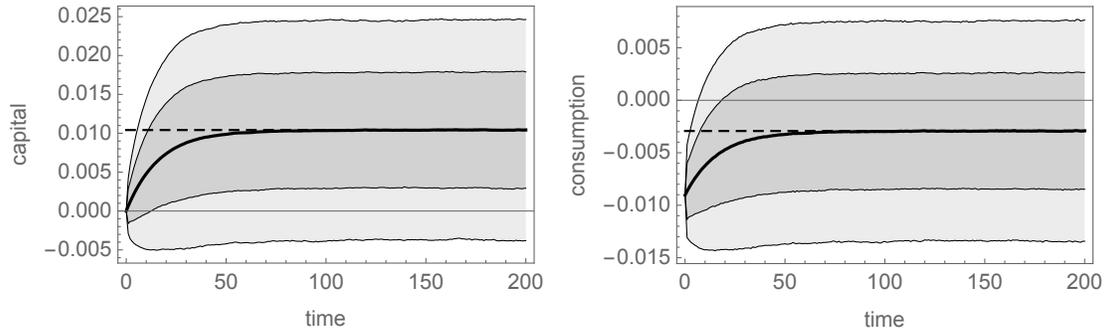

Figure 2: Rational expectations in the linear model

The RE dynamics resulting from a surprise, permanent spending rise are familiar. The policy change raises the present value of the tax burden and thus reduces expected lifetime income. Since consumption and leisure are normal goods, less of each is chosen is the long-run; thus the new long-run steady state of consumption is lower and of labor is higher. The increase in (long-run) labor ceteris paribus raises the real interest rate, inducing an increase in savings and thus resulting in an increase in the long-run level of capital. Consider now the transition to the new long-run steady state. Consumption over-shoots its new long-run level as agents immediately raise their

---

[11] As noted, the mean, quartiles and deciles are cross-sectional: at each in time they are computed across simulations. In particular, the reader should not conclude, e.g, that 50% of the simulations resulted in paths contained in the darker band.

savings in response to the higher real interest rates. Capital, which is a stock and therefore does not undergo an immediate change, converges monotonically to its new, higher steady state.

3.2 Reduced-form learning

Before discussing shadow-price learning within the context of the RBC model, we first consider *reduced-form learning* (RFL) – so-named because the rational expectations operator E_t in the model's reduced-form system (21) is simply replaced with an adaptive learning counterpart. RFL is the simplest implementation of adaptive learning in models such as the one under examination here. See EM2020 for an extensive discussion of reduced-form learning.[12]

The learning model becomes

$$dc_t = F \cdot E_t^* \begin{pmatrix} dc_{t+1} \\ dk_{t+1} \end{pmatrix} + G \cdot dz_t$$

where E_t^* denotes the as yet unspecified boundedly-rational expectations operator. To close the learning model, we must take a stand on the forecasting rules, or PLMs, used by learning agents when forming expectations. Under reduced-form learning, it is conventional (though not necessary) to assume that the agents' PLMs are consistent with the functional dependences present in the REE, and we adopt this convention here. Thus we assume that, for the purposes of forecasting, agents believe dk_{t+1} and dc_t depend linearly on dk_t, dz_t and ι_t. Letting $dx_t = (1, dk_t, dz_t, \iota_t)'$, we may write the agents' PLMs as

$$E_t^* dk_{t+1} = \psi_k' \cdot dx_t \text{ and } dc_t = \psi_c' \cdot dx_t,$$

and thus the pair $\psi = (\psi_c, \psi_k)$ capture agents' beliefs. We also continue to assume that agents know the process governing the productivity shock.

To obtain the implied data-generating process we may use the agents' PLM to write $dx_{t+1}^e = \psi_x \cdot dx_t$, where $\psi_x = \psi_x(\psi_k)$ is a 4×4 matrix function of beliefs, and thus

$$E_t^* dc_{t+1} = \psi_c' \cdot \psi_x \cdot dx_t.$$

Letting e_k be the kth-coordinate vector, and using (21) - (22), we conclude that

$$dc_t = \left(F \begin{pmatrix} \psi_c' \cdot \psi_x \\ \psi_k' \end{pmatrix} + G \cdot e_3' \right) dx_t \equiv T_c(\psi)' dx_t$$

$$dk_{t+1} = \left(\theta_c \cdot T_c(\psi)' + (-d\tau, \theta_k, \theta_z, \theta_\delta) \right) dx_t \equiv T_k(\psi)' dx_t.$$
(24)

Letting $T(\psi) = (T_k(\psi), T_c(\psi))$, we have that $(dk_{t+1}, dc_t)' = T(\psi)' dx_t$. As always, T-map takes the perceived parameters to the implied, or actual parameters. The fixed point of this T-map will align with the model's REE.

---

[12]Giannitsarou (2006) provides an example of a reduced-form learning approach to analyzing a change in tax policy in an RBC model.

20

Equations (24) can be coupled with the RLS updating equations for the agents' forecasting rules to provide the economy's dynamic path. Let ψ_t represent the beliefs of agents determined using data through period t, and let R_t be the agents' corresponding estimate of the second-moment matrix of dx_t. The state of the economy at the beginning of period t may be taken as $(dk_t, dz_{t-1}, \psi_{t-1}, R_{t-1})$. The economy's dynamics are then given by

$$
\begin{aligned}
dz_t &= \rho dz_{t-1} + v_t \\
dx_t &= (1, dk_t, dz_t, \iota_t)' \\
dc_t &= T_c(\psi_{t-1})' dx_t \\
dk_{t+1} &= T_k(\psi_{t-1})' dx_t \\
R_t &= R_{t-1} + \gamma \left(dx_t \otimes dx_t' - R_{t-1} \right) \\
\psi_{c,t} &= \psi_{c,t-1} + \gamma R_t^{-1} dx_t \left(dc_t - \psi_{c,t-1}' dx_t \right) \\
\psi_{k,t} &= \psi_{k,t-1} + \gamma R_t^{-1} dx_t \left(dk_{t+1} - \psi_{k,t-1}' dx_t \right).
\end{aligned}
\tag{25}
$$

Given initial beliefs, capital stock and productivity shock, the system (25) fully determines the evolution of the economy under reduced-form learning.

The system (25) is easily simulated, which is a very important property for applied and empirically-minded researchers and policy makers. The asymptotic properties of (25) are also reachable via the theory of stochastic recursive algorithms (SRAs). The application of this theory to learning models similar in form to (25) has been studied extensively – see EM2020 for details. Importantly, whether the asymptotic behavior of (25) leads to a (possibly degenerate) stationary distribution centered on the model's REE, can be easily assessed using the E-stability principle – a principle that arises from the careful study of SRAs.

The idea behind the principle is to associate with the algorithm the differential equation

$$
\dot{\psi} = T(\psi) - \psi,
\tag{26}
$$

and note that an REE corresponds to a rest point ψ^* of (26).[13] If this rest point is Lyapunov stable then we say that the associated REE is *E-stable*. The E-stability principle says that E-stable REE are locally stable under least-squares learning. It is important to emphasize the simplicity this principle affords stability analysis: if the real parts of eigenvalues of $DT(\psi^*)$ are less than one the dynamic system (25) can be expected to converge, in an appropriate sense, to the REE. Of course there are many details omitted in this brief discussion: see EM2020 for an overview and Evans and Honkapohja (2001) for complete details. Finally, we note that standard calibrations of the RBC model result in unique E-stable REE.

Figure 3 reproduces the policy experiment of Figure 1 in the linear model under real-time learning, and for the same calibration. The dark bands provide cross-sectional quartiles and the lighter bands give the outer deciles. We will discuss details of the learning paths for agent-level implementations as they are more amendable to useful interpretation. It is worth noting here,

---

[13]Here $\dot{\psi}$ represents the derivative of ψ with respect to "notional time." See Section 2 of EM2020.

though, some differences between this figure and its rational counterpart. First observe that the time scale is almost an order-of-magnitude larger for the learning model; second, whereas the post-shock REE mean time paths rise monotonically to their new long-run values, the mean learning paths display more complex, hump-shaped behavior; indeed, their initial post-shock behaviors comprise monotonic declines, which provide an empirically testable distinction between rational expectations and adaptive learning. Finally, the cross-sectional variation is roughly the same for the REE and adaptive learning outcomes.
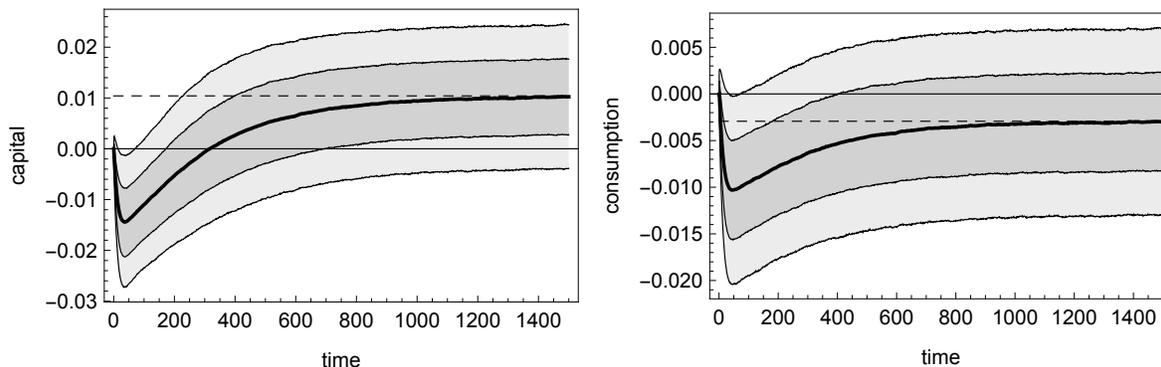

Figure 3: Reduced-form learning in the linear model

3.3 Shadow-price learning in the linear model

Reduced-form learning, as implemented in the RBC model Section 3.2, provides practical access to stability analysis and adaptive-learning dynamics, and this practicality is a principal and appropriate reason for the popularity of the approach; however, the validity of its use rests on several unexamined assumptions. After all, RF-learning introduces bounded rationality only after the collective behavior of agents has been aggregated, simplified (i.e. reduced), and then approximated to first order. Agent-level learning reverses this order-of-operations by introducing bounded rationality at the agent-level and only then aggregating and simplifying, and through this process thus examines the unexamined assumptions.

As emphasized in the previous section, agent-level learning requires that careful attention be paid to the relationships between agents' decisions and the subsequent economic aggregates. In RE models, the alignment of expectations plays a fundamental role in the coordination of agents' behavior; under bounded rationality, expectations are no longer necessarily model consistent, and an explicit modeling of price determination and market clearing – i.e. temporary equilibrium analysis – is needed to understand how actions are coordinated.

We begin by developing the decision-making process of agents under shadow-price learning in the linearized model. Intuitively, just as in the non-linear environment, each period, agents are assumed to make control decisions by balancing estimated measures of the impacts these decisions have on today and tomorrow. This intuition is operationalized by considering the linearizations of

agent i's behavioral equations (15):

$$\text{Euler:} \quad dc_t(i) = \left(\frac{\beta}{u''}\right) d\lambda^e_{t+1}(i)$$

$$\text{Labor:} \quad dn_t(i) = \left(\frac{u'}{\chi''}\right) dw_t + \left(\frac{wu'}{\chi''}\right) dc_t(i) \tag{27}$$

$$\text{Envelope:} \quad d\lambda_t(i) = \beta^{-1} u'' dc_t(i) + u' dr_t$$

$$\text{Budget:} \quad da_t(i) = \beta^{-1} da_{t-1}(i) + k \cdot dr_t + n \cdot dw_t + w \cdot dn_t(i) - dc_t(i) - d\tau.$$

To form forecasts of prices, we let $dx_t = (1, dk_t, dz_t, \iota_t)$ and adopt the same forecasting models used in the non-linear environment of Section 2.4, i.e. agent i holds the following beliefs:

$$dk_{t+1} = \psi'_k(i) \cdot dx_t, \; dr_{t+1} = \psi'_r(i) \cdot dx_{t+1}, \text{ and } dw_{t+1} = \psi'_w(i) \cdot dx_{t+1}, \tag{28}$$

where we continue to assume that agents know the processes governing dz_t and ι_t. Consistent with our work in the non-linear model, we assume agents think that the shadow price of savings depends on their own savings stock as well as on prices; thus we propose the following PLM for shadow prices:

$$\begin{aligned} d\lambda_t(i) &= \psi'_\lambda(i) \cdot (1, da_{t-1}(i), dr_t, dw_t) \\ d\lambda^e_{t+1}(i) &= \psi'_\lambda(i) \cdot \left(1, da_t(i), dr^e_{t+1}, dw^e_{t+1}\right) \equiv \psi(i) \cdot (1, da_t(i), dx_t). \end{aligned} \tag{29}$$

Equation (29) emphasizes we may write $\lambda^e_{t+1}(i)$ as the product of a beliefs' matrix $\psi(i)$, which itself depends on the agent's beliefs $\psi_\star(i)$ for $\star \in \{k, r, w, \lambda\}$, and a vector comprised of the time t control $da_t(i)$ and the time t prices and states dk_t, dz_t and ι_t, which are taken as observable.

Equations (27) - (29) comprise the behavioral primitives of shadow-price learning in the linear model and operationalize boundedly optimal decision-making in this framework. In particular, we may solve these equations for period t decisions to obtain the following schedules:

$$\begin{aligned} dc_t(i) &= dc\left(da_{t-1}(i), dw_t, dr_t, dx_t, \psi(i)\right) \\ dn_t(i) &= dn\left(da_{t-1}(i), dw_t, dr_t, dx_t, \psi(i)\right) \\ da_t(i) &= da\left(da_{t-1}(i), dw_t, dr_t, dx_t, \psi(i)\right) \\ d\lambda_t(i) &= d\lambda\left(da_{t-1}(i), dw_t, dr_t, dx_t, \psi(i)\right). \end{aligned} \tag{30}$$

Thus, given last period's savings and current prices and states, and given beliefs, these equations determine agent i's period t decisions.

Now let
$$dn^d_t = dn^d(dk_t, dw_t, dr_t, dz_t) \text{ and } dk^d_t = dk^d(dn_t, dw_t, dr_t, dz_t)$$

be the linearized labor and capital demands of the representative firm. Prices dr_t and dw_t are then determined in temporary equilibrium via market clearing:

$$dk^d(dn^d_t, dw_t, dr_t, dz_t) = \int_{\mathscr{I}} da_{t-1}(i) di$$

$$dn^d(dk^d_t, dw_t, dr_t, dz_t) = \int_{\mathscr{I}} dn\left(da_{t-1}(i), dw_t, dr_t, dx_t, \psi(i)\right) di. \tag{31}$$

Equations (31) may be used in conjunction with equations (30) to determine the period t temporary-equilibrium decisions of each agent.

Proceeding as before impose, we impose homogeneity, so that $\psi(i) = \psi$ and $da_{t-1}(i) = dk_t$, and then use capital and labor market clearing to obtain the temporary equilibrium vectors $\mathscr{TE}_\star(\psi)$ such that

$$\star_t = \mathscr{TE}_\star(\psi) \cdot dx_t \text{ for } \star \in \{dr, dw, da, dc, dn, d\lambda\}. \tag{32}$$

Since $da_t = dk_{t+1}$, the vector $\mathscr{TE}_{da}(\psi)$, which we also denote $\mathscr{TE}_{dk}(\psi)$ to simplify later notation, together with the exogenous processes dz_t and ι_t, completely determine the evolution of the economy for given beliefs ψ.

The temporary equilibrium vectors may be used to write dz_t and ι_t as linear functions of $d\hat{x}_t \equiv (1, dk_t, dr_t, dw_t)$, and thus we may write $d\lambda_t = T_{d\lambda}(\psi) \cdot d\hat{x}_t$. Letting $T_\star(\psi) = \mathscr{TE}_\star(\psi)$ for $\star \in \{dr, dw, dk\}$ completes the definition of the T map associated with shadow-price learning in the linear model. A fixed point of this T-map will correspond to the REE of the linearized model, and the T-map can be used to assess E-stability.

Real time learning within this environment may be examined by assuming agents use RLS to estimate their forecast models. The state of the economy at the beginning of period t may be taken as $(dk_t, dz_{t-1}, \psi_{t-1}, R_{\hat{x},t-1}, R_{x,t-1})$. The economy's dynamics are then given by

$$
\begin{aligned}
dz_t &= \rho dz_{t-1} + v_t \\
dx_t &= (1, dk_t, dz_t, \iota)' \\
d\star_t &= T_\star(\psi_{t-1})' dx_t \text{ for } \star \in \{c, r, w, a\}, \text{ and } dk_{t+1} = da_t \\
d\hat{x}_t &= (1, dk_t, dr_t, dw_t)' \\
d\lambda_t &= T_\lambda(\psi_{t-1})' d\hat{x}_t \\
R_{x,t} &= R_{x,t-1} + \gamma\left(dx_t \otimes dx_t' - R_{x,t-1}\right) \\
R_{\hat{x},t} &= R_{\hat{x},t-1} + \gamma\left(d\hat{x}_t \otimes d\hat{x}_t' - R_{\hat{x},t-1}\right) \\
\psi_{\lambda,t} &= \psi_{\lambda,t-1} + \gamma R_{\hat{x},t}^{-1} d\hat{x}_t \left(\lambda_t - \psi_{t-1}' d\hat{x}_t\right) \\
\psi_{\star,t} &= \psi_{\star,t-1} + \gamma R_{x,t}^{-1} dx_t \left(\star_t - \psi_{\star,t-1}' dx_t\right) \text{ for } \star \in \{r, w, a\}.
\end{aligned}
\tag{33}
$$

Given initial beliefs, capital stock and productivity shock, the system (33) fully determines the evolution of the economy under shadow-price learning.

Figure 4 reproduces the policy experiment of Figure 1 in the linear model, and for the same calibration. As expected, convergence of the mean paths of consumption and capital to their new long-run values obtains.

There are a number of comments to make regarding Figure 4, and its comparison to the figures produced under non-linear SP-learning and RF-learning. We begin with a discussion of the economics underlying the witnessed behavior before turning to comparisons. Under SP learning, the long-run behavior of the economy is the same as in REE: this is to be expected as it is
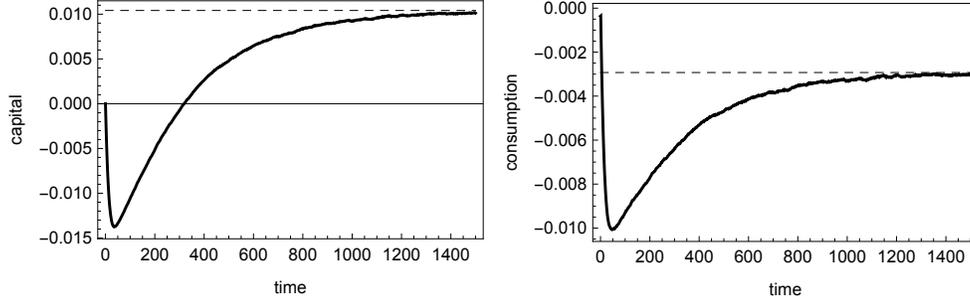
24

Figure 4: shadow-price learning in the linear model

well-known that in the RBC model the REE is stable under reasonable forms of adaptive learning, and direct computation of the T-map's derivative shows that our implementation retains this stability. The transition dynamics under SP learning, however, are markedly different from the REE counterparts.

To explain this difference it is helpful to first recall the REE model and its underlying mechanisms: rational agents fully and immediately incorporate the present value of their increased tax burden as well as the current and future economy-wide impact this policy change will have on aggregate dynamics; this causes agents to sharply decrease consumption – overshooting the long-run steady state – in order to increase savings while interest rates are high, before converging to the new, lower, long-run steady state. Correspondingly, capital monotonically rises to its long-run steady state.

Now consider the consumption and capital paths under SP learning in either the linear or non-linear setting. Initially, capital falls for a number of periods as a direct consequence of the rise in public spending. This fall in capital takes the form of a decrease in savings from the perspective of the agents, thereby raising saving's shadow price; further, the fall in capital increases the expected real interest rate. Both effects induce agents to slowly shift consumption towards savings, with the inertial response reflecting the gradual adjustment of agents' forecast parameters ψ. Only over time do the beliefs of agents adjust to the permanent policy change, allowing the economy to converge to its new long-run steady state.

Turning now to comparisons across models of learning, the mean paths obtained under SP-learning are quite similar in the linear and the non-linear models. This is not surprising in a model with small shocks and the modest curvature induced by log utility and Cobb-Douglass production. Also, the mean paths obtained under SP-learning and RF-learning in the linear model are almost identical, which reflects that the representative agent's FOCs are almost exactly the reduced-form equations characterizing the REE. We will have more to say about this in the next section on Euler-equation learning. There is, however, an important difference between the time paths, even in this simple modeling environment: under RF-learning there is *no* immediate impact on consumption, which reflects the omission of the effect, at the agent level, of a rise in public spending.

25

3.4 Euler-equation learning

As mentioned in the introduction, a number of agent-level implementations have been proposed within the linearized model framework, and among them, the two most prominent are *Euler-equation learning* and *Long-horizon learning*. In this section we develop Euler-equation learning, and show that in this simple environment it is equivalent to shadow-price learning. In the next section we develop long-horizon learning, discuss its relationship with the anticipated-utility approach, and finally show that it is analogous to an iterated version of shadow-price learning that itself replicates anticipated utility in the linear model.

Under shadow-price learning, agents use estimated shadow prices to measure tradeoffs and inform decisions. Under Euler-equation learning agents make decisions in a manner consistent with their Euler equation, modified to include their boundedly rational forecasts, as well as consistent with their intra-temporal FOC. To operationalize this proposal, recall the linearized first order conditions of agent i:

$$
\begin{aligned}
\text{Euler:} \quad & dc_t(i) = dc^e_{t+1}(i) - \beta c\sigma^{-1}dr^e_{t+1} \\
\text{Labor:} \quad & dn_t(i) = \left(\frac{u'}{\chi''}\right)dw_t + \left(\frac{wu'}{\chi''}\right)dc_t(i) \\
\text{Budget:} \quad & da_t(i) = \beta^{-1}da_{t-1}(i) + k\cdot dr_t + n\cdot dw_t + w\cdot dn_t(i) - dc_t(i) - d\tau.
\end{aligned}
\tag{34}
$$

To form forecasts of prices, we again let $dx_t = (1, dk_t, dz_t, \iota_t)$ and adopt the same forecasting models as above: agent i holds the following beliefs:

$$
dk_{t+1} = \psi'_k(i)\cdot dx_t, \ dr_{t+1} = \psi'_r(i)\cdot dx_{t+1}, \text{ and } dw_{t+1} = \psi'_w(i)\cdot dx_{t+1},
\tag{35}
$$

where we continue to assume that agents know the processes governing dz_t and ι_t. As is evident from the Euler equation, agent i must forecast his own future consumption. Analogous to shadow-price learning, we assume the agent uses a linear forecast model of the form

$$
\begin{aligned}
dc_t(i) &= \psi'_c(i)\cdot(1, da_{t-1}(i), dr_t, dw_t), \\
dc^e_{t+1}(i) &= \psi'_c(i)\cdot(1, da_t(i), dr^e_{t+1}, dw^e_{t+1}).
\end{aligned}
\tag{36}
$$

Using these forecasting models, we may write $c^e_{t+1}(i)$ as a linear function of the time t control $da_t(i)$ and the time t states dk_t, dz_t and ι, which are taken as observable.

Equations (34) - (35) comprise the behavioral primitives of Euler-equation learning in the linear model and operationalize boundedly optimal decision-making in this framework. In particular, we may solve these equations for period t decisions to obtain the following schedules:

$$
\begin{aligned}
dc_t(i) &= dc\left(da_{t-1}(i), dw_t, dr_t, dx_t, \psi(i)\right) \\
dn_t(i) &= dn\left(da_{t-1}(i), dw_t, dr_t, dx_t, \psi(i)\right) \\
da_t(i) &= da\left(da_{t-1}(i), dw_t, dr_t, dx_t, \psi(i)\right).
\end{aligned}
\tag{37}
$$

26

Thus, given last period's savings and current prices and states, and given belief parameters, these equations determine agent i's period t decisions. Note that $dc_{t+1}^e(i)$ is simultaneously determined with the agent's controls.

Continuing as before, we impose homogeneity, so that $\psi(i) = \psi$ and $da_{t-1}(i) = dk_t$, and use capital and labor market clearing to obtain the temporary equilibrium vectors $\mathscr{T}\mathscr{E}_\star(\psi)$ such that

$$\star_t = \mathscr{T}\mathscr{E}_\star(\psi) \cdot dx_t \text{ for } \star \in \{dr, dw, da, dc, dn\}.$$

Since $da_t = dk_{t+1}$, the vector $\mathscr{T}\mathscr{E}_{da}(\psi)$, which we also denote $\mathscr{T}\mathscr{E}_{dk}(\psi)$ to simplify later notation, together with the exogenous processes dz_t and ι_t, completely determine the evolution of the economy for fixed beliefs ψ.

The temporary equilibrium vectors may be used to write dz_t and ι_t as linear functions of $d\hat{x}_t \equiv (1, dk_t, dr_t, dw_t)$, and thus we may write $dc = T_{dc}(\psi) \cdot d\hat{x}_t$. Letting $T_\star(\psi) = \mathscr{T}\mathscr{E}_\star(\psi)$ for $\star \in \{dr, dw, dk\}$ completes the definition of the T map associated with shadow-price learning in the linear model. A fixed point of this T-map will correspond to the REE of the linearized model, and the T-map can be used to assess E-stability.

The policy experiment under Euler-equation learning results in mean paths that are indistinguishable from their shadow-price counterparts; hence we do not include an associated figure. This tight relationship reflects that the shadow price measures the interest-rate adjusted marginal utility of consumption, whence acting to meet the Euler equation is qualitatively equivalent to acting optimally against a shadow-price forecast. Minor quantitative differences reflect asymmetries in the way different shock realizations impact specific estimates. In more complex models, particularly in models with more endogenous states than controls, e.g. habit persistence, shadow-price learning and Euler-equation learning are both quantitatively and qualitatively distinct. The precise conditions under, and manner in which, shadow-price learning and Euler-equation learning are equivalent is detailed in Evans and McGough (2018b).

Finally, we return to the observation that the paths of consumption and capital – indeed the paths of all aggregates – under reduced-form learning are almost identical to the those obtained under shadow-price learning, and thus under Euler-equation learning. As noted by Evans and Honkapohja (2006) and Honkapohja, Mitra, and Evans (2013) in their exposition of Euler-equation learning, this close association is commonly used as a justification for adopting the comparatively simple implementation of reduced-form learning.

3.5 Long-horizon learning

Our implementation of Euler-equation learning took the per-period linearized FOCs as behavioral primitives: agents made decisions based on trade-offs measured via one-period-ahead forecasts. long-horizon learning adopts a more sophisticated behavioral view of decision-making: agents are assumed to take their estimated forecast rules as indefinitely valid, and to make fully optimal

decisions, at least to first order, given their validity.[14] In this way, long-horizon learning can be viewed as a first-order implementation of the anticipated-utility approach.[15]

For the environment under consideration, this plays out as follows: by combining their linearized life-time budget constraint (LTBC) with their linearized Euler equations at all iterations, households determine consumption, labor supply and savings schedules conditional on last period's savings and the expected time paths of interest rates, wages and taxes.[16]

A detailed development of the agent-level schedules is somewhat tedious, and we will omit steps, but with the intention of providing enough details that the reader, with reasonable effort, could reproduce the needed formulae. To make matters easier, we impose straight-away the parametric functional forms (19) used in the simulation examples.

Let R_t^n be the *expected* n-step-ahead discount rate at time t:

$$R_t^n = \prod_{k=1}^{n}(1 + r_{t+k}^e)^{-1}, \text{ with } R_t^0 = 1.$$

Agent i's iterated Euler equations may be written $\beta^n c_t(i) = R_t^n(i)c_{t+n}^e(i)$, and his intra-temporal FOC implies

$$w_{t+i}^e(i)n_{t+i}^e(i) = w_{t+i}^e(i) - \xi c_{t+i}^e(i).$$

His expected LTBC is

$$\sum_{i \geq 0} R_t^i(i)c_{t+i}^e(i) = (1 + r_t)a_{t-1}(i) + \sum_{i \geq 0} R_t^i(i)\left(w_{t+i}^e(i)n_{t+i}^e(i) - \tau_{t+i}^e(i)\right), \text{ or}$$

$$\left(\frac{1+\xi}{1-\beta}\right)c_t(i) = (1 + r_t)a_{t-1}(i) + w_t - \tau_t + \text{PV}_t^e(w,i) - \text{PV}_t^e(\tau,i), \tag{38}$$

where the second equation simplifies the LTBC using the iterated Euler equations and intra-temporal FOC, and where $\text{PV}_t^e(\star, i)$ is the expected present value of the variable $\star$ against the discount rates $R_t^n(i)$.

Using $dR_t^n(i) = \beta^{n+1}\sum_{i=1}^{n} dr_{t+i}^e(i)$, we may linearize (38) and then join it with linearized

---

[14]In models with infinitely-lived agents, decision-making based on planning horizons H, where $1 < H < \infty$, are developed in Branch, Evans, and McGough (2013).

[15]We note that long-horizon learning allows for the analysis of anticipated policy changes. See Evans, Honkapohja, and Mitra (2009).

[16]Our treatment here of the long-horizon version of the RBC model corresponds to that given in Mitra, Evans, and Honkapohja (2013) and Mitra, Evans, and Honkapohja (2019). See also Evans, Honkapohja, and Mitra (2009).

versions of the intra-temporal constraint and flow budget constraint to get the following schedules

$$dc_t(i) = dc\left(da_{t-1}(i), dw_t, dr_t, \sum_{i\geq 1}\beta^i dr^e_{t+i}(i), \sum_{i\geq 1}\beta^i dw^e_{t+i}(i)\right)$$

$$dn_t(i) = dn\left(da_{t-1}(i), dw_t, dr_t, \sum_{i\geq 1}\beta^i dr^e_{t+i}(i), \sum_{i\geq 1}\beta^i dw^e_{t+i}(i)\right) \tag{39}$$

$$da_t(i) = da\left(da_{t-1}(i), dw_t, dr_t, \sum_{i\geq 1}\beta^i dr^e_{t+i}(i), \sum_{i\geq 1}\beta^i dw^e_{t+i}(i)\right).$$

These behavioral rules are first-order approximations to the fully optimal decision plans associated with the expected future path of prices; said differently, they are linear approximations to optimal decisions conditional on current beliefs. In this way, long-horizon learning can be viewed as an anticipated-utility approach applied to a linearized model.

To form forecasts of prices, we again let $dx_t = (1, dk_t, dz_t, \iota_t)$ and adopt the same forecasting models as above: agent i holds the following beliefs:

$$dk_{t+1} = \psi'_k(i) \cdot dx_t, \; dr_{t+1} = \psi'_r(i) \cdot dx_{t+1}, \text{ and } dw_{t+1} = \psi'_w(i) \cdot dx_{t+1}, \tag{40}$$

where we continue to assume that agents know the processes governing dz_t and ι_t. Using these forecasting models, we can write $\sum_{i\geq 1}\beta^i d\star^e_{t+i}(i)$ for $\star = r, w$ as linear functions of dx_t.

Equations (39) - (40) comprise the behavioral primitives of long-horizon learning in the linear model and operationalize boundedly optimal decision-making in this framework. In particular, we may solve these equations for period t decisions to obtain the following schedules:

$$dc_t(i) = dc(da_{t-1}(i), dw_t, dr_t, dx_t, \psi(i))$$
$$dn_t(i) = dn(da_{t-1}(i), dw_t, dr_t, dx_t, \psi(i)) \tag{41}$$
$$da_t(i) = da(da_{t-1}(i), dw_t, dr_t, dx_t, \psi(i)).$$

Continuing as before, we impose homogeneity, so that $\psi(i) = \psi$ and $da_{t-1}(i) = dk_t$, and then use capital and labor market clearing to obtain the temporary equilibrium vectors $\mathscr{TE}_\star(\psi)$ such that

$$\star_t = \mathscr{TE}_\star(\psi) \cdot dx_t \text{ for } \star \in \{dr, dw, da, dc, dn\}.$$

Since $da_t = dk_{t+1}$, the vector $\mathscr{TE}_{da}(\psi)$, which we also denote $\mathscr{TE}_{dk}(\psi)$ to simplify later notation, together with the exogenous processes dz_t and ι_t, completely determine the evolution of the economy for fixed beliefs ψ. Letting $T_\star(\psi) = \mathscr{TE}_\star(\psi)$ for $\star \in \{dr, dw, dk\}$ provides the definition of the T map associated with long-horizon learning in the linear model. A fixed point of this T-map will correspond to the REE of the linearized model, and the T-map can be used to assess E-stability. The associated development of the the real-time learning dynamics is routine.

Figure 5 considers our policy experiment in the linear model under long-horizon (LH) learning. A practical issue that is particularly important within the context of LH-learning is the choice

of the gain parameter. The dependence of agents's decisions on long-horizon forecasts often induces very strong negative expectational feedback, which can be destabilizing unless the gain is quite small. This is in contrast to shadow-price, Euler-equation, and reduced-form learning, where the induced expectational feedback is typically positive.[17] For example, in the model at hand, under Euler-equation learning, the dominant eigenvalue has real part equal to 0.9167 whereas under long-horizon learning the dominant non-zero eigenvalue is less than -4.0.[18] Following Eusepi and Preston (2011), we set the gain at 0.002 for our LH-learning experiment.

The behavior witnessed here is somewhat different from the time paths associated with shadow-price and Euler-equation learning. Three comments are warranted here. First, in contrast to SP-learning, the impact-effect of the spending increase results in over-shooting of consumption, which in part reflects that the agent's behavioral rule (39) does explicitly incorporate the true present value of the increased tax burden. This distinction from behavior implied by shadow-price learning is empirically testable. Second, the agent still has to learn how the policy change will alter the dynamics of aggregate capital, and here his behavior is analogous to his SP learner counterpart: as capital initially falls the agent becomes overly pessimistic, resulting in the sharp decline in consumption; as his beliefs adjust to mitigate the pessimism the economy returns to the REE.

Under long-horizon learning, agents move somewhat more quickly toward the new equilibrium, which reflects that through their long-horizon impacts, changes in beliefs result in larger changes in behaviors than would be induced in the analogous SP-learner. At a mechanical level, this difference reflects the eigenvalues mentioned above. Intuitively, eigenvalues with real part near one correspond to beliefs that are almost self-fulfilling; thus they result in smaller forecast errors and slower adjustment. Analogously, large negative eigenvalues correspond to beliefs that are contradicted by outcomes and thus result in larger forecast errors and more rapid adjustment. That such eigenvalues arise under long-horizon learning reflects that estimated forecast models are used to form expectations far into the future, thus compounding the impact of poorly estimated forecast-model parameters.
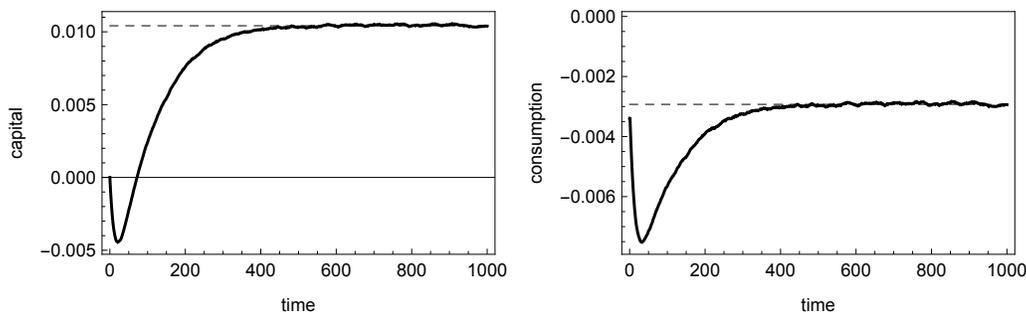

Figure 5: long-horizon learning in the linear model, with gain $\gamma = 0.002$

---

[17]The analytical results underlying this phenomenon is provided in Evans, Guesnerie, and McGough (2019).

[18]These contrasting stability properties are magnified by larger gains. For example, with the gain set to 0.1, of $50,000$ LH-learning simulations 1.44% were unstable; the corresponding proportion for SP-learning was 0.006%.

3.6 Shadow-price learning and anticipated utility

The shadow-price approach can be modified to accord fully with the paradigm of anticipated utility. Under the standard implementation of SP learning agents choose to neglect that the realized dependence of their shadow price on savings and prices differs from the estimated dependence that they used to form forecasts. There are many reasons why the agents might make this choice, including cognitive abilities and costs, and that the discrepancy might reflect errors in their state-variable forecast models, and therefore correcting the discrepancy might result in poorer decisions. On the other hand, if we adopt the anticipated-utility perspective that agents view their state-variable forecast models as accurate and that they are cognitively unconstrained then it makes sense to assume agents do note and correct the discrepancy.

The corrective behavior induced by the anticipated-utility approach may be implemented by finding a fixed point of the T-map associated with shadow-price perceptions. More specifically, using the notation of Section 3.3 and considering only the homogeneous case, we let $\hat{\psi} = (\psi_k, \psi_r, \psi_w)$ denote the agent's transition beliefs, and we define $\psi_\lambda^* (\hat{\psi})$ as the solution to

$$\psi_\lambda^* (\hat{\psi}) = T_{d\lambda} \left(\psi_\lambda^* (\hat{\psi}), \hat{\psi} \right).$$

Then $\psi_\lambda^* (\hat{\psi})$ corresponds to the time-invariant shadow price forecasting model consistent with the savings behavior it induces, given the fixed transition beliefs $\hat{\psi}$. By assuming agents use $\psi_\lambda^* (\hat{\psi})$ to forecast shadow prices, we are, in effect, assuming agents act as if their transition beliefs are correct and will hold indefinitely, and so they take the time (and have the sophistication) to compute the implied the time-invariant forecast model for shadow prices.
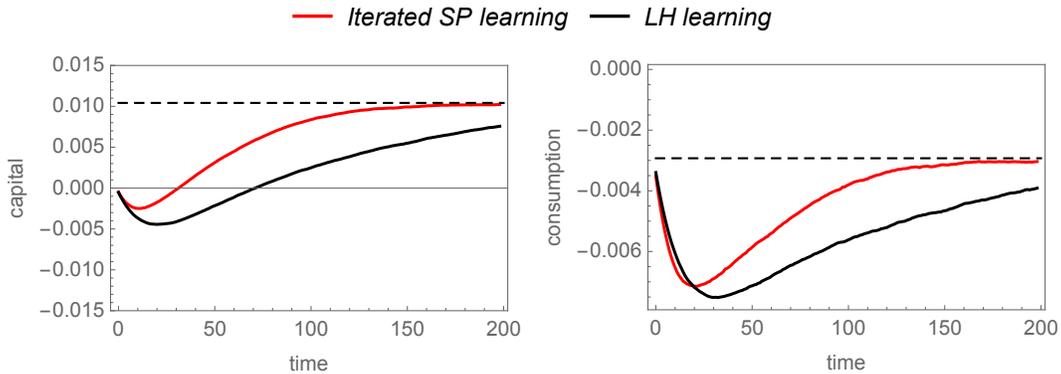

Figure 6: Iterated shadow-price learning in the linear model

Adopting the assumption that agents use ψ_λ^* to forecast shadow prices removes the explicit dependence of the model's temporary equilibrium maps on ψ_λ (because it is now determined by $\hat{\psi}$), and so the maps may be written

$$\star_t = \mathscr{T}\mathscr{E}_\star (\hat{\psi}) \cdot dx_t \text{ for } \star \in \{dr, dw, da, dc, dn, d\lambda\}.$$

With this modification, the remainder of the modeling and analysis proceeds just as it has in previous treatments.

Figure 6 compares the outcomes of the policy experiment under long-horizon and iterated shadow-price learning.[19] Focusing on the consumption panel, the impact effect of the spending increase is qualitatively and quantitatively the same for both implementations: similar to RE and in contrast to SP-learning, both consumption paths over-shoot the new steady-state level. In contrast to RE and similar to SP-learning, in the periods following the impact effect, both consumption paths continue to decline as agents' forecasting models are updated to asymptotically align with the new public spending level. Finally, after reaching similar troughs, both consumption paths converge monotonically to the new steady-state level, and at a rate that is faster than the standard implementation of SP learning.

4   Conclusions

In EM2020 we survey the central ideas of adaptive learning, focusing on the standard reduced-form analysis, within a range of macroeconomic models. This companion article has examined the agent-level approach to adaptive learning, emphasizing the interaction between boundedly optimal decision-making and boundedly rational forecasting, as well as its aggregate implications. The SP approach to agent-level learning has a major advantage in its ability to be implemented in non-linear environments. Though not addressed in this article, SP-learning is also able to accommodate complex heterogeneity.

In a linearized DSGE environment, we have examined several alternative implementations of the agent-level approach: SP-learning, Euler-equation learning, long-horizon learning, and iterated SP-learning. We compare and contrast these approaches within a linearized RBC model, focusing on a simple policy change involving a permanent increase in government spending.

Within the context of linearized representative-agent models, reduced-form learning provides tractable access to many central considerations of adaptive learning including asymptotic stability, transition dynamics and empirical analysis. Furthermore, it is often possible to justify reduced-form learning using a particular implementation of the agent-level approach. Agent-level learning may be more appropriate for questions of policy analysis, particularly when the Lucas critique is a concern, and is likely to be of increasing importance because of its ease of applicability to heterogeneous-agent models.

---

[19]For the graphs in Figure 6, the gain used for simulating iterated SP learning is $\gamma = 0.04$ as above, while for comparison purposes the gain used for simulating LH learning is $\gamma = 0.004$.

References

ARROW, K. J., AND M. INTRILIGATOR (eds.) (1982): *Handbook of Mathematical Economics, Volume II*. North-Holland, Amsterdam.

BRANCH, W., AND B. MCGOUGH (2016): "Heterogeneous beliefs and trading inefficiencies," *Journal of Economic Theory*, 163, 786 – 818.

BRANCH, W. A., G. W. EVANS, AND B. MCGOUGH (2013): "Finite Horizon Learning," in Sargent and Vilmunen (2013), chap. 9.

EUSEPI, S., AND B. PRESTON (2010): "Central Bank Communication and Expectations Stabilization," *American Economic Journal: Macroeconomics*, 2, 235–271.

——— (2011): "Expectations, Learning and Business Cycle Fluctuations," *American Economic Review*, 101, 2844–2872.

——— (2012): "Debt, Policy Uncertainty and Expectations Stabilization," *Journal of the European Economic Association*, 10, 860–886.

EVANS, D., G. W. EVANS, AND B. MCGOUGH (2019): "Learning When to Say No," Working paper, University of Oregon.

EVANS, G. W., R. GUESNERIE, AND B. MCGOUGH (2019): "Eductive Stability in Real Business Cycle Models," *Economic Journal*, 129, 821–852.

EVANS, G. W., AND S. HONKAPOHJA (2001): *Learning and Expectations in Macroeconomics*. Princeton University Press, Princeton, New Jersey.

——— (2006): "Monetary Policy, Expectations and Commitment," *Scandinavian Journal of Economics*, 108, 15–38.

EVANS, G. W., S. HONKAPOHJA, AND K. MITRA (2009): "Anticipated Fiscal Policy and Learning," *Journal of Monetary Economics*, 56, 930–953.

EVANS, G. W., AND B. MCGOUGH (2018a): "Interest-Rate Pegs in New Keynesian Models," *Journal of Money, Credit and Banking*, 50, 939–965.

——— (2018b): "Learning to Optimize," mimeo, University of Oregon.

——— (2020): "Adaptive Learning and Macroeconomics," *Oxford Research Encyclopedia of Economics and Finance,* in preparation.

GIANNITSAROU, C. (2006): "Supply-Side Reforms and Learning Dynamics," *Journal of Monetary Economics*, 53, 291–309.

GIUSTO, A. (2014): "Adaptive Learning and Distributional Dynamics in an Incomplete Markets Model," *Journal of Economic Dynamics and Control*, 40, 317–333.

HAHN, F. (1982): "Stability," in Arrow and Intriligator (1982), chap. 16, pp. 746–793.

HICKS, J. R. (1946): *Value and Capital, Second edition*. Oxford University Press, Oxford UK.

HONKAPOHJA, S., K. MITRA, AND G. W. EVANS (2013): "Notes on Agents' Behavioral Rules Under Adaptive Learning and Studies of Monetary Policy," in Sargent and Vilmunen (2013), chap. 5.

JACOBS, D., E. KALAI, AND M. KAMIEN (eds.) (1998): *Frontiers of Research in Economic Theory*. Cambridge University Press, Cambridge.

KREPS, D. M. (1998): "Anticipated Utility and Dynamic Choice," in Jacobs, Kalai, and Kamien (1998), pp. 242–274.

KRUSELL, P., AND A. SMITH (1998): "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, (106), 867–896.

MITRA, K., G. W. EVANS, AND S. HONKAPOHJA (2013): "Policy Change and Learning in the RBC Model," *Journal of Economic Dynamics and Control*, 37, 1947–1971.

——— (2019): "Fiscal Policy Multipliers in an RBC Model with Learning," *Macroeconomic Dynamics*, 23, 240–283.

PRESTON, B. (2005): "Learning about Monetary Policy Rules when Long-Horizon Expectations Matter," *International Journal of Central Banking*, 1, 81–126.

SARGENT, T. J., AND J. VILMUNEN (eds.) (2013): *Macroeconomics at the Service of Public Policy*. Oxford University Press.