Centre Interuniversitaire sur le Risque,
les Politiques Économiques et l'Emploi

Cahier de recherche/Working Paper **13-36**

On Learning and Growth

Leonard J. Mirman
Kevin Reffet
Marc Santugini

Octobre/October 2013
Version revisée/revised : Octobre/October 2015

Mirman: Department of Economics, University of Virginia, USA
lm8h@virginia.edu
Reffett: Department of Economics, Arizona State University, USA
kevin.reffett@asu.edu
Santugini: Department of Applied Economics and CIRPÉE, HEC Montréal, Canada
marc.santugini@hec.ca

Abstract:
We study optimal growth under learning. We extend the Mirman-Zilcha stochastic growth results characterizing optimal programs for general utility and production functions to the case of learning. We then use recursive methods to study the effect of learning on the dynamic program by considering the case of iso-elastic utility and linear production, for general distributions of the random shocks and beliefs (i.e., without the use of conjugate priors), for any horizon. Finally, we address the issue of experimentation by providing a solution to an infinite-horizon optimal dynamic program.

Keywords: Brock-Mirman environment, Dynamic programming, Euler equation, Experimentation, Learning, Optimal growth

JEL Classification: D8, D9, E2

1 Introduction

Economic agents make optimal decisions without complete knowledge of their environment. This is particularly relevant to dynamic problems. To understand behavior under uncertainty in a dynamic model, random shocks were introduced as part of the future outcomes.[1] In the optimal growth model under uncertainty studied in Brock and Mirman (1972) and Mirman and Zilcha (1975), it is future output that is assumed to be unknown. Specifically, the agent has no knowledge of the realizations of future shocks, which affects future output. Although the realization of future shocks are unknown in stochastic optimal growth models, the distribution of these shocks is assumed to be *known*. In other words, the structure of the economy is known. The agent uses this knowledge in making optimal decisions to form rational expectations over future outcomes.

When the structure of the production process is unknown (e.g., a parameter of the production process or the distribution function generating the random shock is unknown), the ability to learn arises since the sequence of outputs yields information about the structure of the economy. The possibility of learning adds a layer of complexity to the dynamic maximization problem, well beyond that of adding more uncertainty. Indeed, a learning agent does more than just react to uncertainty, he also modifies the uncertainty faced in future decision-making. For instance, suppose that in addition to not knowing the realization of future output, the agent does not know but has beliefs about the distribution generating the production shocks. In that case, observing past and present levels of output provides information about the unknown distribution, which allows the agent to update beliefs about the structure of the economy as new information is observed. Hence, in a learning environment, the agent makes optimal consumption and investment decisions using prior beliefs to form rational expectations over future shocks as well as their effects on *both* future output and future beliefs. In particular, the anticipation of beliefs updating prevents any separation between

[1] See Mirman (1970) for an early analysis of uncertainty (i.e., a random production function) in the Solow model.

decision-making and learning and, thus, has a profound effect on optimal behavior.

Unlike the literature on stochastic optimal growth, there is little work on optimal growth in which agents learn about the structure of the economy. Recently, Koulovatianos, Mirman, and Santugini (2009) (KMS) studied the effect of learning on optimal consumption in the Mirman and Zilcha (1975) (MZ) growth model (log utility and Cobb-Douglas production with exponential uncertainty). The implication of the KMS analysis is that learning involves more than merely replacing the objective distribution of future outcomes by the subjective distribution of future outcomes. Indeed, learning has a deep effect on the dynamic maximization problem. Although KMS characterizes optimal behavior for general distributions of shocks and beliefs and studies the effect of learning on optimal behavior, it does not capture the entire effect of learning on the dynamic program. A characterization of optimal growth using the Euler equation approach, yielding general results, remains to be done.

Moreover, little is known about the link between learning and dynamic programming. Indeed, in KMS, the value function is conjectured and verified. This prevents a complete analysis of the effect of learning on the dynamic maximization problem. There are several important aspects of learning that are hidden in the analysis of KMS. In a general model of information and learning, we need to understand how information is generated, i.e., what signal is observed and what information is conveyed by the signal. The next issue is how the information is extracted in order to learn about the unknown parameter. Finally, we must consider what power the agent has to influence the amount of information in the signal. More precisely, is it possible to experiment? If so, is it optimal to experiment or is passive learning optimal? It is the object of this paper to study these links between learning and the dynamic maximization problem.

We consider a Brock-Mirman (BM) growth model in which the structure of the economy is unknown, in the sense that a parameter of the distribution of future output is unknown. Hence, not only are the future outcomes random (as in BM), but the underlying structure of the economy generating these random outcomes is unknown.

First, we focus on the case of learning without experimentation. That is, the agent has no opportunity to generate more information by altering consumption.[2] Using Bayesian methods to update beliefs, we embed the learning process in the dynamic program and identify the beliefs component and the updating component of learning in the value function. We then extend the Mirman-Zilcha results of stochastic optimal growth to learning without experimentation. Namely, the value function is strictly concave and differentiable. The solution to the optimal dynamic program is unique and completely characterized by the Euler equation. In other words, the general model of learning without experimentation can be characterized using the Euler equation in the same way that the BM stochastic growth model can be characterized, so that it is still the case that the Euler equation (along with the transversality condition) are both necessary and sufficient conditions for the optimal policy function. On the other hand, the Euler equation does not make apparent the effect of learning on the dynamic maximization problem as it hides the intricacies of learning.

To understand the influence of learning on the dynamic maximization problem, we specialize the model to the case of iso-elastic utility and linear production with multiplicative uncertainty. In order to clarify the effect of learning, we use the recursive approach by solving for the optimal policy function for any finite horizon, beginning with the one-period horizon and extending the program to the infinite horizon. For any horizon, optimal behavior is characterized for general distributions of the random shock and beliefs (i.e., there is no use of conjugate priors). The solution is unique, linear in the output and the optimal fraction of output consumed depends on beliefs. The recursive procedure allows us to see how learning and especially the updating component alters the structure of future payoffs. We show that the beliefs component of learning is present in the one-period-horizon

[2] In that case, changing the decision has no influence on the amount of information generated, i.e., experimentation is not possible. However, learning without experimentation has a significant effect on optimal behavior because optimal growth models are signal-dependent problems. A dynamic maximization problem is *signal-dependent* when the signal has an effect on future payoffs other than through posterior beliefs.

dynamic maximization problem whereas the updating component comes into play when the horizon is equal to two periods. It is the presence of the beliefs components in the one-period horizon that leads to the dependence of the one-period-horizon value function on prior beliefs. This dependence means that the anticipation of beliefs updating begins to be relevant at the two-period-horizon maximization problem. Both components of learning remain present in maximization problems with a horizon longer than two periods.

We then apply the recursive approach to the MZ growth model studied in KMS.[3] The MZ growth model with log utility and Cobb-Douglas production functions offers a preliminary insight of the effect of learning in a general optimal growth model. Although learning (compared to the stochastic case) has an effect in each model, we show that there are some major differences between the growth model studied in this paper and the MZ model.

Finally, we consider the case of learning with experimentation. Optimal growth in the case of learning with experimentation is a much more complicated dynamic problem. In fact, in the literature, there are no examples in an infinite-period horizon.[4] We first present the BM set up under learning with experimentation. Instead of an unknown distribution of the production shock, we consider an unknown parameter influencing the production function. This leads to the opportunity for experimentation. We then provide an example of an infinite-horizon growth model in which it is optimal to experiment. Specifically, in the case of linear utility and Cobb-Douglas production function, we derive the optimal policy function when the production shock is uniformly distributed. The agent, in this model, is able to alter consumption in order to increase information, i.e., to experiment.

The paper is organized as follows. Sections 2, 3, and 4 consider the case of learning without experimentation. Specifically, in section 2, we describe the growth model under learning and we characterize optimal behavior by deriving the Euler equation. Section 3 studies the effect of learning on the dynamic program using the recursive approach. In section 4, we discuss the

[3]There is no recursive analysis in KMS.

[4]Experimentation has been studied in signal-dependent dynamic models with at most a three-period horizon (Bertocchi and Spagat, 1998; Datta et al., 2002; El-Gamal and Sundaram, 1993; Huffman and Kiefer, 1994).

differences about the effect of learning on the dynamic maximization problem between our model (with iso-elastic utility and linear production) and the MZ model (with log utility and Cobb-Douglas production). Section 5 addresses the issue of learning with experimentation in optimal growth. We provide concluding remarks in Section 6.

2 Model

In this section, we present the general problem of optimal growth with learning. We first recall the Brock-Mirman (BM) framework in the stochastic case, i.e., without learning. We then consider the learning case by presenting the learning process and the Bellman equation. We also derive properties of the value function as well as the optimal policy function in the learning case. In subsequent sections, we use a recursive approach to clarify the effect of learning on optimal behavior, i.e., we highlight the influence of learning on the maximization problem.

2.1 Preliminaries

Consider a BM framework in which, in period $t = 0, 1, \ldots$, an agent divides output y_t between consumption $c_t \in [0, y_t]$ and investment $k_t = y_t - c_t$. Investment k_t is used for the production of output in period $t + 1$, i.e.,

$$y_{t+1} = f(y_t - c_t, r_t) \tag{1}$$

where $r_t \in (\underline{r}, \overline{r})$ is a production shock, $0 \leq \underline{r} < \overline{r}$. The production function $f(k_t, r_t)$ has the usual neoclassical properties. The Inada conditions are also assumed.

Assumption 2.1. *For all $r_t \in (\underline{r}, \overline{r})$,*

1. *For all $k_t \geq 0$, $f_1(k_t, r_t) \geq 0, f_{11}(k_t, r_t) \leq 0$.*

2. *For $k_t = 0$, $f(0, r_t) = 0$, and for $k_t > 0$, $f(k_t, r_t) > 0$.*

3. *$f_1(0, r_t) = \infty$ and $f_1(\infty, r_t) = 0$.*

The production shock r_t is a realization of the random variable $\tilde{r}_t$ with p.d.f. $\phi(r_t|\theta^*)$, i.e., the p.d.f. of $\tilde{r}_t$ depends on a parameter $\theta^* \in \Theta \subset \mathbb{R}^N, N \in \mathbb{N}$. The distribution of $\tilde{r}_t$ is parametric and fully characterized by the vector θ^*. Moreover, $\{\tilde{r}_t\}_{t=0}^{\infty}$ are independently and identically distributed (i.i.d.). The objective is to maximize the expected discounted sum of utilities subject to (1). The utility function is endowed with the following properties.

Assumption 2.2. *For all $c_t \geq 0$, $u'(c_t) > 0, u''(c_t) < 0$. Moreover, $u'(0) = \infty$.*

Brock and Mirman (1972), and Mirman and Zilcha (1975) study optimal growth in the *stochastic* case in which the agent faces uncertainty about the realizations of future production shocks, but knows the true distribution of $\{\tilde{r}_t\}_{t=0}^{\infty}$, i.e., θ^* is known. At time t, given y_t, the infinite-horizon value function is

$$V^S(y_t) = \max_{c_t \in [0, y_t]} \left\{ u(c_t) + \delta \int_{\underline{r}}^{\overline{r}} V^S(f(y_t - c_t, r_t)) \ \phi(r_t|\theta^*) \ \mathrm{d}r_t \right\} \tag{2}$$

where the discount factor is $\delta \in (0, 1)$. Here, the superscript S refers to the stochastic case.

Mirman and Zilcha (1975) shows that the value function is differentiable and derives necessary and sufficient conditions for the optimal policy function. Specifically, the maximum is obtained at a unique point $c_t = \rho^S(y_t)$ since the maximand is strictly concave. Moreover, $\rho^S(y_t) \in (0, y_t)$ since the following conditions hold.

Conditions 2.3. *[Interior Solutions]*

1. *There is at least one future period.*

2. *For all $r_t \in (\underline{r}, \overline{r})$, $f(0, r_t) = 0$.*

3. *$u'(0) = \infty$.*

Hence, for the infinite horizon, a maximizing program from any initial point never exhausts the stock and is therefore infinite. Finally, $\rho^S(y_t)$ is increasing in y_t and satisfies the Euler equation

$$u'(\rho^S(y_t)) = \delta \int_{\underline{r}}^{\overline{r}} f_1(y_t - \rho^S(y_t), r_t)\ u'(\rho^S(y^S_{t+1}(r_t)))\ \phi(r_t|\theta^*)\ \mathrm{d}r_t, \tag{3}$$

$y^S_{t+1}(r_t) \equiv f(y_t - \rho^S(y_t), r_t)$, as well as the transversality condition.

The Euler equation provides a complete characterization of the optimal policy function. To illustrate the Euler approach, consider the stochastic growth models for which the optimal policies are linear functions of the output, i.e., $\rho^S(y_t) = \omega y_t$ where $\omega \in (0, 1)$ is the constant optimal fraction of output consumed. First, suppose that utility is logarithmic and production is Cobb-Douglas with a power shock as studied in Mirman and Zilcha (1975). Plugging $u(c_t) = \ln c_t$, $f(y_t - c_t, r_t) = (y_t - c_t)^{r_t}, r_t \in (0, 1)$, and $c_t = \omega y_t$ into (3) yields[5]

$$\omega = 1 - \delta \int_0^1 r_t\ \phi(r_t|\theta^*)\ \mathrm{d}r_t. \tag{4}$$

From (4), uncertainty affects the optimal fraction of output consumed only through the mean of the production shock, i.e., certainty equivalence prevails.

Next, from Example 1 in Mirman and Zilcha (1977), the utility is isoelastic and production is linear with a multiplicative shock. Then, using the Euler equation, the optimal fraction of output consumed is[6]

$$\omega = 1 - \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r_t^\alpha\ \phi(r_t|\theta^*)\ \mathrm{d}r_t \right)^{\frac{1}{1-\alpha}}. \tag{5}$$

From (5), this example does not display certainty equivalence as in (4), i.e., higher moments of $\tilde{r}_t$ influence optimal behavior.

Having recalled the stochastic case, we present next optimal growth with learning. We show that, as in the stochastic case, the Euler equation corresponding to the learning case provides a complete characterization of the optimal policy function.

[5]From (3), the Euler equation $\frac{1}{\omega y_t} = \delta \int_0^1 \frac{r_t (y_t - \omega y_t)^{r_t - 1}}{\omega (y_t - \omega y_t)^{r_t}}\ \phi(r_t|\theta^*)\ \mathrm{d}r_t$ yields (4).

[6]When $u(c_t) = c_t^\alpha$, $f(y_t - c_t, r_t) = r_t(y_t - c_t), r_t \in (0, 1)$ and $c_t = \omega y_t$, expression (3) is rewritten as $\alpha \omega^{\alpha-1} y_t^{\alpha-1} = \delta \int_0^1 r_t^\alpha \alpha \omega^{\alpha-1} (y_t - \omega y_t)^{\alpha-1}\ \phi(r_t|\theta^*)\ \mathrm{d}r_t$, which yields (5).

2.2 Learning Process

In the *learning* case, the agent faces uncertainty about the distribution of future levels of output. We consider the case in which the agent does not know the distribution of future production shocks, i.e., θ^* is unknown and is thus treated as a random variable by the agent. Although the agent is uninformed about the distribution of future shocks (and thus the distribution of future output), observing past and present levels of output provides information about the value of the unknown parameter. This information is used to update beliefs via Bayesian methods.

To embed learning in a dynamic program, we first discuss the learning process. Let $\xi_t(\theta), \theta \in \Theta$ be the beliefs about θ^* at time t.[7] At time $t+1$, given y_t and c_t, the agent observes the signal y_{t+1}. The beliefs at time $t+1$ are then updated by Bayes' theorem to

$$\xi_{t+1}(\theta|y_{t+1}, y_t - c_t) = \frac{\phi_Y(y_{t+1}|y_t - c_t, \theta)\xi_t(\theta)}{\int_{\theta' \in \Theta} \phi_Y(y_{t+1}|y_t - c_t, \theta')\xi_t(\theta')\mathrm{d}\theta'}, \quad \theta \in \Theta \qquad (6)$$

where $\phi_Y(y_{t+1}|y_t - c_t, \theta)$ is the p.d.f. of $\tilde{y}_{t+1}$ conditional on θ. At time t, ξ_t and $\xi_{t+1}(\cdot|y_{t+1}, y_t - c_t)$ are referred to as prior and posterior beliefs, respectively.[8]

If the agent has no opportunity to experiment (i.e., no opportunity to alter consumption in order to generate more information), then expression (6) may be simplified further.

Assumption 2.4. *For all $k_t \geq 0$ and $r_t \in (\underline{r}, \overline{r})$, $f_2(k_t, r_t) \neq 0$.*

Assumption 2.4 removes the opportunity for experimentation, which allows us to focus on the case of learning without experimentation, often referred to as *passive learning*.[9] Indeed, since $f_2(k_t, r_t) \neq 0$, it follows that r_t can be inferred exactly from observing y_{t+1}.[10] Moreover, the signal $\tilde{r}_t$ is sufficient for the signal $\tilde{y}_{t+1}$, i.e., all information about θ^* contained in $\tilde{y}_{t+1}$ is

[7] That is, for any $S \subset \Theta$, the probability that $\theta^* \in S$ at time t is $\int_{\theta \in S} \xi_t(\theta)\mathrm{d}\theta$.

[8] In order to simplify notation, from the point of view of time t, we do not make explicit the dependence of ξ_t on past and present signals $\{y_s\}_{s=0}^t$.

[9] Learning with experimentation adds another layer of complexity in the maximization problem, which we address in Section 5.

[10] The assumption $f_2(k_t, r_t) \neq 0$ implies either $f_2(k_t, r_t) > 0$ or $f_2(k_t, r_t) < 0$.

summarized in $\tilde{r}_t$. That is, regardless of the consumption decision, the flow of information about θ^* remains the same since it emanates only from the random shock.

Formally, using the fact that $y_{t+1} = f(y_t - c_t, r_t)$ is a function of a continuous random variable $r_t \in (\underline{r}, \overline{r})$, the p.d.f. of the signal $y_{t+1} \in (f(y_t - c_t, \underline{r}), f(y_t - c_t, \overline{r}))$ is given by

$$\phi_Y(y_{t+1}|y_t - c_t, \theta) = \phi(r_t|\theta)|f_2(y_t - c_t, r_t)|^{-1} \tag{7}$$

where $\phi(r_t|\theta)$ is the p.d.f. of $\tilde{r}_t$ conditional on θ and $|f_2(y_t - c_t, r_t)|^{-1}$ is well-defined. Moreover, $f_2(y_t - c_t, r_t)$ is independent of θ, and thus, the information about the unknown parameter is embodied by only the p.d.f. of $\tilde{r}_t$.

Plugging (7) into (6) yields

$$\xi_{t+1}(\theta|y_{t+1}, y_t - c_t) = \frac{\phi(r_t|\theta)|f_2(y_t - c_t, r_t)|^{-1}\xi_t(\theta)}{\int_{\theta' \in \Theta} \phi(r_t|\theta')|f_2(y_t - c_t, r_t)|^{-1}\xi_t(\theta')\mathrm{d}\theta'}, \tag{8}$$

$$= \frac{\phi(r_t|\theta)\xi_t(\theta)}{\int_{\theta' \in \Theta} \phi(r_t|\theta')\xi_t(\theta')\mathrm{d}\theta'}, \tag{9}$$

$\theta \in \Theta$. From (9), the consumption decision has no effect on posterior beliefs because the distribution of $\tilde{r}_t$ is independent of c_t. When learning about θ^*, the agent cannot influence future beliefs through his present consumption decision.

Remark 2.5. *Given Assumption 2.4, beliefs at time $t+1$ are independent of the consumption decision at time t. That is, for $\theta \in \Theta$, $\xi_{t+1}(\theta|y_{t+1}, y_t - c_t) = \xi_{t+1}(\theta|r_t)$.*

2.3 Control and Learning

Having described the learning process, we study next the agent's consumption decisions while learning about θ^*. That is, endowed with initial stock and prior beliefs, consumption c_t is chosen. The production shock r_t is then realized and the output, in the subsequent period, is determined from (1).

Information is gleaned from observing the signal y_{t+1} and thus r_t, which, from (9), affects beliefs about θ^*. In the learning case, the infinite-horizon value function is

$$V^L(y_t, \xi_t) = \max_{c_t \in (0, y_t)} \left\{ u(c_t) + \delta \int_{\underline{r}}^{\overline{r}} V^L(f(y_t - c_t, r_t), \xi_{t+1}(\cdot | r_t)) \left[\int_{\theta \in \Theta} \phi(r_t|\theta)\xi_t(\theta)\mathrm{d}\theta \right] \mathrm{d}r_t \right\} \tag{10}$$

where the superscript L refers to the learning case. Here, $\int_{\theta \in \Theta} \phi(r_t|\theta)\xi_t(\theta)\mathrm{d}\theta$ is the expected p.d.f. of the production shock given beliefs at time t.

A few comments about (10) are warranted. First, optimal growth with learning falls into the category of signal-dependent problems. That is, the signal y_{t+1} (or the sufficient signal r_t) has a direct effect on future payoffs, which is different from its effect on posterior beliefs. Specifically, the consumption choice set in the subsequent period depends on the signal y_{t+1}, i.e., $c_{t+1} \in (0, y_{t+1})$.[11]

Second, there are two components of Bayesian learning embedded in the value function. A prior beliefs due to structural uncertainty and an updating rule to go from prior to posterior beliefs. More precisely, the effect of learning on optimal behavior depends on a beliefs component and an updating component.[12] From (10), the beliefs component is captured by the expected p.d.f. $\int_{\theta \in \Theta} \phi(r_t|\theta)\xi_t(\theta)\mathrm{d}\theta$ and the updating component is due to the dependence of the continuation value function on the posterior beliefs, i.e., $\xi_{t+1}(\cdot|r_t)$.

Third, the agent anticipates the effect of the production shock on future output as well as posterior beliefs.[13] In a dynamic and learning context, rational expectations imply that the information contained in the future production shock is anticipated. The anticipation of the acquisition of information is integrated into (10) through the posterior $\xi_{t+1}(\cdot|r_t)$. The dual role of

[11] In other words, the consumption choice set in the subsequent period depends on the signal r_t, i.e., $c_{t+1} \in (0, f(y_t - c_t, r_t))$.

[12] In KMS, the updating component is called the anticipation component. The label *updating* is more precise.

[13] Even with i.i.d. shocks in production, a learning environment implies that the agent faces Markov shocks through the updating of the beliefs. See Hopenhayn and Prescott (1992) and Mirman et al. (2008) for stochastic growth models with Markov processes.

the production shock on future payoffs via future output and posterior beliefs implies that, in general, there is no separation between decision-making and learning. In particular, the agent's decisions are influenced by the learning process, i.e., the anticipation of beliefs updating changes the uncertainty faced by the agent, which alters decision-making.

Fourth, Bayesian dynamics complicates the maximization problem because the evolution of beliefs must be taken into account in the optimal dynamic program. That is, the continuation value function encompasses beliefs that are updated *infinitely* many times.

2.4 Euler Equation

The MZ results stated for the stochastic case can be extended to the learning case. Indeed, the maximum is obtained at a unique point $c_t = \rho^L(y_t, \xi_t)$ since the maximand is strictly concave. Moreover, Conditions 2.3 are satisfied, and thus, there cannot be any corner solutions, i.e., $\rho^L(y_t, \xi_t) \in (0, y_t)$. Lemma 2.6 states that the value function is differentiable and is equal to the marginal utility evaluated at the maximizer, i.e., the envelope theorem. This result is then used in Proposition 2.7 to derive the Euler equation under learning.

Lemma 2.6. *For all y_t and ξ_t, $\partial V^L(y_t, \xi_t)/\partial y_t$ exists such that*

$$\frac{\partial V^L(y_t, \xi_t)}{\partial y_t} = u'(\rho^L(y_t, \xi_t)). \tag{11}$$

Proof. See Appendix A. □

Theorem 1 in Mirman and Zilcha (1975) applies also to the learning environment. That is, necessary and sufficient conditions that $\rho^L(y_t, \xi_t)$ be the optimal policy function are the Euler equation and the transversality condition. Proposition 2.7 states the Euler equation in the learning case.

Proposition 2.7. *For all y_t and ξ_t, $\rho^L(y_t, \xi_t)$ satisfies the Euler equation*

$$u'(\rho^L(y_t, \xi_t)) = \delta \int_{\underline{r}}^{\overline{r}} f_1(y_t - \rho^L(y_t, \xi_t), r_t)\, u'(\rho^L(y^L_{t+1}(r_t), \xi_{t+1}(\cdot|r_t))) \left[\int_{\theta\in\Theta} \phi(r_t|\theta)\xi_t(\theta)\mathrm{d}\theta\right] \mathrm{d}r_t \tag{12}$$

where $y^L_{t+1}(r_t) \equiv f(y_t - \rho^L(y_t, \xi_t), r_t)$.

Proof. The first-order condition corresponding to (10) is

$$u'(c_t) = \delta \int_{\underline{r}}^{\overline{r}} f_1(y_t - c_t, r_t) \frac{\partial V^L(y_{t+1}, \xi_{t+1}(\cdot|r_t))}{\partial y_{t+1}} \left[\int_{\theta\in\Theta} \phi(r_t|\theta)\xi_t(\theta)\mathrm{d}\theta \right] \mathrm{d}r_t \tag{13}$$

evaluated at $c_t = \rho^L(y_t, \xi_t)$. Since (11) holds for all y_t and ξ_t, plugging (11) (updated to the next period) into (13) yields (12). □

From (12), the Euler equation provides a complete characterization of the optimal policy function in the learning case. Hence, the Euler equation allows us to compare the optimal policy functions between the stochastic and learning cases. Specifically, observe that the differences between the Euler equation for the stochastic case (defined by (3)) and the Euler equation under learning (defined by (12)) are subtle but important. In particular, learning does more than change the distribution of the production shock from an objective distribution used in (3) to a subjective distribution used in (12).[14] It also alters the marginal utility evaluated at next period's consumption through the uncertainty of future beliefs, i.e., the term $u'(\rho^L(y^L_{t+1}(r_t), \xi_{t+1}(\cdot|r_t)))$ in (12). Indeed, the anticipation of beliefs updating implies that the dynamics in output and beliefs are entwined through the production shock as shown in (12).

As in the stochastic case, the Euler approach can be used to solve for optimal policies. To see this, we embed learning in the two growth models previously presented in the stochastic case. In the learning case, the optimal policy is a linear function of the output. However, unlike the stochastic case, the optimal fraction of output consumed in the learning case is not a constant, but depends on beliefs. Specifically, the optimal policy function has the form $\rho^L(y_t, \xi_t) = \omega(\xi_t)y_t$ where the optimal fraction of output consumed is a function of prior beliefs.

For the MZ growth model, plugging $u(c_t) = \ln c_t$, $f(y_t - c_t, r_t) = (y_t -$

[14] That is, learning does more than replacing the p.d.f. $\phi(r_t|\theta^*)$ by the expected p.d.f. $\int_{\theta\in\Theta} \phi(r_t|\theta)\xi_t(\theta)\mathrm{d}\theta$.

$c_t)^{r_t}, r_t \in (0,1)$, and $\rho^L(y_t, \xi_t) = \omega(\xi_t) y_t$ into (12) yields

$$\frac{1}{\omega(\xi_t) y_t} = \delta \int_0^1 \frac{r_t (y_t - \omega(\xi_t) y_t)^{r_t - 1}}{\omega(\xi_{t+1}(\cdot|r_t)) y^L_{t+1}(r_t)} \left[\int_{\theta \in \Theta} \phi(r_t|\theta) \xi_t(\theta) \mathrm{d}\theta \right] \mathrm{d}r_t, \tag{14}$$

$y^L_{t+1}(r_t) = (y_t - \omega(\xi_t) y_t)^{r_t}$, which simplifies to

$$\frac{1 - \omega(\xi_t)}{\omega(\xi_t)} = \delta \int_0^1 \frac{r_t}{\omega(\xi_{t+1}(\cdot|r_t))} \left[\int_{\theta \in \Theta} \phi(r_t|\theta) \xi_t(\theta) \mathrm{d}\theta \right] \mathrm{d}r_t. \tag{15}$$

By conjecturing and verifying the value function, KMS show that

$$\omega(\xi_t) = \left(\int_{\theta \in \Theta} \frac{\xi_t(\theta) \mathrm{d}\theta}{1 - \delta \int_0^1 r_t \phi(r_t|\theta) \mathrm{d}r_t} \right)^{-1}. \tag{16}$$

Note that from (4) and (16), the inclusion of learning removes the certainty equivalence property.

Next, suppose that $u(c_t) = c_t^\alpha, \alpha \in (0,1)$ and $f(y_t - c_t, r_t) = r_t(y_t - c_t), r_t \in (0,1)$. Conjecturing that $\rho^L(y_t, \xi_t) = \omega(\xi_t) y_t$, the Euler equation is rewritten as

$$(\omega(\xi_t))^{\alpha - 1} y_t^{\alpha - 1} = \delta \int_0^1 r_t \left(\omega(\xi_{t+1}(\cdot|r_t)) \right)^{\alpha - 1} (y^L_{t+1}(r_t))^{\alpha - 1} \left[\int_{\theta \in \Theta} \phi(r_t|\theta) \xi_t(\theta) \mathrm{d}\theta \right] \mathrm{d}r_t, \tag{17}$$

$y^L_{t+1}(r_t) = r_t(y_t - \omega(\xi_t) y_t)$, which yields an implicit characterization of the function $\omega(\cdot)$, i.e.,

$$\frac{1 - \omega(\xi_t)}{\omega(\xi_t)} = \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 \frac{r_t^\alpha}{\left(\omega(\xi_{t+1}(\cdot|r_t)) \right)^{1-\alpha}} \left[\int_{\theta \in \Theta} \phi(r_t|\theta) \xi(\theta) \mathrm{d}\theta \right] \mathrm{d}r_t \right)^{\frac{1}{1-\alpha}}. \tag{18}$$

In the next section, we show that $\omega(\xi_t)$ exists and is unique.

Although (12) characterizes optimal behavior under learning, the Euler equation hides the intricacies of learning and prevents a thorough analysis of the effect of learning on the maximization problem. For instance, comparing (15) and (18) provides no clear insights about how learning alters the dynamic program. In the subsequent sections, we show explicitly the influence of learning. First, we focus on the case of iso-elastic utility and linear production using a recursive analysis. Then, we compare optimal behavior between the MZ model and the case of iso-elastic utility and linear production.

3 Recursive Analysis

In this section, we use the recursive approach to shed light on optimal behavior in a learning environment. Specifically, we show how learning works to change the optimal policy, i.e., the influence of learning on the maximization problem. To simplify notation, the t-subscript for indexing time is removed and the hat sign is used to indicate the value of a variable in the subsequent period, i.e., y is output today and

$$\hat{y} = f(y - c, r) \tag{19}$$

is output in the next period. Moreover, prior beliefs are denoted as ξ and, from (9), posterior beliefs are

$$\hat{\xi}(\theta|r) = \frac{\phi(r|\theta)\xi(\theta)}{\int_{\theta' \in \Theta} \phi(r|\theta')\xi(\theta')\mathrm{d}\theta'}, \quad \theta \in \Theta. \tag{20}$$

To distinguish among different horizons of the dynamic program, we use the index $\tau = 0, 1, \ldots$. Hence, in the learning case, for $\tau = 1, 2, \ldots$, the τ-period-horizon value function is

$$V_\tau^L(y, \xi) = \max_{c \in (0,y)} \left\{ u(c) + \delta \int_0^\infty V_{\tau-1}^L(f(y-c, r), \hat{\xi}(\cdot|r)) \left[\int_{\theta \in \Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\} \tag{21}$$

where $c \in (0, y)$ since Conditions 2.3 are satisfied.

We make further assumptions on the utility and production functions, but we retain *general distributions* of the production shock and prior beliefs. In particular, we make no restriction on the evolution of beliefs and we do not prevent the prior and posterior p.d.f.'s ξ and $\hat{\xi}(\cdot|r)$ from belonging to different families. Specifically, we characterize and analyze optimal behavior in the class of optimal stochastic growth models with an iso-elastic utility function and a linear production function under multiplicative uncertainty. It turns out that there exists a unique optimal solution that is linear, but cannot be made explicit. The solution depends upon the prior and is valid for all distributions even those that are outside of families that are closed under sampling.

Assumption 3.1. *The utility function is iso-elastic:* $u(c) = c^\alpha, \alpha \in (0, 1)$.

Assumption 3.2. *The production function is linear:* $f(k, r) = rk, r \in (0, 1)$.

Given Assumptions 3.1 and 3.2, (21) is rewritten as

$$V_\tau^L(y, \xi) = \max_{c \in (0,y)} \left\{ c^\alpha + \delta \int_0^1 V_{\tau-1}^L(r(y-c), \hat{\xi}(\cdot|r)) \left[\int_{\theta \in \Theta} \phi(r|\theta)\xi(\theta) \mathrm{d}\theta \right] \mathrm{d}r \right\} \tag{22}$$

and $\rho_\tau^L(y, \xi)$ is the corresponding optimal policy. Using (22), the recursive analysis has two parts. First, we consider the one-period and two-period horizons to show how learning alters the maximization problem. Second, we characterize optimal behavior for any finite horizon and provide the infinite-horizon optimal solution by showing that the limit of the finite-horizon optimal solution (as the horizon tends to infinity) exists.

3.1 One- and Two-Period Horizons

To show how learning alters the maximization problem, we first solve the optimal one-period-horizon dynamic program, which is used to solve the optimal two-period-horizon dynamic program.

Note that when there is no horizon (in the last period), it is optimal for the agent to consume the entire output and thus the value function does not depend on beliefs, i.e., $\rho_0^L(y, \xi) = y$ and $V_0^L(y, \xi) = y^\alpha$. Hence, there is no effect of learning in the no-horizon case, i.e., both the beliefs and updating components of learning are absent.

One-Period Horizon. Using (22) and the fact that $V_0^L(y, \xi) = y^\alpha$, the one-period-horizon value function is

$$V_1^L(y,\xi) = \max_{c\in(0,y)} \left\{ c^\alpha + \delta \int_0^1 V_0^L(r(y-c),\hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \tag{23}$$

$$= \max_{c\in(0,y)} \left\{ c^\alpha + \delta \left(\int_0^1 r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right) (y-c)^\alpha \right\}. \tag{24}$$

In a one-period-horizon value function, the effect of learning takes place only through the beliefs component, i.e., prior beliefs are used only to anticipate the future shock through the *expected* p.d.f. of $\tilde{r}$, i.e., $\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta$. Indeed, when there is only one remaining period, the agent anticipates consuming the entire output in the next period, which makes beliefs in the next period irrelevant. It follows that, for a one-period horizon, there is no need to anticipate the updating of beliefs, i.e., there is no updating component of learning.

The first-order condition $c^{\alpha-1} = \delta \left(\int_0^1 r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right) (y-c)^{\alpha-1}$ yields

$$\rho_1^L(y,\xi) = \frac{y}{1+\delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right)^{\frac{1}{1-\alpha}}}. \tag{25}$$

Plugging (25) into (24) yields

$$V_1^L(y,\xi) = \left(1+\delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right)^{\frac{1}{1-\alpha}} \right)^{1-\alpha} y^\alpha. \tag{26}$$

Two-Period Horizon. Using (26), the two-period-horizon value function is

$$
\begin{aligned}
&V_2^L(y,\xi) \\
&= \max_{c\in(0,y)} \left\{ c^\alpha + \delta \int_0^1 V_1^L(r(y-c), \hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \qquad (27) \\
&= \max_{c\in(0,y)} \Bigg\{ c^\alpha \\
&\quad + \delta \int_0^1 \left[\left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_{r'\in(0,1)} r'^\alpha \left[\int_{\theta'\in\Theta} \phi(r'|\theta')\hat{\xi}(\theta'|r)\mathrm{d}\theta' \right] \mathrm{d}r' \right)^{\frac{1}{1-\alpha}} \right)^{1-\alpha} r^\alpha (y-c)^\alpha \right] \\
&\quad \cdot \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r, \Bigg\} \qquad (28)
\end{aligned}
$$

where $\hat{\xi}(\theta'|r)$ is defined by (20). Expression (28) involves multiple integrals for the production shock. The outer integral with dummy r reflects the uncertainty faced by the agent today (i.e., in period 1) about today's production shock, the value of which is revealed next period (i.e., in period 2). The uncertainty emanating from today's yet-to-be-realized production shock has an effect on the stock in the next period (through the term $r^\alpha(y-c)^\alpha$) and on posterior beliefs (through the term $\hat{\xi}(\theta|r)$). The effect through posterior beliefs complicates the maximization problem because updating beliefs has an effect on the inner integral with dummy r' that refers to the expectation that the agent takes in the next period (i.e., in period 2) for the next-period production shock, the value of which is revealed in the last period (i.e., in period 3). The production shock in the next period has an effect on the expected distribution of output in the last period.

To understand further the effect of learning in the two-period-horizon dynamic program, we rewrite the continuation value function defined in (26). Using (25), $\rho_1^L(y,\xi) = \omega_1(\xi)y$ where

$$
\omega_1(\xi) = \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right)^{\frac{1}{1-\alpha}} \right)^{-1} \qquad (29)
$$

is the optimal one-period-horizon fraction of output consumed. Plugging

$\rho_1^L(y,\xi) = \omega_1(\xi)y$ into (24) yields

$$V_1^L(y,\xi) = (\omega_1(\xi))^{\alpha-1}\, y^\alpha. \tag{30}$$

From (30), prior beliefs enter the one-period-horizon value function through the optimal fraction of output consumed.

Now, consider (30) instead of (26) so that the two-period-horizon value function is rewritten as

$$V_2^L(y,\xi) = \max_{c\in(0,y)} \left\{ c^\alpha + \delta \int_0^1 V_1^L(r(y-c), \hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \tag{31}$$

$$= \max_{c\in(0,y)} \left\{ c^\alpha + \delta \int_0^1 \left(\omega_1(\hat{\xi}(\cdot|r)) \right)^{\alpha-1} r^\alpha (y-c)^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\} \tag{32}$$

where, using (29),

$$\omega_1(\hat{\xi}(\cdot|r)) = \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_{r'\in(0,1)} r'^\alpha \left[\int_{\theta'\in\Theta} \phi(r'|\theta')\hat{\xi}(\theta'|r)\mathrm{d}\theta' \right] \mathrm{d}r' \right)^{\frac{1}{1-\alpha}} \right)^{-1} \tag{33}$$

is the optimal one-period-horizon fraction of output consumed as a function of posterior beliefs, which is a random variable from the vantage point of the present period.

Expressions (32) and (33) provide a general picture of how both the beliefs component and the updating component affect the two-period-horizon dynamic program. First, not knowing the true distribution of the random production shock implies that the agent uses the prior beliefs ξ to form an expectation of the distribution of the random production shock. As in the one-period-horizon maximization problem, the beliefs component is reflected by the expected p.d.f. of $\tilde{r}$ given prior beliefs, $\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta$ in (32).

Next, learning about the true distribution of the random production shocks means that the agent anticipates updating prior beliefs to the posterior $\hat{\xi}(\cdot|r)$ in order to *update* the expectation of future output and, thus, future optimal payoffs. Posterior beliefs is a random variable from today's perspective that depends upon the shock. Hence, the continuation value function is also a random variable that depends upon posterior beliefs, i.e., the random continuation value function embedded in (32) is $V_1^L(\hat{y}, \hat{\xi}(\cdot|r)) = \left(\omega_1(\hat{\xi}(\cdot|r))\right)^{\alpha-1} \hat{y}^\alpha$, which is derived from (30). Note that the agent anticipates updating beliefs, which affects the expectation of the mean of $\tilde{r}^\alpha$, i.e.,

$$\int_{r'\in(0,1)} r'^\alpha \left[\int_{\theta'\in\Theta} \phi(r'|\theta)\hat{\xi}(\theta'|r)\mathrm{d}\theta'\right] \mathrm{d}r' \tag{34}$$

in (33). Updating the expectation about the mean of $\tilde{r}^\alpha$ has an effect on the continuation value function in (32) *through* the optimal allocation of output between consumption and saving for a one-period-horizon maximization problem. In other words, the updating component of learning alters the expected mean of $\tilde{r}^\alpha$, which, in turn, affects the optimal fraction of output consumed for a one-period-horizon dynamic program as defined by (33).

To see this, consider the first-order condition corresponding to (32),

$$c^{\alpha-1} - \delta(y-c)^{\alpha-1} \int_0^1 \left(\omega_1(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r = 0 \tag{35}$$

where the posterior beliefs affects optimal behavior through the term $\omega_1(\hat{\xi}(\cdot|r))$. From (35), the anticipation of beliefs updating embedded in the optimal fraction of output consumed in the one-period-horizon program has an effect on the optimal fraction of output consumed in the two-period-horizon program.

In summary, the updating component of learning comes into play when the horizon is equal to two periods. It is the presence of the beliefs component in the one-period horizon that leads to the dependence of the one-period-horizon value function on prior beliefs, which implies the presence of the updating component in the two-period-horizon case. Hence, the influence of learning on the dynamic program is gradual beginning with the beliefs component in a one-period horizon, joined by the updating component in the two-period horizon. Next, we show that both beliefs and updating components remain active in finite horizons greater than 2 and when the horizon is infinite.

3.2 Finite and Infinite Horizons

From the previous discussion, the zero-period- and one-period-horizon value functions are linear in y^α with the multiplicative term depending on beliefs. The two-period-horizon value function also retains the same functional form. Indeed, solving (35) yields the two-period-horizon optimal consumption

$$\rho_2^L(y,\xi) = \frac{y}{1+\delta^{\frac{1}{1-\alpha}}\left(\int_0^1 \left(\omega_1(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right)^{\frac{1}{1-\alpha}}}. \tag{36}$$

Note that (36) retains the linearity of optimal consumption in y, and thus, the two-period-horizon value function also retains the linearity in y^α.

We now solve for the τ-period-horizon optimal consumption function. Consider a τ-period horizon for which the continuation value function is of the form $V_{\tau-1}^L(y,\xi) = \left(\omega_{\tau-1}(\xi)\right)^{\alpha-1} y^\alpha$ where $\omega_{\tau-1}(\xi)\in(0,1)$ depends on beliefs. For $\tau = 2,3,\ldots$, the τ-period-horizon value function is

$$V_\tau^L(y,\xi) = \max_{c\in(0,y)}\left\{c^\alpha + \delta\int_0^1 V_{\tau-1}^L(r(y-c),\hat{\xi}(\cdot|r))\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right\}, \tag{37}$$

$$= \max_{c\in(0,y)}\left\{c^\alpha + \delta\int_0^1 \left(\omega_{\tau-1}(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha(y-c)^\alpha\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right\}. \tag{38}$$

Note that the effect of anticipating beliefs updating identified in the discussion following equation (32) remains valid here. That is, for a τ-period-horizon optimal dynamic program, the anticipation of beliefs updating alters future payoffs through the optimal fraction of output consumed corresponding to a $(\tau-1)$-period-horizon dynamic program. In fact, the influence of the updating component of learning on the maximization problem with horizon greater than or equal to 2 is contained in the optimal fraction of output consumed for the next period. This is assured by the recursive nature of the dynamic program.[15] In other words, the optimal policy function for a τ-period horizon depends on expectations of functions of the optimal fraction of output consumed for *all* shorter horizons $t < \tau$, which are contained in $\omega_{\tau-1}(\cdot)$.[16]

The first-order condition corresponding to (38)

$$c^{\alpha-1} = \delta(y-c)^{\alpha-1} \int_0^1 \left(\omega_{\tau-1}(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \tag{39}$$

yields $\rho_\tau^L(y,\xi) = \omega_\tau(\xi)y$ where $\omega_\tau(\xi)$ is defined implicitly by

$$\omega_\tau(\xi)^{\alpha-1} = \delta(1-\omega_\tau(\xi))^{\alpha-1} \int_0^1 \left(\omega_{\tau-1}(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r. \tag{40}$$

Plugging $\rho_\tau^L(y,\xi) = \omega_\tau(\xi)y$ into (38) yields

$$\begin{aligned} V_\tau^L(y,\xi) = &\left(\omega_\tau(\xi)^\alpha + \delta(1-\omega_\tau(\xi))^\alpha \int_0^1 \left(\omega_{\tau-1}(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \right. \\ &\left. \cdot \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right) y^\alpha, \qquad (41) \\ \equiv &\left(\omega_\tau(\xi)\right)^{\alpha-1} y^\alpha, \qquad (42) \end{aligned}$$

so that the optimal fraction of output consumed for the τ-period horizon is consistent with the functional form of the continuation value function and is

[15]For the one-period-horizon optimal dynamic program, there is no updating component of learning since the agent consumes all remaining output in the last period.

[16]If the horizon is two-period, then from (32) and (33), only the optimal one-period-horizon fraction of output consumed matters since the agent consumes all output remaining in the last period.

implicitly characterized by

$$\omega_\tau(\xi)^{\alpha-1} = \omega_\tau(\xi)^\alpha + \delta(1-\omega_\tau(\xi))^\alpha \int_0^1 \left(\omega_{\tau-1}(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \cdot \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r, \tag{43}$$

$$\omega_\tau(\xi) = \frac{1}{1+\delta^{\frac{1}{1-\alpha}}\left(\int_0^1 \left(\omega_{\tau-1}(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right)^{\frac{1}{1-\alpha}}}, \tag{44}$$

with, from (25), initial condition

$$\omega_1(\xi) = \left(1+\delta^{\frac{1}{1-\alpha}}\left(\int_0^1 r^\alpha \left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{-1}. \tag{45}$$

Proposition 3.3 provides the optimal policy function for any finite-horizon problem. Optimal consumption is linear in output and the optimal fraction of output consumed depends on beliefs. Expression (46) summarizes the anticipation of beliefs updating that influences optimal behavior for any finite horizon, i.e., the updating component of learning. As pointed out, the term $\omega_\tau(\xi)$ depends on the entire learning process through the recursive nature of the dynamic program defined by (46).[17] This procedure takes account of the fact that in every period beliefs are updated after the output is observed and the production shock is then deduced.

Proposition 3.3. *In the learning case, for* $\tau = 0, 1, \ldots,$ $\rho^L_\tau(y,\xi) = \omega_\tau(\xi)y$ *where* $\omega_\tau(\xi)$ *is defined recursively by*

$$\omega_\tau(\xi)^{\alpha-1} = \left(1+\delta^{\frac{1}{1-\alpha}}\left(\int_0^1 \left(\omega_{\tau-1}(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha}, \tag{46}$$

[17] Using the same recursive approach, Section 4 provides further insights about the effect of learning on the optimal dynamic program by relating this growth model with the MZ growth model (with log utility and Cobb-Douglas production).

with initial condition

$$\omega_1(\xi) = \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right)^{\frac{1}{1-\alpha}} \right)^{-1}. \tag{47}$$

Proof. Rearranging (44) yields (46). □

Proposition 3.4 provides the optimal policy function for an infinite-horizon problem. Consistent with (46), (48) summarizes the anticipation of beliefs updating that influences optimal behavior for the infinite horizon, i.e., the updating component of learning. That is, present optimal behavior depends on expected future optimal behavior, taking into account that beliefs are updated infinitely many times. This learning process is complicated by the fact that each period, the agent anticipates rationally the effect that the shock has on both beliefs and optimal behavior infinitely many times in the future.

Proposition 3.4. *From* (46)*,* $\lim_{\tau\to\infty} \rho^L_\tau(y,\xi) \equiv \rho^L_\infty(y,\xi) = \omega^L_\infty(\xi)y$ *exists and* $\omega^L_\infty(\xi) \in (0,1)$ *is defined implicitly by*

$$\omega^L_\infty(\xi)^{\alpha-1} = \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 \left(\omega^L_\infty(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right)^{\frac{1}{1-\alpha}} \right)^{1-\alpha}. \tag{48}$$

Proof. See Appendix B. □

4 Comparisons

In addition to the growth model with iso-elastic utility and linear production, there is only one growth model dealing with the issue of learning in an infinite horizon and with general distributions of the production shocks as well as beliefs, the MZ model with log utility and Cobb-Douglas production studied in KMS. To provide further insights into the effect of learning on the dynamic program, we apply the recursive approach to the MZ model and compare it to the case of iso-elastic utility and linear production.[18]

We begin with the one- and two-period-horizon maximization problems of the MZ model. We then generalize the analysis to any horizon.[19] For the MZ model, $u(c) = \ln c$ and $\hat{y} = (y-c)^r, y \in (0,1), r \in (0,1)$. Note that the restriction $y < 1$ ensures that for any $c \in (0, y)$, the production function is increasing in r, i.e., Assumption 2.4 holds. Indeed, without this restriction, $f(y-c, r) = (y-c)^r$ is strictly increasing in r for $y - c < 1$ and strictly decreasing for $y - c > 1$. Moreover, for $y - c = 1$, r cannot be inferred from observing $\hat{y}$, and thus, the signal $\hat{y}$ is uninformative about θ^*.

A Cobb-Douglas production function without any restriction on the initial output falls into the category of *learning with experimentation*. The choice of consumption alters drastically the informational content of the output-signal. To see this, suppose that $y > 1$. Then, choosing $c' > 0$ such that $y - c' = 1$ yields no information about θ^* whereas choosing $c'' > 0$ such that $y - c'' \neq 1$ provides partial information about θ. For the remainder of this section, we assume that $y \in (0,1)$, which prevents the agent from experimenting.[20]

Using (21), the value function corresponding to the MZ model is

$$V_\tau^L(y,\xi) = \max_{c \in (0,y)} \left\{ \ln c + \delta \int_0^1 V_{\tau-1}^L((y-c)^r, \hat{\xi}(\cdot|r)) \left[\int_{\theta \in \Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\} \tag{49}$$

where

$$V_0^L(y,\xi) = \ln y. \tag{50}$$

[18] KMS compares the optimal policy functions of the MZ model between stochastic and learning cases. The analogous comparison between stochastic and learning environments in the case of iso-elastic utility and linear production is found in Appendix C.

[19] In KMS, there is no recursive analysis of the MZ model. Rather, the value function in the learning case is conjectured and verified.

[20] Another way to focus on learning without experimentation (i.e., passive learning) is to assume that the production shock is directly observable as done in KMS. The direct observability of the production shock removes the opportunity to experiment as well.

4.1 One- and Two-Period Horizons

One-Period Horizon. Using (50), the one-period-horizon value function is

$$V_1^L(y,\xi) = \max_{c\in(0,y)} \left\{ \ln c + \delta \int_0^1 V_0^L((y-c)^r, \hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \tag{51}$$

$$= \max_{c\in(0,y)} \left\{ \ln c + \delta \int_0^1 r \ln(y-c) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \tag{52}$$

$$= \max_{c\in(0,y)} \left\{ \ln c + \delta \ln(y-c) \int_0^1 r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}. \tag{53}$$

The first-order condition $\frac{1}{c} = \frac{\delta \int_0^1 r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r}{y-c}$ yields the optimal one-period-horizon consumption $\rho_1^L(y,\xi) = \varphi_1(\xi)y$ where

$$\varphi_1(\xi) = \frac{1}{1 + \delta \int_0^1 r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r} \tag{54}$$

is the optimal one-period-horizon fraction of output consumed.[21] Plugging $\rho_1^L(y,\xi) = \varphi_1(\xi)y$ into (53) and using (54) yields

$$V_1^L(y,\xi) = \ln \rho_1^L(y,\xi) + \delta \left(\int_0^1 r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right) \ln(y - \rho_1^L(y,\xi)), \tag{55}$$

$$= \ln \varphi_1(\xi)y + \delta \left(\int_0^1 r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right) \ln(y - \varphi_1(\xi)y), \tag{56}$$

$$= \left(1 + \delta \int_0^1 r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right) \ln y + b_1(\xi) \tag{57}$$

[21] In order to distinguish between the two growth models, we use the notation $\varphi_\tau(\xi)$ for the MZ model as opposed to the notation $\omega_\tau(\xi)$ for the case of iso-elastic utility and linear production.

where

$$\begin{aligned} b_1(\xi) = & -\ln\left(1+\delta\int_0^1 r\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right) \\ & +\delta\int_0^1 r\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r \\ & \cdot\ln\left(\frac{\delta\int_0^1 r\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r}{1+\delta\int_0^1 r\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r}\right). \end{aligned} \tag{58}$$

As in the case of iso-elastic utility and linear production, the one-period-horizon value function contains only the beliefs component of learning. However, there is a subtle difference between (26) and (57). From (26), the derivative of the value function with respect to output depends *nonlinearly* on the integral $\int_0^1 r^\alpha \left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r$. However, from (57), the derivative of the value function with respect to output is linear in the integral $\int_0^1 r\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r$. This difference bears consequences in how the updating component of learning influences the two-period-horizon maximization problem.

Two-Period Horizon. To simplify notation, let $\mu(\theta) = \int_0^1 r\phi(r|\theta)\mathrm{d}r$ be the mean of $\tilde{r}$ conditional on θ. Using (57), the two-period-horizon value function is

$$\begin{aligned} & V_2^L(y,\xi) \\ & = \max_{c\in(0,y)}\left\{\ln c+\delta\int_0^1 V_1^L((y-c)^r,\hat{\xi}(\cdot|r))\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right\}, \quad (59) \\ & = \max_{c\in(0,y)}\left\{\ln c+\delta\ln(y-c)\int_0^1 r\left(1+\delta\int_{\theta'\in\Theta}\mu(\theta')\hat{\xi}(\theta'|r)\mathrm{d}\theta'\right)\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right. \\ & \quad \left. +\delta\cdot\int_0^1 b_1(\hat{\xi}(\cdot|r))\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right\} \quad (60) \end{aligned}$$

so that the first-order condition is

$$\frac{1}{c} = \frac{\delta\int_0^1 r\left(1+\delta\int_{\theta'\in\Theta}\mu(\theta')\hat{\xi}(\theta'|r)\mathrm{d}\theta'\right)\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r}{y-c}. \tag{61}$$

From (61), the term

$$\int_0^1 r\left(1+\delta\int_{\theta'\in\Theta}\mu(\theta')\hat{\xi}(\theta'|r)\mathrm{d}\theta'\right)\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r \tag{62}$$

contains both beliefs and updating components. Due to the linearity of the integral identified in the one-period-horizon value function, (62) can be further simplified. Using (9), plugging $\hat{\xi}(\theta'|r) = \frac{\phi(r|\theta')\xi(\theta')}{\int_{\theta''\in\Theta}\phi(r|\theta'')\xi(\theta'')\mathrm{d}\theta''}$ and $\mu(\theta) = \int_0^1 r\phi(r|\theta)\mathrm{d}r$ into (62) yields

$$\int_0^1 r\left(1+\delta\int_{\theta'\in\Theta}\mu(\theta')\frac{\phi(r|\theta')\xi(\theta')}{\int_{\theta''\in\Theta}\phi(r|\theta'')\xi(\theta'')\mathrm{d}\theta''}\mathrm{d}\theta'\right)\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r, \tag{63}$$

$$= \int_{\theta\in\Theta}(1+\delta\mu(\theta))\left(\int_0^1 r\phi(r|\theta)\mathrm{d}r\right)\xi(\theta)\mathrm{d}\theta, \tag{64}$$

$$= \int_{\theta\in\Theta}(1+\delta\mu(\theta))\mu(\theta)\xi(\theta)\mathrm{d}\theta, \tag{65}$$

$$= \int_{\theta\in\Theta}(\mu(\theta)+\delta[\mu(\theta)]^2)\xi(\theta)\mathrm{d}\theta. \tag{66}$$

The step from (63) and (64) shows how the beliefs component (defined by the integral $\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta$) is merged with the updating component (defined by the posterior beliefs). That is, from (63), given r and θ', we observe that

$$\frac{\phi(r|\theta')\xi(\theta')}{\int_{\theta''\in\Theta}\phi(r|\theta'')\xi(\theta'')\mathrm{d}\theta''}\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r = \phi(r|\theta')\xi(\theta'). \tag{67}$$

This simplification is due to the linearity of the value function with respect to the integral. Using (66), (61) is rewritten as

$$\frac{1}{c} = \frac{\delta\int_{\theta\in\Theta}(\mu(\theta)+\delta[\mu(\theta)]^2)\xi(\theta)\mathrm{d}\theta}{y-c}. \tag{68}$$

Note that in the case of iso-elastic utility and linear production, the first-order condition for the two-period-horizon maximization problem cannot be simplified in a similar way. Indeed, plugging (33) into (35) and using (9) does not combine beliefs and updating components. In other words, given the shock r, the beliefs component defined by $\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]$ in (35) cannot be combined with the posterior beliefs in (33) due to the nonlinear dependence of (33) on the integral $\int_{r'\in(0,1)} r'^{\alpha}\left[\int_{\theta'\in\Theta}\phi(r'|\theta')\hat{\xi}(\theta'|r)\mathrm{d}\theta'\right]\mathrm{d}r'$.

4.2 Finite Horizons

Next, we consider an arbitrary finite horizon to show that in general the difference between these two growth models is due to differences in the determination of the value function and the optimal policy function. In particular, for any horizon, the MZ model retains the linear dependence of the value function on the integral whereas the model with iso-elastic utility and linear production does not.

Propositions 4.1 and 4.2 provide the value function and the optimal policy function corresponding to, respectively, the MZ model and the case of iso-elastic utility and linear production. Although Proposition 3.3 already provides the optimal policy function for the case of iso-elastic utility and linear production, Proposition 4.2 highlights the joint determination of the value function and optimal policy function, which facilitates the comparison with Proposition 4.1.[22]

Proposition 4.1. *Suppose that* $u(c) = \ln c$ *and* $\hat{y} = (y-c)^r, y \in (0,1), r \in (0,1)$. *Then, for* $\tau = 1, 2, \ldots$, *the value function is*

$$V_\tau^L(y,\xi) = a_\tau(\xi)\ln y + b_\tau(\xi) \tag{69}$$

and the optimal policy function is

$$\rho_\tau^L(y,\xi) = \varphi_\tau(\xi)y \tag{70}$$

[22]Note that in both Propositions 4.1 and 4.2, the initial conditions are independent of beliefs since the entire output is consumed when there is no horizon.

where

$$\varphi_\tau(\xi) = \frac{1}{1 + \delta \int_0^1 a_{\tau-1}(\hat{\xi}(\cdot|r)) r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r} \tag{71}$$

$$a_\tau(\xi) = 1 + \delta \int_0^1 a_{\tau-1}(\hat{\xi}(\cdot|r)) r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \tag{72}$$

$$\begin{aligned} b_\tau(\xi) = \ln \varphi_\tau(\xi) &+ \delta \int_0^1 a_{\tau-1}(\hat{\xi}(\cdot|r)) r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \ln(1-\varphi_\tau(\xi)) \\ &+ \delta \int_0^1 b_{\tau-1}(\hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \end{aligned} \tag{73}$$

with initial conditions $(\varphi_0(\xi), a_0(\xi), b_0(\xi)) = (1, 1, 0)$.

Proof. See Appendix B. □

Proposition 4.2. *Suppose that* $u(c) = c^\alpha$ *and* $\hat{y} = r(y-c), r \in (0,1)$. *Then, for* $\tau = 1, 2, \ldots$, *the value function is*

$$V_\tau^L(y, \xi) = \kappa_\tau(\xi) y^\alpha \tag{74}$$

and the optimal policy function is

$$\rho_\tau^L(y, \xi) = \omega_\tau(\xi) y \tag{75}$$

where

$$\omega_\tau(\xi) = \frac{1}{1 + \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 \kappa_{\tau-1}(\hat{\xi}(\cdot|r)) r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right)^{\frac{1}{1-\alpha}}}, \tag{76}$$

$$\kappa_\tau(\xi) = (\omega_\tau(\xi))^\alpha + \delta(1-\omega_\tau(\xi))^\alpha \int_0^1 \kappa_{\tau-1}(\hat{\xi}(\cdot|r)) r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \tag{77}$$

with initial condition $(\omega_0(\xi), \kappa_0(\xi)) = (1, 1)$.

Proof. See Appendix B. □

From Proposition 4.1, the optimal policy function is defined by (71) whereas the value function is defined by (72) and (73). From (72), the term

$a_\tau(\cdot)$ is solved independently of (71) and (73). Moreover, for any horizon, the value function remains linear in the integral defined recursively by (72) such that, for $\tau = 1, 2, \ldots$,

$$a_\tau(\xi) = 1 + \int_{\theta\in\Theta} \left(\sum\nolimits_{t=1}^{\tau} \delta^t [\mu(\theta)]^t\right) \xi(\theta)\mathrm{d}\theta \tag{78}$$

where $\mu(\theta) = \int_0^1 r\phi(r|\theta)\mathrm{d}r$.

Next, from Proposition 4.2, the optimal policy function is defined by (76) whereas the value function is defined by (77). The nonlinearity of the integral with respect to prior beliefs remains valid in the case of iso-elastic utility and linear production for any horizon. In particular, plugging (76) into (77) shows that $\kappa_\tau(\cdot)$ is a nonlinear function of the integral.

5 Experimentation

So far, we have considered a maximization problem with Bayesian learning in which the agent has no opportunity to experiment, i.e., passive learning is *de facto* optimal since the agent cannot manipulate beliefs. Indeed, all information about θ^* is summarized in the signal r_t (inferable from observing y_{t+1}), which is independent of c_t. In this section, we augment the learning environment to situations in which the agent has the opportunity to experiment, i.e., to change his consumption decision in order to make the signal more informative. We present an example in which it is optimal to experiment.[23]

To address the joint issue of learning and experimentation in a growth model, the evolution of output is rewritten as

$$y_{t+1} = f(y_t - c_t, r_t, \gamma^*) \tag{79}$$

where $\gamma^* \in \Gamma$ is a parameter of the production function, which is unknown to the agent. The production shock r_t is a realization of the random vari-

[23] When the agent has the opportunity to experiment, it might be optimal not to experiment. See Mirman et al. (1993) for a discussion on conditions regarding the optimality of experimentation and passive learning in the case of a signal-independent model of experimentation in which the only link between periods is beliefs.

ables $\tilde{r}_t$ with p.d.f. $\phi(r_t)$.[24] Note that $\tilde{r}_t$ is independent of γ^*. Using (79) instead of (1), we consider the case in which the agent does not know γ^* and learns about it by observing the output signal. To analyze experimentation in a signal-dependent dynamic program, we first describe the learning process and show the dependence of posterior beliefs on decisions. We then embed learning and experimentation in the dynamic program. Finally, we characterize optimal behavior for the infinite-horizon dynamic program with a linear utility and a Cobb-Douglas production.

When the agent is uninformed about the value of γ^*, the consumption decision has an effect on the learning process. Hence, there is an opportunity to experiment, i.e., different levels of consumption imply different informational content of the signal. To see this, let $\xi_t(\gamma), \gamma \in \Gamma$ be the prior p.d.f. for γ^* at time t. At time $t+1$, given c_t, the agent observes the signal y_{t+1}, a realization of the random output $\tilde{y}_{t+1} = f(y_t - c_t, \tilde{r}_t, \gamma^*)$. The agent's posterior beliefs are then

$$\xi_{t+1}(\gamma | y_{t+1}, y_t, c_t) = \frac{\phi_Y(y_{t+1}|y_t - c_t, \gamma)\xi_t(\gamma)}{\int_{\gamma' \in \Gamma} \phi_Y(y_{t+1}|y_t - c_t, \gamma')\xi_t(\gamma')\mathrm{d}\gamma'}, \quad \gamma \in \Gamma \tag{80}$$

where $\phi_Y(y_{t+1}|y_t - c_t, \gamma)$ is the p.d.f. of $\tilde{y}_{t+1}$ conditional on γ.

When learning about γ^*, the consumption decision has an effect on posterior beliefs, and thus, the agent has the opportunity to engage in experimentation. That is, in this section, Remark 2.5 is replaced by the following remark.

Remark 5.1. *Beliefs at time* $t+1$ *are dependent on the consumption decision at time* t.

In the experimentation case, the infinite-horizon value function is

$$\begin{aligned} V^E(y_t, \xi_t) = \max_{c_t \in [0, y_t]} \Big\{ u(c_t) + \delta \int_{y_{t+1} \in Y_{t+1}} & V^E(y_{t+1}, \xi_{t+1}(\cdot | y_{t+1}, y_t - c_t)) \\ & \cdot \left[\int_{\gamma \in \Gamma} \phi_Y(y_{t+1}|y_t - c_t, \gamma)\xi_t(\gamma)\mathrm{d}\gamma \right] \mathrm{d}y_{t+1} \Big\} \end{aligned} \tag{81}$$

[24] To simplify notation, we omit the parameter θ^* in the p.d.f. of $\tilde{r}_t$.

where the superscript E stands for *experimentation*. From (81), the effect of c_t on future payoffs is two-fold, i.e., on the future stock and the posterior beliefs. Optimal growth with experimentation has never been completely solved. Indeed, in the literature, experimentation has been studied in signal-dependent dynamic models with at most a three-period horizon (Bertocchi and Spagat, 1998; Datta et al., 2002; El-Gamal and Sundaram, 1993; Huffman and Kiefer, 1994).[25]

Next, we characterize the optimal policy for a specific infinite-horizon program with experimentation. To simplify notation, we drop the time subscript and use the hat sign to distinguish between present and future values of a variable. Specifically, we assume that utility is $u(c) = c$ and the stock evolves as $\hat{y} = \gamma(y - c)^{\beta} + r$ where $\beta \in (0, 1)$ and r is a realization of a uniformly-distributed random variable, i.e., $\tilde{r} \sim U[-1, 1]$. The production shocks are independently and identically distributed (i.i.d.). Proposition 5.2 provides the optimal policy function in the stochastic case, i.e., when the parameter $\gamma^* \in \Gamma$ is known. We suppose that the initial stock is sufficiently large to ensure an interior solution.

Proposition 5.2. *Suppose that, for $\gamma^* \in \Gamma$, $y > \beta^{\frac{1}{1-\beta}} \delta^{\frac{1}{1-\beta}} (\gamma^*)^{\frac{1}{1-\beta}}$. In the stochastic case, the infinite-horizon optimal policy function is*

$$g^S(y) = y - \beta^{\frac{1}{1-\beta}} \delta^{\frac{1}{1-\beta}} (\gamma^*)^{\frac{1}{1-\beta}}. \tag{82}$$

Proof. See Appendix B. □

Next, consider the case in which γ^* is unknown. Suppose further that the true value of the unknown parameter is either $\overline{\gamma}$ or $\underline{\gamma}$, $\overline{\gamma} > \underline{\gamma} > 0$. Let

[25] For dynamic signal-*independent* models of experimentation in which the only link between periods is beliefs, see Prescott (1972), Grossman et al. (1977), Easley and Kiefer (1988, 1989), Kiefer and Nyarko (1989), Balvers and Cosimano (1990), Aghion et al. (1991), Fusselman and Mirman (1993), Mirman et al. (1993), Trefler (1993), Creane (1994), Fishman and Gandal (1994), Keller and Rady (1999), and Wieland (2000).

$\xi \in (0,1)$ be the prior probability that $\gamma = \overline{\gamma}$. Hence, (81) is rewritten as

$$V^E(y,\xi) = \max_{c\in[0,y]} \Big\{ c + \delta\xi \int_{-1}^{1} (1/2) V^E(\overline{\gamma}(y-c)^\beta + r, \hat{\xi}(\overline{\gamma}(y-c)^\beta + r))\mathrm{d}r + \delta(1-\xi) \int_{-1}^{1} (1/2) V^E(\underline{\gamma}(y-c)^\beta + r, \hat{\xi}(\underline{\gamma}(y-c)^\beta + r))\mathrm{d}r \Big\} \tag{83}$$

where with probability ξ, the random signal is $\hat{y} = \overline{\gamma}(y-c)^\beta + r$ leading to random posterior $\hat{\xi}(\overline{\gamma}(y-c)^\beta + r)$ and with probability $1-\xi$, the random signal is $\hat{y} = \underline{\gamma}(y-c)^\beta + r$ with the corresponding random posterior $\hat{\xi}(\underline{\gamma}(y-c)^\beta + r)$. By iterating on the value function, the conjecture of the value function is found to be

$$V^E(y,\xi) = y + A(\xi) \tag{84}$$

where $A(\xi)$ is a function of beliefs. Plugging (84) into (83) yields

$$V^E(y,\xi) = \max_{c\in[0,y]} \big\{ c + \delta(\xi\overline{\gamma} + (1-\xi)\underline{\gamma})(y-c)^\beta + \delta \int_{-1}^{1} (1/2) \Big(\xi A(\hat{\xi}(\overline{\gamma}(y-c)^\beta + r)) + (1-\xi) A(\hat{\xi}(\underline{\gamma}(y-c)^\beta + r)) \Big) \mathrm{d}r \Big\} \tag{85}$$

where the derivative of the term $(\xi\overline{\gamma} + (1-\xi)\underline{\gamma})(y-c)^\beta$ with respect to c is the effect of consumption on future payoffs through the future stock, whereas the derivative of the term

$$\int_{-1}^{1} (1/2) \Big(\xi A(\hat{\xi}(\overline{\gamma}(y-c)^\beta + r)) + (1-\xi) A(\hat{\xi}(\underline{\gamma}(y-c)^\beta + r)) \Big) \mathrm{d}r \tag{86}$$

with respect to c is the experimentation effect, i.e., the effect of consumption on future payoffs via posterior beliefs.

In order to study the value function in the case of learning with experimentation, we need to explain how the updating of beliefs is anticipated by focusing on (86). To that end, we perform a change of variable by working directly with the *random posterior beliefs*. The uniform distribution assumption implies that the support for the posterior beliefs has only three elements. Indeed, conditional on $\{y, c\}$, the support of the output-signal is

$$\hat{y}|\underline{\gamma} \in [\underline{\gamma}(y-c)^{\beta} - 1, \underline{\gamma}(y-c)^{\beta} + 1] \tag{87}$$

when $\gamma^* = \underline{\gamma}$, and

$$\hat{y}|\overline{\gamma} \in [\overline{\gamma}(y-c)^{\beta} - 1, \overline{\gamma}(y-c)^{\beta} + 1] \tag{88}$$

when $\gamma^* = \overline{\gamma}$. First, suppose that we observe $\hat{y} \in [\underline{\gamma}(y-c)^{\beta} - 1, \overline{\gamma}(y-c)^{\beta} - 1)]$. Then, the signal cannot be generated under $\gamma^* = \overline{\gamma}$ and thus $\hat{\xi} = 0$. Next, when the realization is $\hat{y} \in (\underline{\gamma}(y-c)^{\beta} + 1, \overline{\gamma}(y-c)^{\beta} + 1]$, the signal cannot be generated by $\gamma^* = \underline{\gamma}$ so that $\hat{\xi} = 1$. Finally, if $\hat{y} \in [\overline{\gamma}(y-c)^{\beta} - 1, \underline{\gamma}(y-c)^{\beta} + 1]$ with $\overline{\gamma}(y-c)^{\beta} - 1 < \underline{\gamma}(y-c)^{\beta} + 1$, then it is equally probable that the signal is generated by either $\gamma^* = \overline{\gamma}$ or $\gamma^* = \underline{\gamma}$, which provides no information, i.e., $\hat{\xi} = \frac{\frac{1}{2}\xi}{\frac{\xi}{2} + \frac{1-\xi}{2}} = \xi$. In our example, the consumption decision has an effect on the support of the output signal.

Having established that the support of the posterior has three elements, i.e., $\hat{\xi} \in \{0, 1, \xi\}$, we turn to calculating the probabilities to observe each of these elements. Specifically,

$$\begin{aligned} \Pr[\hat{\xi} = 1|\xi] = \; & \Pr[\gamma^* = \overline{\gamma}] \cdot \Pr[\hat{y} \in (\underline{\gamma}(y-c)^{\beta} + 1, \overline{\gamma}(y-c)^{\beta} + 1]|\gamma^* = \overline{\gamma}] \\ & + \Pr[\gamma^* = \underline{\gamma}] \cdot \Pr[\hat{y} \in (\underline{\gamma}(y-c)^{\beta} + 1, \overline{\gamma}(y-c)^{\beta} + 1]|\gamma^* = \underline{\gamma}] \end{aligned} \tag{89}$$

where $\Pr[\hat{y} \in (\underline{\gamma}(y-c)^{\beta} + 1, \overline{\gamma}(y-c)^{\beta} + 1)|\gamma^* = \underline{\gamma}] = 0$ since it is impossible for the signal conditional on $\gamma^* = \underline{\gamma}$ to fall between $\underline{\gamma}(y-c)^{\beta} + 1$ and $\overline{\gamma}(y-c)^{\beta} + 1$. Thus,

$$\begin{aligned} \Pr[\hat{\xi} = 1|\xi] = \; & \xi \cdot \frac{\overline{\gamma}(y-c)^{\beta} - 1 - (\underline{\gamma}(y-c)^{\beta} - 1)}{\underline{\gamma}(y-c)^{\beta} + 1 - (\underline{\gamma}(y-c)^{\beta} - 1)} \\ & + (1-\xi) \cdot 0, && (90) \\ = \; & \xi \cdot \frac{(\overline{\gamma} - \underline{\gamma})(y-c)^{\beta}}{2} > 0. && (91) \end{aligned}$$

Similarly,

$$\Pr[\hat{\xi} = 0|\xi] = (1-\xi) \cdot \frac{(\overline{\gamma} - \underline{\gamma})(y-c)^{\beta}}{2} > 0, \tag{92}$$

and thus, whenever $(\overline{\gamma} - \underline{\gamma})(y-c)^{\beta}/2 < 1$,

$$\Pr[\hat{\xi} = \xi] = 1 - \frac{(\overline{\gamma} - \underline{\gamma})(y-c)^{\beta}}{2} > 0. \tag{93}$$

Note that a decrease in c makes the signal more informative, i.e., the probability to learn nothing decreases, i.e., $\Pr[\hat{\xi} = \xi]$ is decreasing along with an increase in c.

Hence, posterior beliefs about γ^* evolve as a random variable with support $\hat{\xi} \in \{0, 1, \xi\}$ and corresponding probabilities

$$\left\{(1-\xi) \cdot \frac{(\overline{\gamma} - \underline{\gamma})(y-c)^{\beta}}{2}, \xi \cdot \frac{\left(\overline{\gamma} - \underline{\gamma}\right)(y-c)^{\beta}}{2}, 1 - \frac{(\overline{\gamma} - \underline{\gamma})(y-c)^{\beta}}{2}\right\}. \tag{94}$$

Next, using (91), (92), and (93), and the conjecture $V^E(y,\xi) = y + A(\xi)$, for $\xi \in (0,1)$, (85) is rewritten as,[26]

$$\begin{aligned} V^E(y,\xi) = \quad & \max_{c \in [0,y]} \Big\{ c + \delta(\xi\overline{\gamma} + (1-\xi)\underline{\gamma})(y-c)^{\beta} \\ & + \delta\xi \frac{(\overline{\gamma} - \underline{\gamma})(y-c)^{\beta}}{2} A(1) \\ & + \delta(1-\xi)\frac{(\overline{\gamma} - \underline{\gamma})(y-c)^{\beta}}{2} A(0) \\ & + \delta\left(1 - \frac{(\overline{\gamma} - \underline{\gamma})(y-c)^{\beta}}{2}\right) A(\xi) \Big\}. \end{aligned} \tag{95}$$

Proposition 5.3 provides the optimal policy function in the experimentation case. As in the stochastic case, we suppose that the initial stock is sufficiently large to ensure an interior solution. Moreover, we suppose that parameter values satisfy $\frac{(\overline{\gamma} - \underline{\gamma})^{\frac{1}{\beta}}(1-\delta)\beta A(\xi)}{2^{\frac{1}{\beta}}(1-\beta)} < 1$ so that at the optimum, $\Pr[\hat{\xi} = 1|\xi], \Pr[\hat{\xi} = 0|\xi], \Pr[\hat{\xi} = \xi|\xi] \in (0,1)$.

[26] Note that (85) cannot be rewritten as (95) for $\xi \in \{0, 1\}$.

Proposition 5.3. *Suppose that* $y > \max_{\xi\in[0,1]}\left\{\frac{(1-\delta)\beta A(\xi)}{1-\beta}\right\}$ *and* $\frac{(\overline{\gamma}-\underline{\gamma})^{\frac{1}{\beta}}(1-\delta)\beta A(\xi)}{2^{\frac{1}{\beta}}(1-\beta)} < 1$. *In the experimentation case, the infinite-horizon optimal policy function is*

$$g^E(y,\xi) = y - \frac{(1-\delta)\beta A(\xi)}{1-\beta} \tag{96}$$

where $A(\xi)$ *is a function of prior beliefs, and, for* $\xi \in (0,1)$, *uniquely defined by*[27]

$$A(\xi) = \frac{\delta^{\frac{1}{1-\beta}}\beta^{\frac{\beta}{1-\beta}}}{2^{\frac{1}{1-\beta}}}\frac{1-\beta}{1-\delta} \cdot \left(2(\xi\overline{\gamma}+(1-\xi)\underline{\gamma}) + (\overline{\gamma}-\underline{\gamma})\left(\delta^{\frac{1}{1-\beta}}\beta^{\frac{\beta}{1-\beta}}\frac{1-\beta}{1-\delta}\left(\xi\overline{\gamma}^{\frac{1}{1-\beta}} + (1-\xi)\underline{\gamma}^{\frac{1}{1-\beta}}\right) - A(\xi)\right)\right)^{\frac{1}{1-\beta}}. \tag{97}$$

Proof. See Appendix B. □

Note that the difference between (82) and (96) captures the full effect of learning with experimentation on optimal behavior, i.e.,

$$g^E(y,\xi) - g^S(y) = \beta^{\frac{1}{1-\beta}}\delta^{\frac{1}{1-\beta}}(\gamma^*)^{\frac{1}{1-\beta}} - \frac{(1-\delta)\beta A(\xi)}{1-\beta}. \tag{98}$$

6 Final Remarks

The relation between present decisions and future outcomes forms the basis for studying a dynamic economy. The inclusion of uncertainty is essential to address dynamic issues in economics. In this paper, we have gone beyond the classical analysis of optimal growth under uncertainty by combining control and learning in dynamic programming. Although learning has a profound effect on behavior, optimal growth with learning falls in the category of rational expectations model. In other words, the agent forms expectations that are consistent with his beliefs and with his imperfect knowledge of the dynamics. In particular, the agent forms rational expectations about his future output and his future beliefs.

[27] For $\xi \in \{0,1\}$, $A(1)$ and $A(0)$ are defined by (158) and (159), respectively.

As shown throughout the paper, optimal growth with learning is not a simple extension of the stochastic case. Indeed, control and learning are not separable, and, thus, treating such problems sequentially by first solving a standard stochastic growth model with known distribution and then merely replacing the known distribution by a believed distribution yields non-optimal solutions. Such approach minimizes the profound effect of learning on decision-making.

This paper has provided general results for the case of learning without experimentation along with characterizing the optimal policy function for another class of growth models without making use of prior conjugates. Finally, it has also provided a first example of optimal behavior in an infinite horizon in the case of learning with experimentation. It would be interesting to study whether and when the Euler equation is relevant in the case of learning with experimentation.

A On the Differentiability of the Value Function

In this section, we prove Lemma 2.6. The proof follows closely the proof of Lemma 1 in Mirman and Zilcha (1975). We begin by showing boundedness and strict concavity of the value function.

Notation. Let $\mathbf{k} = (k_0, k_1, \ldots)$ and $\mathbf{c} = (c_0, c_1, \ldots)$ such that, for $t = 1, 2, \ldots$,

$$\begin{aligned} c_0 + k_0 &= y_0 \\ c_t + k_t &= f(k_{t-1}, r_{t-1}) \end{aligned} \tag{99}$$

is satisfied. The pair $(\mathbf{k}, \mathbf{c})$ is said to be a feasible program from initial stock y_0 and a given sequence $\{r_t\}_{t=0}^{\infty}$.

Boundedness. There exists a unique $y_{\overline{r}}$ such that $f(y_{\overline{r}}, \overline{r}) = y_{\overline{r}}$. For a given sequence $\{r_t\}_{t=0}^{\infty}$, any feasible consumption plans $C = \{c_t\}_{t=0}^{\infty}$ yields

$$\sum_{t=0}^{\infty} \delta^t u(c_t) \leq \frac{\max\{u(y_0), u(y_{\overline{r}})\}}{1-\delta} < \infty. \tag{100}$$

Taking expectations over $\{\tilde{r}_t\}_{t=0}^{\infty}$ and thus over the future beliefs $\{\tilde{\xi}_t\}_{t=0}^{\infty}$,[28] it follows that

$$\sum_{t=0}^{\infty} \delta^t \mathbb{E} u(c_t) \leq \frac{\max\{u(y_0), u(y_{\overline{r}})\}}{1-\delta} < \infty. \tag{101}$$

Hence, for any stock y_t and beliefs ξ_t, the value function is well-defined, i.e.,

$$V(y_t, \xi_t) = \sup_{\{c_t, c_{t+1}, \ldots\}} \sum_{t=0}^{\infty} \delta^t \mathbb{E} u(c_t) < \infty \tag{102}$$

subject to (99).

Strict Concavity. We now show strict concavity of of the value function. Consider $y_t^1 > 0$ and $y_t^2 > 0$ and $0 < \lambda < 1$.

[28] Given Assumption 2.4, the sequence $\{r_t\}_{t=0}^{\infty}$ implies a sequence of beliefs $\{\xi_t\}_{t=1}^{\infty}$ about the distribution of the production shocks.

First, for a given sequence $\{r_\tau\}_{\tau=t}^{\infty}$ and thus a sequence of beliefs $\{\xi_t\}_{\tau=t+1}^{\infty}$, to each $\varepsilon > 0$, there exists feasible consumption plans $C^1 = \{c_\tau^1\}_{\tau=t}^{\infty}$ and $C^2 = \{c_\tau^2\}_{\tau=t}^{\infty}$ from y_t^1 and y_t^2 such that

$$V^L(y_t^1, \xi_t) \leq \sum\nolimits_{\tau=t}^{\infty} \delta^t \mathbb{E}u(c_\tau^1) + \varepsilon \tag{103}$$

and

$$V^L(y_t^2, \xi_t) \leq \sum\nolimits_{\tau=t}^{\infty} \delta^t \mathbb{E}u(c_\tau^2) + \varepsilon. \tag{104}$$

Second, the path $C = \{\lambda c_\tau^1 + (1-\lambda)c_\tau^2\}_{\tau=t}^{\infty}$ is feasible from $\lambda y_t^1 + (1-\lambda)y_t^2$. Hence,

$$V^L(\lambda y_t^1 + (1-\lambda)y_t^2, \xi_t) \geq \sum\nolimits_{\tau=t}^{\infty} \delta^t \mathbb{E}u(\lambda c_\tau^1 + (1-\lambda)c_\tau^2) \tag{105}$$

Third, since $f(k_\tau, r_\tau)$ is strictly increasing in r_τ, learning is passive and for a given sequence of $\{r_\tau\}_{\tau=t}^{\infty}$, $C = \{\lambda c_\tau^1 + (1-\lambda)c_\tau^2\}_{\tau=t}^{\infty}$ yields the same information as the path $C^1 = \{c_\tau^1\}_{\tau=t}^{\infty}$ and $C^2 = \{c_\tau^2\}_{\tau=t}^{\infty}$. In other words, changing the consumption path from C^1 or C^2 to C does not alter future beliefs since the sequence of $\{r_\tau\}_{\tau=t}^{\infty}$ is the same and thus the information obtained is the same, i.e., the sequence of beliefs $\{\xi_t\}_{\tau=t+1}^{\infty}$ is the same. Hence, given the strict concavity of u and the fact that expectations are the same, the following inequality holds, i.e.,

$$\sum\nolimits_{\tau=t}^{\infty} \delta^t \mathbb{E}u(\lambda c_\tau^1 + (1-\lambda)c_\tau^2) \geq \sum\nolimits_{\tau=t}^{\infty} \delta^t [\lambda \mathbb{E}u(c_\tau^1) + (1-\lambda)\mathbb{E}u(c_\tau^2)]. \tag{106}$$

Finally, we know that

$$\sum\nolimits_{\tau=t}^{\infty} \delta^t [\lambda \mathbb{E}u(c_\tau^1) + (1-\lambda)\mathbb{E}u(c_\tau^2)] \geq \lambda V^L(y_t^1, \xi_t) + (1-\lambda)V^L(y_t^2, \xi_t) - \varepsilon. \tag{107}$$

Combining (105), (106), and (107), $V^L(y_t, \xi_t)$ is strictly concave in y_t since ε is arbitrary.

Continuity follows from concavity.

Existence and Uniqueness. Continuity implies that there exists an optimum on $c_t \in [0, y_t]$. Given that Conditions 2.3 are satisfied, we have an interior solution, which is unique by strict concavity.

Proof of Lemma 2.6. Using a dynamic programming argument, given y_t, the infinite-horizon value function is rewritten as in (10).

Let $y_t > 0$ and $\rho^L(y_\tau, \xi_\tau) \in (0, y_\tau)$ be the optimal consumption plan from time $\tau = t, t+1, \ldots$. Define a feasible consumption program from $y_t + \Delta y_t$ where $\Delta y_t > 0$ as follows: $c_t = \rho^L(y_t, \xi_t) + \Delta y_t$ and $c_{t+\tau} = \rho^L(y_{t+\tau}, \xi_{t+\tau})$ for $\tau = 1, 2, \ldots$ Note that by increasing both output and consumption by Δy_t, this feasible consumption program does not altered the future path of output, i.e., it does not affect the amount saved,

$$y_t + \Delta y_t - \rho^L(y_t, \xi_t) - \Delta y_t = y_t - \rho^L(y_t, \xi_t). \tag{108}$$

Hence, the expected continuation value function is the same and in particular, the updating component of learning is not affected, i.e., the distributions of the posterior beliefs remain the same.

Using (10),

$$\begin{aligned}
V^L(y_t + \Delta y_t, \xi_t) \geq\ & u(\rho^L(y_t, \xi_t) + \Delta y_t) \\
& + \delta \int_{\underline{r}}^{\overline{r}} V^L(f(y_t + \Delta y_t - \rho^L(y_t, \xi_t) - \Delta y_t, r_t), \xi_{t+1}(\cdot|r_t)) \\
& \cdot \left[\int_{\theta \in \Theta} \phi(r_t|\theta) \xi_t(\theta) \mathrm{d}\theta \right] \mathrm{d}r_t, \qquad (109) \\
=\ & u(\rho^L(y_t, \xi_t) + \Delta y_t) \\
& + \delta \int_{\underline{r}}^{\overline{r}} V^L(f(y_t - \rho^L(y_t, \xi_t), r_t), \xi_{t+1}(\cdot|r_t)) \\
& \cdot \left[\int_{\theta \in \Theta} \phi(r_t|\theta) \xi_t(\theta) \mathrm{d}\theta \right] \mathrm{d}r_t. \qquad (110)
\end{aligned}$$

It follows that

$$
\begin{aligned}
V^L(y_t+\Delta y_t,\xi_t)-V^L(y_t,\xi_t) \geq\ & u(\rho^L(y_t,\xi_t)+\Delta y_t) \\
& +\delta\int_{\underline{r}}^{\overline{r}} V^L(f(y_t-\rho^L(y_t,\xi_t),r_t),\xi_{t+1}(\cdot|r_t)) \\
& \cdot\left[\int_{\theta\in\Theta}\phi(r_t|\theta)\xi_t(\theta)\mathrm{d}\theta\right]\mathrm{d}r_t \\
& -u(\rho^L(y_t,\xi_t)) \\
& -\delta\int_{\underline{r}}^{\overline{r}} V^L(f(y_t-\rho^L(y_t,\xi_t),r_t),\xi_{t+1}(\cdot|r_t)) \\
& \cdot\left[\int_{\theta\in\Theta}\phi(r_t|\theta)\xi_t(\theta)\mathrm{d}\theta\right]\mathrm{d}r_t \qquad (111) \\
=\ & u(\rho^L(y_t,\xi_t)+\Delta y_t)-u(\rho^L(y_t,\xi_t)) \qquad (112) \\
=\ & u'(\rho^L(y_t,\xi_t))\Delta y_t+o(\Delta y_t) \qquad (113)
\end{aligned}
$$

Therefore, $\lim_{\Delta y_t\to 0}\frac{V^L(y_t+\Delta y_t,\xi_t)-V^L(y_t,\xi_t)}{\Delta y_t}\geq u'(\rho^L(y_t,\xi_t))$.

Similarly, let $0<\Delta y_t<\rho^L(y_t,\xi_t)$. Define a feasible program from $y_t-\Delta y_t$ as follows: $c_t=\rho^L(y_t,\xi_t)-\Delta y_t$ and $c_{t+\tau}=\rho^L(y_{t+\tau},\xi_{t+\tau})$ for $\tau=1,2,\ldots$ Using the same argument as above,

$$
V^L(y_t-\Delta y_t,\xi_t)-V^L(y_t,\xi_t)\geq -u'(\rho^L(y_t,\xi_t))\Delta y_t+o(\Delta y_t). \qquad (114)
$$

Hence, $\lim_{\Delta y_t\to 0}-\frac{V^L(y_t-\Delta y_t,\xi_t)-V^L(y_t,\xi_t)}{\Delta y_t}\leq u'(\rho^S(y_t))$.

Therefore, $\frac{\partial V^L(y_t,\xi_t)}{\partial y_t}=u'(\rho^L(y_t,\xi_t))$.

B Proofs

Proof of Proposition 3.4. Let $\kappa_\tau(\xi)\equiv\omega_\tau(\xi)^{\alpha-1}$ so that (46) is rewritten as

$$
\kappa_\tau(\xi)=\left(1+\delta^{\frac{1}{1-\alpha}}\left(\int_0^1\kappa_{\tau-1}(\hat{\xi}(\cdot|r))r^\alpha\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha} \qquad (115)
$$

with initial condition $\kappa_0(\xi) = 1$ since $V_0^L(y, \xi) = y^\alpha$.

1. Monotonicity of $\kappa_\tau(\xi)$. Using $\kappa_0(\xi) = 1$ and (115), we know that

$$\kappa_0(\xi) = 1 < \kappa_1(\xi) = \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha}. \tag{116}$$

Suppose next that $\kappa_\tau(\xi) > \kappa_{\tau-1}(\xi) > 0$. Then,

$$\kappa_{\tau+1}(\xi) = \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 \kappa_\tau(\hat{\xi}(\cdot|r)) r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha} \tag{117}$$

$$> \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 \kappa_{\tau-1}(\hat{\xi}(\cdot|r)) r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha} \tag{118}$$

$$= \kappa_\tau(\xi). \tag{119}$$

2. Boundedness of $\kappa_\tau(\xi)$. Let

$$M = \left(\frac{1}{1 - \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right)^{\frac{1}{1-\alpha}}}\right)^{1-\alpha} > 1 \tag{120}$$

since $\int_0^1 r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \in (0, 1)$. Hence, $\kappa_0(\xi) = 1 < M$. Suppose next that $\kappa_\tau(\xi) < M$. Then,

$$\kappa_{\tau+1}(\xi) = \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 \kappa_\tau(\hat{\xi}(\cdot|r)) r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha} \tag{121}$$

$$< \left(1 + \delta^{\frac{1}{1-\alpha}} M^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha} \tag{122}$$

$$= M. \tag{123}$$

where the last equality comes from (120).

3. Since $\lim_{\tau\to\infty} \kappa_\tau(\xi)$ exists, so does $\lim_{\tau\to\infty} \omega^L_\tau(\xi)$. Since $\kappa_\tau(\xi) \equiv \omega_\tau(\xi)^{1-\alpha}$, taking the limits on both sides of (115) yields (48).

Proof of Proposition 4.1. Suppose that

$$V^L_{\tau-1}(y,\xi) = a_{\tau-1}(\xi)\ln y + b_{\tau-1}(\xi). \tag{124}$$

Then, the τ-period-horizon value function is

$$V^L_\tau(y,\xi) = \max_{c\in(0,y)} \left\{ \ln c + \delta \int_0^1 V^L_{\tau-1}((y-c)^r, \hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \right\} \tag{125}$$

$$\begin{aligned} &= \max_{c\in(0,y)} \left\{ \ln c + \delta \int_0^1 a_{\tau-1}(\hat{\xi}(\cdot|r)) \ln(y-c)^r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \right. \\ &\quad \left. + \delta \int_0^1 b_{\tau-1}(\hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \right\} \qquad (126) \\ &= \max_{c\in(0,y)} \left\{ \ln c + \delta \ln(y-c) \int_0^1 a_{\tau-1}(\hat{\xi}(\cdot|r)) r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \right. \\ &\quad \left. + \delta \int_0^1 b_{\tau-1}(\hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \right\}. \qquad (127) \end{aligned}$$

The first-order condition is

$$\frac{1}{c} = \frac{\delta \int_0^1 a_{\tau-1}(\hat{\xi}(\cdot|r)) r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r}{y-c} \tag{128}$$

so that the τ-period-horizon optimal consumption is

$$\rho^L_\tau(y,\xi) = \varphi_\tau(\xi) y \tag{129}$$

where

$$\varphi_\tau(\xi) = \frac{1}{1 + \delta \int_0^1 a_{\tau-1}(\hat{\xi}(\cdot|r)) r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r} \tag{130}$$

is the τ-period-horizon optimal fraction of output consumed. Hence,

$$V_\tau^L(y,\xi) = \ln \varphi_\tau(\xi) y + \delta \int_0^1 a_{\tau-1}(\hat{\xi}(\cdot|r)) r \ln(y - \varphi_\tau(\xi) y) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r$$

$$+ \delta \int_0^1 b_{\tau-1}(\hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \tag{131}$$

$$= \left(1 + \delta \int_0^1 a_{\tau-1}(\hat{\xi}(\cdot|r)) r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right) \ln y$$

$$+ \ln \varphi_\tau(\xi) + \delta \ln(1 - \varphi_\tau(\xi)) \int_0^1 a_{\tau-1}(\hat{\xi}(\cdot|r)) r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r$$

$$+ \delta \int_0^1 b_{\tau-1}(\hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \tag{132}$$

$$\equiv a_\tau(\xi) \ln y + b_\tau(\xi) \tag{133}$$

so that

$$a_\tau(\xi) = 1 + \delta \int_0^1 a_{\tau-1}(\hat{\xi}(\cdot|r)) r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \tag{134}$$

$$b_\tau(\xi) = \ln \varphi_\tau(\xi) + \delta \ln(1 - \varphi_\tau(\xi)) \int_0^1 a_{\tau-1}(\hat{\xi}(\cdot|r)) r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r$$

$$+ \delta \int_0^1 b_{\tau-1}(\hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r. \tag{135}$$

Expressions (130), (134), and (135) combined with the initial condition $(\varphi_0(\xi), a_0(\xi), b_0(\xi)) = (1, 1, 0)$ yields (71), (72), and (73).

Proof of Proposition 4.2. Suppose that

$$V_{\tau-1}^L(y,\xi) = \kappa_{\tau-1}(\xi) y^\alpha. \tag{136}$$

Then, the τ-period-horizon value function is

$$V_\tau^L(y,\xi) = \max_{c\in(0,y)} \left\{ c^\alpha + \delta \int_0^1 V_{\tau-1}^L(r(y-c), \hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\} \tag{137}$$

$$= \max_{c\in(0,y)} \left\{ c^\alpha + \delta(y-c)^\alpha \int_0^1 \kappa_{\tau-1}(\hat{\xi}(\cdot|r)) r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}. \tag{138}$$

The first-order condition is

$$\alpha c^{\alpha-1} = \alpha\delta(y-c)^{\alpha-1} \int_0^1 \kappa_{\tau-1}(\hat{\xi}(\cdot|r)) r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \tag{139}$$

so that

$$\rho_\tau^L(y,\xi) = \omega_\tau(\xi) y \tag{140}$$

where

$$\omega_\tau(\xi) = \frac{1}{1 + \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 \kappa_{\tau-1}(\hat{\xi}(\cdot|r)) r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right)^{\frac{1}{1-\alpha}}}. \tag{141}$$

Hence, the τ-period-horizon value function is

$$V_\tau^L(y,\xi) = \left((\omega_\tau(\xi))^\alpha + \delta\,(1-\omega_\tau(\xi))^\alpha \int_0^1 \kappa_{\tau-1}(\hat{\xi}(\cdot|r)) r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right) y^\alpha \tag{142}$$

$$\equiv \kappa_\tau(\xi) y^\alpha \tag{143}$$

so that

$$\kappa_\tau(\xi) = (\omega_\tau(\xi))^\alpha + \delta\,(1-\omega_\tau(\xi))^\alpha \int_0^1 \kappa_{\tau-1}(\hat{\xi}(\cdot|r)) r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r. \tag{144}$$

Expressions (141) and (144) combined with the initial condition $(\omega_0(\xi), \kappa_0(\xi)) = (1,1)$ yields (76) and (77).

Proof of Proposition 5.2. Here, to simplify notation, we omit the asterisk sign on γ. The value function is

$$V^S(y) = \max_{c\in[0,y]}\{c + \delta \int_{-1}^{1} (1/2)V^S(\gamma(y-c)^\beta + r)\mathrm{d}r\}. \tag{145}$$

We conjecture that $V^S(y) = y + A$ where A is a parameter. Plugging the conjecture into (145) yields

$$V^S(y) = \max_{c\in[0,y]}\{c + \delta\gamma(y-c)^\beta + \delta A\}. \tag{146}$$

The first-order condition $1 - \beta\delta\gamma(y-c)^{\beta-1} = 0$ yields (82). Plugging (82) into the maximand of (145) yields

$$\begin{aligned} V^S(y) &= y - \beta^{\frac{1}{1-\beta}}\delta^{\frac{1}{1-\beta}}\gamma^{\frac{1}{1-\beta}} + \beta^{\frac{\beta}{1-\beta}}\delta^{\frac{1}{1-\beta}}\gamma^{\frac{1}{1-\beta}} + \delta A && (147)\\ &= y + \delta^{\frac{1}{1-\beta}}\beta^{\frac{\beta}{1-\beta}}(1-\beta)\gamma^{\frac{1}{1-\beta}} + \delta A && (148)\\ &\equiv y + A && (149)\end{aligned}$$

so that the conjecture is verified when

$$A = \frac{\delta^{\frac{1}{1-\beta}}\beta^{\frac{\beta}{1-\beta}}(1-\beta)\gamma^{\frac{1}{1-\beta}}}{1-\delta}, \tag{150}$$

which implies that (82) is the optimal policy function for the infinite-horizon dynamic program in the stochastic case.

Proof of Proposition 5.3. The first-order condition corresponding to (95) is

$$\begin{aligned} &1 - \delta\beta(\xi\overline{\gamma} + (1-\xi)\underline{\gamma})(y-c)^{\beta-1}\\ &- \delta\beta(\overline{\gamma} - \underline{\gamma})(y-c)^{\beta-1}(\xi A(1) + (1-\xi)A(0) - A(\xi))/2 = 0 \end{aligned} \tag{151}$$

so that

$$g^E(y,\xi) = y - \frac{\delta^{\frac{1}{1-\beta}}\beta^{\frac{1}{1-\beta}}}{2^{\frac{1}{1-\beta}}}\left(2(\xi\overline{\gamma} + (1-\xi)\underline{\gamma}) + (\overline{\gamma}-\underline{\gamma})(\xi A(1) + (1-\xi)A(0) - A(\xi))\right)^{\frac{1}{1-\beta}}. \tag{152}$$

Plugging (152) into the objective function of (95) yields

$$\begin{aligned}
&V^E(y,\xi) \\
&= y + \delta A(\xi) \\
&\quad - \frac{\delta^{\frac{1}{1-\beta}}\beta^{\frac{1}{1-\beta}}}{2^{\frac{1}{1-\beta}}}\left(2(\xi\overline{\gamma} + (1-\xi)\underline{\gamma}) + (\overline{\gamma}-\underline{\gamma})(\xi A(1) + (1-\xi)A(0) - A(\xi))\right)^{\frac{1}{1-\beta}} \\
&\quad + \delta(\xi\overline{\gamma} + (1-\xi)\underline{\gamma})\frac{\delta^{\frac{\beta}{1-\beta}}\beta^{\frac{\beta}{1-\beta}}}{2^{\frac{\beta}{1-\beta}}}\left(2(\xi\overline{\gamma} + (1-\xi)\underline{\gamma}) + (\overline{\gamma}-\underline{\gamma})(\xi A(1) + (1-\xi)A(0) - A(\xi))\right)^{\frac{\beta}{1-\beta}} \\
&\quad + \frac{\delta(\overline{\gamma}-\underline{\gamma})}{2}(\xi A(1) + (1-\xi)A(0) - A(\xi))\frac{\delta^{\frac{\beta}{1-\beta}}\beta^{\frac{\beta}{1-\beta}}}{2^{\frac{\beta}{1-\beta}}} \\
&\quad \cdot \left(2(\xi\overline{\gamma} + (1-\xi)\underline{\gamma}) + (\overline{\gamma}-\underline{\gamma})(\xi A(1) + (1-\xi)A(0) - A(\xi))\right)^{\frac{\beta}{1-\beta}} && (153) \\
&= y + \delta A(\xi) \\
&\quad - \frac{\delta^{\frac{1}{1-\beta}}\beta^{\frac{1}{1-\beta}}}{2^{\frac{1}{1-\beta}}}\left(2(\xi\overline{\gamma} + (1-\xi)\underline{\gamma}) + (\overline{\gamma}-\underline{\gamma})(\xi A(1) + (1-\xi)A(0) - A(\xi))\right)^{\frac{1}{1-\beta}} \\
&\quad + \frac{\delta^{\frac{1}{1-\beta}}\beta^{\frac{\beta}{1-\beta}}}{2^{\frac{1}{1-\beta}}}\left(2(\xi\overline{\gamma} + (1-\xi)\underline{\gamma}) + (\overline{\gamma}-\underline{\gamma})(\xi A(1) + (1-\xi)A(0) - A(\xi))\right)^{\frac{1}{1-\beta}} && (154) \\
&= y + \delta A(\xi) \\
&\quad + \frac{\delta^{\frac{1}{1-\beta}}\beta^{\frac{\beta}{1-\beta}}}{2^{\frac{1}{1-\beta}}}(1-\beta)\left(2(\xi\overline{\gamma} + (1-\xi)\underline{\gamma}) + (\overline{\gamma}-\underline{\gamma})(\xi A(1) + (1-\xi)A(0) - A(\xi))\right)^{\frac{1}{1-\beta}} && (155) \\
&\equiv y + A(\xi). && (156)
\end{aligned}$$

Hence, $A(\xi)$ satisfies

$$A(\xi) = \frac{\delta^{\frac{1}{1-\beta}}\beta^{\frac{\beta}{1-\beta}}}{2^{\frac{1}{1-\beta}}}\frac{1-\beta}{1-\delta}\left(2(\xi\overline{\gamma} + (1-\xi)\underline{\gamma}) + (\overline{\gamma}-\underline{\gamma})(\xi A(1) + (1-\xi)A(0) - A(\xi))\right)^{\frac{1}{1-\beta}} \tag{157}$$

where, from (150),

$$A(1) = A|_{\gamma=\overline{\gamma}} = \delta^{\frac{1}{1-\beta}}\beta^{\frac{\beta}{1-\beta}}\frac{1-\beta}{1-\delta}\overline{\gamma}^{\frac{1}{1-\beta}} \tag{158}$$

$$A(0) = A|_{\gamma=\underline{\gamma}} = \delta^{\frac{1}{1-\beta}}\beta^{\frac{\beta}{1-\beta}}\frac{1-\beta}{1-\delta}\underline{\gamma}^{\frac{1}{1-\beta}}. \tag{159}$$

Plugging (158) and (159) into (152) and (157) yields

$$g^E(y,\xi) = y - \frac{\delta^{\frac{1}{1-\beta}}\beta^{\frac{1}{1-\beta}}}{2^{\frac{1}{1-\beta}}} \cdot \left(2(\xi\overline{\gamma} + (1-\xi)\underline{\gamma}) + (\overline{\gamma}-\underline{\gamma})\left(\delta^{\frac{1}{1-\beta}}\beta^{\frac{\beta}{1-\beta}}\frac{1-\beta}{1-\delta}\left(\xi\overline{\gamma}^{\frac{1}{1-\beta}} + (1-\xi)\underline{\gamma}^{\frac{1}{1-\beta}}\right) - A(\xi)\right)\right)^{\frac{1}{1-\beta}} \tag{160}$$

where, for $\xi \in (0,1)$, $A(\xi)$ is implicitly and uniquely defined by[29]

$$A(\xi) = \frac{\delta^{\frac{1}{1-\beta}}\beta^{\frac{\beta}{1-\beta}}}{2^{\frac{1}{1-\beta}}}\frac{1-\beta}{1-\delta} \cdot \left(2(\xi\overline{\gamma} + (1-\xi)\underline{\gamma}) + (\overline{\gamma}-\underline{\gamma})\left(\delta^{\frac{1}{1-\beta}}\beta^{\frac{\beta}{1-\beta}}\frac{1-\beta}{1-\delta}\left(\xi\overline{\gamma}^{\frac{1}{1-\beta}} + (1-\xi)\underline{\gamma}^{\frac{1}{1-\beta}}\right) - A(\xi)\right)\right)^{\frac{1}{1-\beta}}. \tag{161}$$

Using (161), (160) is rewritten as (96).

C Learning vs. Benchmark Models

In the model with an iso-elastic utility and a linear production, we consider two benchmark cases, the deterministic case in which all parameters are known and the stochastic case in which production depends on a random shock with a *known* distribution. We then study the influence of learning on optimal behavior by comparing the optimal policy between the two benchmark models and the learning environment. This comparison reveals the

[29] To see uniqueness, note that, for any $\xi \in (0,1)$, the left-hand side of (161) is the 45 degree line whereas the right-side evaluated at $A(\xi) = 0$ is strictly positive. Moreover, the right-hand side is decreasing such that it crosses the 45-degree line only once.

profound effect that learning has on the optimal policy functions. For both benchmark cases, the solution is explicit, but the introduction of learning complicates the analysis so that the solution can only be solved implicitly. We also show that the effect of learning on the optimal level of consumption is ambiguous in general. Nonetheless, when the prior distribution of the shock is riskier than the actual distribution of the shock, consumption in the learning environment is greater than consumption in the stochastic case. We first derive the optimal policy functions for both benchmark models. We then compare the benchmark models with the learning model.

C.1 Deterministic Case

In the **deterministic** model, the production shock r has constant value $\overline{r}$ over time. For $\tau = 1, 2, \ldots$, the τ-period-horizon value function in a deterministic (D) case is

$$V_\tau^D(y) = \max_{c \in (0,y)} \left\{ u(c) + \delta V_{\tau-1}^D(f(y-c, \overline{r})) \right\}. \tag{162}$$

Given Assumptions 3.1 and 3.2, (162) is rewritten as

$$V_\tau^D(y) = \max_{c \in (0,y)} \left\{ c^\alpha + \delta V_{\tau-1}^D(\overline{r}(y-c)) \right\} \tag{163}$$

Proposition C.1 provides the optimal policy function for any finite horizon.

Proposition C.1. *In a deterministic environment, for* $\tau = 0, 1, \ldots,$

$$\rho_\tau^D(y) = \frac{y}{\sum_{t=0}^{\tau} \delta^{\frac{\tau}{1-\alpha}} \overline{r}^{\frac{\alpha\tau}{1-\alpha}}}. \tag{164}$$

Proof. Using (163) and the fact that $V_0^D(y) = y^\alpha$,[30] the one-period-horizon

[30]When there is no horizon (i.e., $\tau = 0$), it is optimal to consume the entire stock regardless of the environment faced by the agent.

value function is

$$V_1^D(y) = \max_{c\in(0,y)} \left\{c^\alpha + \delta V_0^D(\overline{r}(y-c))\right\} \tag{165}$$

$$= \max_{c\in(0,y)} \{c^\alpha + \delta\overline{r}^\alpha (y-c)^\alpha\} \tag{166}$$

so that the first-order condition $c^{\alpha-1} - \delta\overline{r}^\alpha(y-c)^{\alpha-1} = 0$ yields

$$\rho_1^D(y) = \frac{y}{1+\delta^{\frac{1}{1-\alpha}}\overline{r}^{\frac{\alpha}{1-\alpha}}}. \tag{167}$$

Plugging (167) into (166) yields

$$V_1^D(y) = \left(1+\delta^{\frac{1}{1-\alpha}}\overline{r}^{\frac{\alpha}{1-\alpha}}\right)^{1-\alpha} y^\alpha. \tag{168}$$

Given that, from (168), the one-period-horizon value function is linear in y^α, we now consider a τ-period horizon where the continuation value function is of the form $V_{\tau-1}^D(y) = \kappa_{\tau-1}^D y^\alpha$ with constant parameter $\kappa_{\tau-1}^D > 0$. For $\tau = 2, 3, \ldots$, the τ-period-horizon value function is

$$V_\tau^D(y) = \max_{c\in(0,y)} \left\{c^\alpha + \delta V_{\tau-1}^D(\overline{r}(y-c))\right\} \tag{169}$$

$$= \max_{c\in(0,y)} \left\{c^\alpha + \delta\kappa_{\tau-1}^D\overline{r}^\alpha (y-c)^\alpha\right\} \tag{170}$$

so that the first-order condition $c^{\alpha-1} - \delta\kappa_{\tau-1}^D\overline{r}^\alpha(y-c)^{\alpha-1} = 0$ yields

$$\rho_\tau^D(y) = \frac{y}{1+\left(\kappa_{\tau-1}^D\right)^{\frac{1}{1-\alpha}}\delta^{\frac{1}{1-\alpha}}\overline{r}^{\frac{\alpha}{1-\alpha}}}. \tag{171}$$

Plugging (171) into (170) yields

$$V_\tau^D(y) = \left(1+\left(\kappa_{\tau-1}^D\right)^{\frac{1}{1-\alpha}}\delta^{\frac{1}{1-\alpha}}\overline{r}^{\frac{\alpha}{1-\alpha}}\right)^{1-\alpha} y^\alpha \tag{172}$$

$$\equiv \kappa_\tau^D y^\alpha \tag{173}$$

so that

$$\kappa_\tau^D = \left(1+\left(\kappa_{\tau-1}^D\right)^{\frac{1}{1-\alpha}}\delta^{\frac{1}{1-\alpha}}\overline{r}^{\frac{\alpha}{1-\alpha}}\right)^{1-\alpha} \tag{174}$$

with, from (168), initial condition

$$\kappa_1^D = \left(1 + \delta^{\frac{1}{1-\alpha}} \overline{r}^{\frac{\alpha}{1-\alpha}}\right)^{1-\alpha}. \tag{175}$$

Solving (174) and imposing the initial condition (175) yields

$$\kappa_\tau^D = \left(\sum_{t=0}^{\tau} \delta^{\frac{t}{1-\alpha}} \overline{r}^{\frac{\alpha t}{1-\alpha}}\right)^{1-\alpha}. \tag{176}$$

Plugging (176) into (171) yields (164). □

Proposition C.2 provides the optimal policy function for an infinite horizon in the deterministic model. As in the learning case, the solution is unique and linear in output. Unlike the learning environment, the deterministic model yields an explicit solution for the optimal policy function.

Proposition C.2. *From* (164), $\lim_{\tau\to\infty} \rho_\tau^D(y) \equiv \rho_\infty^D(y)$ *exists such that*

$$\rho_\infty^D(y) = (1 - \delta^{\frac{1}{1-\alpha}} \overline{r}^{\frac{\alpha}{1-\alpha}})y. \tag{177}$$

C.2 Stochastic Case

In the **stochastic** model, the agent faces uncertainty about the future production shocks and knows the true distribution of $\tilde{r}$, i.e., θ^* is known. For $\tau = 1, 2, \ldots$, the τ-period-horizon value function in a stochastic (S) case is

$$V_\tau^S(y) = \max_{c\in(0,y)} \left\{ u(c) + \delta \int_{\underline{r}}^{\overline{r}} V_{\tau-1}^S(f(y-c, r))\phi(r|\theta^*)\mathrm{d}r \right\}. \tag{178}$$

Given Assumptions 3.1 and 3.2, (178) is rewritten as

$$V_\tau^S(y) = \max_{c\in(0,y)} \left\{ c^\alpha + \delta \int_0^1 V_{\tau-1}^S(r(y-c))\phi(r|\theta^*)\mathrm{d}r \right\}. \tag{179}$$

Proposition C.3 provides the optimal policy function for any finite horizon.

Proposition C.3. *In the stochastic case, for $\tau = 0, 1, \ldots$,*

$$\rho_\tau^S(y) = \frac{y}{\sum_{t=0}^{\tau} \delta^{\frac{\tau}{1-\alpha}} \left(\int_0^1 r^\alpha \phi(r|\theta^*)\mathrm{d}r\right)^{\frac{\tau}{1-\alpha}}}. \tag{180}$$

Proof. Using (179) and the fact that $V_0^S(y) = y^\alpha$, the one-period-horizon value function is

$$V_1^S(y) = \max_{c\in(0,y)} \left\{ c^\alpha + \delta \int_0^1 V_0^S(r(y-c))\phi(r|\theta^*)\mathrm{d}r \right\} \tag{181}$$

$$= \max_{c\in(0,y)} \left\{ c^\alpha + \delta \left(\int_0^1 r^\alpha \phi(r|\theta^*)\mathrm{d}r \right) (y-c)^\alpha \right\} \tag{182}$$

so that the first-order condition $c^{\alpha-1} - \delta \left(\int_0^1 r^\alpha \phi(r|\theta^*)\mathrm{d}r\right)(y-c)^{\alpha-1} = 0$ yields

$$\rho_1^S(y) = \frac{y}{1 + \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \phi(r|\theta^*)\mathrm{d}r\right)^{\frac{1}{1-\alpha}}}. \tag{183}$$

Plugging (183) into (182) yields

$$V_1^S(y) = \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \phi(r|\theta^*)\mathrm{d}r \right)^{\frac{1}{1-\alpha}} \right)^{1-\alpha} y^\alpha. \tag{184}$$

Given that, from (184), the one-period-horizon value function is linear in y^α, we now consider a τ-period horizon where the continuation value function is of the form $V_{\tau-1}^S(y) = \kappa_{\tau-1}^S y^\alpha$ with constant parameter $\kappa_{\tau-1}^S > 0$. For $\tau = 2, 3, \ldots$, the τ-period-horizon value function is

$$V_\tau^S(y) = \max_{c\in(0,y)} \left\{ c^\alpha + \delta \int_0^1 V_{\tau-1}^S(r(y-c))\phi(r|\theta^*)\mathrm{d}r \right\} \tag{185}$$

$$= \max_{c\in(0,y)} \left\{ c^\alpha + \delta\kappa_{\tau-1}^S \left(\int_0^1 r^\alpha \phi(r|\theta^*)\mathrm{d}r \right) (y-c)^\alpha \right\} \tag{186}$$

so that the first-order condition $c^{\alpha-1} - \delta\kappa_{\tau-1}^S \left(\int_0^1 r^\alpha \phi(r|\theta^*)\mathrm{d}r\right)(y-c)^{\alpha-1} = 0$

yields

$$\rho_\tau^S(y) = \frac{y}{1 + \left(\kappa_{\tau-1}^S\right)^{\frac{1}{1-\alpha}} \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \phi(r|\theta^*) \mathrm{d}r\right)^{\frac{1}{1-\alpha}}}. \tag{187}$$

Plugging (187) into (186) yields

$$V_\tau^S(y) = \left(1 + \left(\kappa_{\tau-1}^S\right)^{\frac{1}{1-\alpha}} \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \phi(r|\theta^*) \mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha} y^\alpha \tag{188}$$

$$\equiv \kappa_\tau^S y^\alpha \tag{189}$$

so that

$$\kappa_\tau^S = \left(1 + \left(\kappa_{\tau-1}^S\right)^{\frac{1}{1-\alpha}} \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \phi(r|\theta^*) \mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha} \tag{190}$$

with, from (184), initial condition

$$\kappa_1^S = \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \phi(r|\theta^*) \mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha}. \tag{191}$$

Solving (190) and imposing the initial condition (191) yields

$$\kappa_\tau^S = \left(\sum\nolimits_{t=0}^{\tau} \delta^{\frac{t}{1-\alpha}} \left(\int_0^1 r^\alpha \phi(r|\theta^*) \mathrm{d}r\right)^{\frac{t}{1-\alpha}}\right)^{1-\alpha}. \tag{192}$$

Plugging (192) into (187) yields (180). □

Proposition C.4 provide the optimal consumption for an infinite horizon in the stochastic model. As in the learning case, the solution is unique and linear in output. Unlike the learning case, the stochastic model yields explicit solutions for the optimal policy function.

Proposition C.4. *From* (180), $\lim_{\tau\to\infty} \rho_\tau^S(y) \equiv \rho_\infty^S(y)$ *exists such that*

$$\rho_\infty^S(y) = \left(1 - \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \phi(r|\theta^*) \mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right) y. \tag{193}$$

C.3 Comparisons

First, observe from (177) and (193) that there is certainty equivalence between deterministic and stochastic cases. That is, replacing $\overline{r}^\alpha$ by $\int_0^1 r^\alpha \phi(r|\theta^*)\mathrm{d}r$ in (177) yields (193). Second, while adding uncertainty does not alter the functional form of optimal consumption, it does have an effect on the level of optimal consumption. To see this, suppose that $\overline{r} = \int_0^1 r\phi(r|\theta^*)\mathrm{d}r$. Then, from (177) and (193), $\alpha \in (0,1)$ implies that uncertainty makes future payoffs riskier, which increases present consumption, i.e., $\rho^D_\infty(y) < \rho^S_\infty(y)|_{\int_0^1 r\phi(r|\theta^*)\mathrm{d}r=\overline{r}}$.

Next, we compare optimal behavior between the learning case and the stochastic case (in the context of an iso-elastic utility and linear production) for the infinite horizon.[31] In general, the effect of learning is ambiguous. However, when the true distribution of the random shock second-order stochastically dominates the expected distribution of the random shock (conditional on prior beliefs), then learning increases consumption.

Proposition C.5. *Suppose that the p.d.f.* $\phi(r|\theta^*)$ *second-order stochastically dominates the p.d.f.* $\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta$*. Then, learning increases present consumption, i.e.,* $\rho^L_\infty(y,\xi) > \rho^S_\infty(y)$*.*

Proof. First, since $\phi(r|\theta^*)$ second-order stochastically dominates $\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta$ and $\alpha \in (0,1)$, it follows that

$$\int_0^1 r^\alpha \left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r < \int_0^1 r^\alpha\phi(r|\theta^*)\mathrm{d}r,$$

$$1-\delta^{\frac{1}{1-\alpha}}\left(\int_0^1 r^\alpha \left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right)^{\frac{1}{1-\alpha}} > 1-\delta^{\frac{1}{1-\alpha}}\left(\int_0^1 r^\alpha\phi(r|\theta^*)\mathrm{d}r\right)^{\frac{1}{1-\alpha}}. \tag{195}$$

Second, from Proposition 3.4, $\rho^L_\infty(y,\xi) = \omega_\infty(\xi)y$ such that from the proof

[31] For a one-period horizon, the presence of learning does not alter the functional form of optimal consumption. Indeed, replacing $\overline{r}$ in (167) or $\int_0^1 r^\alpha\phi(r|\theta^*)\mathrm{d}r$ in (183) by

$$\int_0^1 r^\alpha \left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r \tag{194}$$

yields (25).

of Proposition 3.4,

$$\omega_\infty(\xi) = \frac{1}{\kappa_\infty(\xi)^{\frac{1}{1-\alpha}}} \tag{196}$$

where from (121), (122), and (123), $\kappa_\infty(\xi) < M$, M defined by (120). Since $\kappa_\infty(\xi) \equiv \omega_\infty(\xi)^{\alpha-1}$, it follows that

$$\omega_\infty(\xi) > \frac{1}{M^{\frac{1}{1-\alpha}}}. \tag{197}$$

Plugging (120) into (197) yields

$$\omega_\infty(\xi) > 1 - \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right)^{\frac{1}{1-\alpha}}. \tag{198}$$

Combining inequalities (195) and (198) with (193) implies that learning increases present consumption, i.e.,

$$\rho^L_\infty(y,\xi) = \omega_\infty(\xi)y \tag{199}$$

$$> \left(1 - \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right)^{\frac{1}{1-\alpha}} \right) y \tag{200}$$

$$> \left(1 - \delta^{\frac{1}{1-\alpha}} \left(\int_0^1 r^\alpha \phi(r|\theta^*)\mathrm{d}r \right)^{\frac{1}{1-\alpha}} \right) y \tag{201}$$

$$= \rho^S_\infty(y). \tag{202}$$

□

References

P. Aghion, P. Bolton, C. Harris, and B. Jullien. Optimal Learning by Experimentation. *Rev. Econ. Stud.*, 58(4):621–654, 1991.

R.J. Balvers and T.F. Cosimano. Actively Learning About Demand and the Dynamics of Price Adjustment. *Econ. J.*, 100(402):882–898, 1990.

G. Bertocchi and M. Spagat. Growth under Uncertainty with Experimentation. *J. Econ. Dynam. Control*, 23(2):209–231, 1998.

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

A. Creane. Experimentation with Heteroskedastic Noise. *Econ. Theory*, 4 (2):275–286, 1994.

M. Datta, L.J. Mirman, and E.E. Schlee. Optimal Experimentation in Signal-Dependent Decision Problems. *Int. Econ. Rev.*, 43(2):577–607, 2002.

D. Easley and N.M. Kiefer. Controlling a Stochastic Process with Unknown Parameters. *Econometrica*, 56(5):1045–1064, 1988.

D. Easley and N.M. Kiefer. Optimal Learning with Endogenous Data. *Int. Econ. Rev.*, 30(4):963–978, 1989.

M.A. El-Gamal and R.K. Sundaram. Bayesian Economists ... Bayesian Agents: An Alternative Approach to Optimal Learning. *J. Econ. Dynam. Control*, 17(3):355–383, 1993.

A. Fishman and N. Gandal. Experimentation and Learning with Network Effects. *Econ. Letters*, 44(1–2):103–108, 1994.

J.M. Fusselman and L.J. Mirman. Experimental Consumption for a General Class of Disturbance Densities. In R. Becker, M. Boldrin, R. Jones, and M. Thompson, editors, *General Equilibrium, Growth, and Trade. The Legacy of Lionel McKenzie. Economic Theory, Econometrics, and Mathematical Economics Series*, volume 2, pages 367–392. Academy Press, London, 1993.

S.J. Grossman, R.E. Kihlstrom, and L.J. Mirman. A Bayesian Approach to the Production of Information and Learning by Doing. *Rev. Econ. Stud.*, 44(3):533–547, 1977.

H.A. Hopenhayn and E.C. Prescott. Stochastic Monotonicity and Stationary Distributions for Dynamic Economies. *Econometrica*, 60(6):1387–1406, 1992.

G.W. Huffman and N.M. Kiefer. Optimal Learning in a One-Sector Growth Model. Working Papers 90-24, Cornell University, Center for Analytic Economics, 1994.

G. Keller and S. Rady. Optimal Experimentation in a Changing World. *Rev. Econ. Stud.*, 66(3):475–507, 1999.

N.M. Kiefer and Y. Nyarko. Optimal Control of an Unknown Linear Process with Learning. *Int. Econ. Rev.*, 30(3):571–586, 1989.

C. Koulovatianos, L.J. Mirman, and M. Santugini. Optimal Growth and Uncertainty: Learning. *J. Econ. Theory*, 144(1):280–295, 2009.

L.J. Mirman. *Two Essays on Uncertainty and Economics*. PhD thesis, University of Rochester, 1970.

L.J. Mirman and I. Zilcha. On Optimal Growth under Uncertainty. *J. Econ. Theory*, 11(3):329–339, 1975.

L.J. Mirman and I. Zilcha. Characterizing Optimal Policies in a One-Sector Model of Economic Growth under Uncertainty. *J. Econ. Theory*, 14(2): 389–401, 1977.

L.J. Mirman, L. Samuelson, and A. Urbano. Monopoly Experimentation. *Int. Econ. Rev.*, 34(3):549–563, 1993.

L.J. Mirman, O. Morand, and K. Reffett. A Qualitative Approach to Markovian Equilibrium in Infinite Horizon Economies with Capital. *J. Econ. Theory*, 139(1):75–98, 2008.

E.C. Prescott. The Multi-Period Control Problem Under Uncertainty. *Econometrica*, 40(6):1043–1058, 1972.

D. Trefler. The Ignorant Monopolist: Optimal Learning with Endogenous Information. *Int. Econ. Rev.*, 34(3):565–581, 1993.

V. Wieland. Learning by Doing and the Value of Optimal Experimentation. *J. Econ. Dynam. Control*, 24(4):501–534, 2000.