CIRPÉE

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 Reffett
Marc Santugini

Octobre/October 2013

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

Abstract:
We study the effect of learning on optimal growth. We first derive the Euler equation in a general learning environment without experimentation. We then consider the case of iso-elastic utility and linear production, for general distributions of the random shocks and beliefs (i.e., no conjugate priors) and for any horizon. We characterize the unique optimal policy function for this *learning* model. We show how learning alters the maximization problem of the social planner. We also compare the learning model with the deterministic and stochastic models.
This work builds on the work on learning and growth in a Brock-Mirman environment initiated by Koulovatianos, Mirman, and Santugini (2009) (KMS) for the Mirman-Zilcha model (with log utility and Cobb-Douglas production). While the Mirman-Zilcha model provides some insights about the effect of learning on growth, it also hides many important features of learning that the model in this paper takes account of. In other words, compared to the Mirman-Zilcha model, we show that the case of iso-elastic utility and linear production yields a more profound effect of learning on dynamic programming and thus optimal behavior.

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

JEL Classification: D8, D9, E2

1 Introduction

Uncertainty is ubiquitous to virtually all economic problems beginning with growth and real business cycles in macroeconomics and continuing with industrial organization and consumer behavior in microeconomics. Indeed, economic agents make optimal decisions without complete knowledge of the environment in which they live. This is particularly relevant to dynamic maximization problems in which a myriad of future variables is unobserved by the decision makers. To analyze optimal behavior under uncertainty, random shocks are included in the objective functions and the constraints.[1] While the agents have no knowledge of the realized shocks, they know their distributions and thus use this knowledge to form expectations over the sum of present and discounted future payoffs subject to constraints.

Uncertainty is particularly relevant in optimal growth. Although agents have a certain control over the evolution of capital such as infrastructures, roads, telecommunications, energy, and common-pool natural resources, the dynamics of capital remains highly uncertain. In the stochastic growth models initially studied in Brock and Mirman (1972) and Mirman and Zilcha (1975), the social planner makes consumption and saving decisions taking account of uncertainty by forming expectations. However, economic agents have the ability to do more than just react to uncertainty. In many cases, they can also alter the uncertainty they face through learning. That is, agents learn about the structure of the economy in order to reduce the uncertainty they face. For instance, suppose that in addition to not observing future shocks, the agents ignore the true distributions generating these shocks. In that case, observing past shocks provides information about these unknown distributions. Hence, the agents not only make decisions of consumption and saving, but at the same time they engage in econometric activities, gathering and analyzing data in order to learn about unknown variables, and, thus, reducing the uncertainty they face. In general, decision making and learning are nonseparable and influence each other.[2]

[1] Random shocks can be embedded in positive models as well. See Mirman (1970) for an early analysis of uncertainty (i.e., random production function) in the Solow model.

[2] There is a two-way interaction between decision making and learning. On the one

Unlike the literature on stochastic optimal growth models, there is little work on *learning* optimal growth models.[3] One exception concerns the study of capital accumulation when the agents have the ability to experiment. However, these models consider at most a three-period horizon or rely heavily on the use of conjugate priors, especially the normal distribution.[4] Recently, Koulovatianos, Mirman, and Santugini (2009) (KMS) provides closed-form solutions for the social planner's optimal policy function in the Mirman-Zilcha class of models when the agents learn about the production function.[5] In the learning growth model studied in KMS, there is no experimentation. Indeed, the full problem (with experimentation) has yet to be solved or studied in optimal growth with infinite horizon. Hence, understanding the problem of learning takes several steps. KMS makes the assumption that the signal is seen so experimentation is not relevant. Yet, the learning activity changes future payoffs.[6] Optimal behavior is characterized for general distributions of the random shock and beliefs (i.e., no use of conjugate priors). The class of Mirman-Zilcha models with log utility and Cobb-Douglas production functions offers a preliminary insight of the effect of learning on optimal growth. Although learning is shown to have a profound effect on the social planner's optimal policy function (there is no equivalence in the function form between the stochastic case and the learning case), it turns out that the log case combined with the Cobb-Douglas

hand, decision making may have an effect on learning, which is referred as experimentation. On the other hand, the presence of learning adds risk which affects future payoffs and thus behavior.

[3]There is however a large literature that has focused on learning in dynamic programming but *abstracting* from the evolution of capital. This was largely studied in the context of 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).

[4]See Bertocchi and Spagat (1998), Datta et al. (2002), El-Gamal and Sundaram (1993), Huffman and Kiefer (1994), and Beck and Wieland (2002).

[5]For non-optimal models, e.g., bounded rationality, there is a literature on adaptive learning. See Evans and Honkapohja (2001).

[6]Even with iid shocks in production, a learning environment implies that the agents face 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.

function removes some of the effects of learning. The reason is that in the Mirman-Zilcha class of models, part of the effect of learning is found in the constant term of the value function, which has no effect on optimal behavior. In this paper, we extend further the analysis of the effect of learning (without experimentation) on another more general class of growth models with iso-elastic utility and linear production functions. The social planner faces multiplicative uncertainty in production and does not know the true distribution of the production shock.

To understand the effect of learning, we do not conjecture and verify the value function as previously done in KMS. Rather, we solve for optimal behavior recursively, which sheds light on how learning alters the social planner's maximization problem. We show that there is a unique solution for optimal behavior for every finite horizon. We also show that the limit exists, yielding a unique solution for the infinite horizon. To clarify the effect of learning, we also 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 compare the benchmark models with the learning model. In particular, with unbiased beliefs about the mean of the production shock, learning increases consumption.

Finally, we compare the effect of learning between the Mirman-Zilcha class of models studied in KMS and the class of models studied in this paper. In general, the effect of learning is two-fold. First, there is a direct effect due to the anticipation of the planner about the stochastic effect of today's production shock on tomorrow's stock as well as the stochastic effect of today's production shock on tomorrow's expectations about the next period production shock. Second, there is an indirect effect of learning that the agent anticipates (stochastically) the effect of observing the production shock on posterior beliefs through future optimal decisions, i.e., what the planner alters behavior upon observing the shock and updating beliefs. In the Mirman-Zilcha class of models, only the first direct effect influences optimal behavior whereas both direct and indirect effects matter in the class of models with iso-elastic utility and linear production functions.

The paper is organized as follows. In section 2, we introduce learning in a general Brock-Mirman environment and derive the Euler equation corresponding to learning growth. Section 3 provides the optimal policy functions for the specific class of models corresponding to the deterministic and stochastic (benchmark) models. Section 4 provides the analysis for the learning growth models. Finally, in Section 5, we compare the effect of learning on the maximization problem between our model (with iso-elastic utility and linear production) and the Mirman-Zilcha model (with log utility and Cobb-Douglas production).

2 Model

In this section, we present the Brock-Mirman environment under a learning environment. We also consider two benchmark models, the deterministic and stochastic environments. For each environment, we derive and compare the Euler equations. In the subsequent sections, we study the effect of learning on dynamic programming in the case of an iso-elastic utility and a linear production.

2.1 Preliminaries

Consider a Brock-Mirman environment in which, in period $t = 0, 1, ...,$ a social planner divides output y_t between consumption c_t and investment $k_t = y_t - c_t$. Investment k_t is then used for the production of the output in the subsequent period, i.e.,

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

where $f(k_t, r_t)$ is the production function with the usual neoclassical properties[7] and r_t is a realization of the random production shock $\tilde{r}_t$ with p.d.f $\phi(r_t|\theta^*), r_t \in \mathcal{H} \subset \mathbb{R}$. The p.d.f. depends on a parameter $\theta^* \in \Theta \subset \mathbb{R}^N, N \in \mathbb{N}$. The distribution of r_t is parametric and fully characterized by the vector

[7]Namely, $f(\cdot, r_t)$ is an increasing concave differentiable function with $f(0, r_t) = 0$ for all r_t. The Inada conditions are also assumed, $f_1(0, r_t) = \infty$ and $f_1(\infty, r_t) = 0$ for all r_t. Finally, $f(k_t, \cdot)$ is an increasing differentiable function for all k_t.

θ^*.

In order to study the effect of learning, we consider the *deterministic* and the *stochastic* benchmark environments. Before proceeding with the analysis, we describe the three environments. Regardless of the environment faced by the planner, his objective is to maximize the expected sum of discounted utilities where the discount factor is $\delta \in (0,1)$ and the utility function is $u(c_t), u' > 0, u'' < 0$ with $u'(0) = \infty$. 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 $\hat{y} = f(y-c, r)$ is output tomorrow when today's production shock is r. To distinguish among different horizons of the dynamic program, we use the index $\tau = 0, 1, \ldots, \infty$.

2.2 Benchmark Models

In the **deterministic environment**, today's production shock r is known to have constant value $\bar{r}$. For $\tau = 1, 2, \ldots$, the τ-period-horizon value function in a deterministic (D) environment is

$$V_\tau^D(y; \bar{r}) = \max_{c \in [0,y]} \left\{ u(c) + \delta V_{\tau-1}^D(f(y-c, \bar{r}); \bar{r}) \right\}. \tag{2}$$

The maximum is obtained at a unique point $c = \rho_\tau^D(y; \bar{r})$ since the maximand is strictly concave. Moreover, $\rho_\tau^D(y; \bar{r}) \in (0, y)$ since $u'(0) = \infty$. For $\tau = \infty$, the optimal policy $c^D(y) \equiv \rho_\infty^D(y; \bar{r})$ satisfies the Euler equation

$$u'(c^D(y)) = \delta f_1(y - c^D(y), \bar{r}) \cdot u'(c^D(\hat{y}^D)), \tag{3}$$

where $\hat{y}^D \equiv f(y - c^D(y), \bar{r})$.

In the **stochastic environment**, the planner faces uncertainty about the future production shocks while knowing 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) environment is

$$V_\tau^S(y; \theta^*) = \max_{c \in [0,y]} \left\{ u(c) + \delta \int_{r \in \mathcal{H}} V_{\tau-1}^S(f(y-c, r); \theta^*) \phi(r|\theta^*) \mathrm{d}r \right\}. \tag{4}$$

The maximum is obtained at a unique point $c = \rho^S_\tau(y;\theta^*)$ since the maximand is strictly concave. Moreover, $\rho^S_\tau(y;\theta^*) \in (0, y)$ since $u'(0) = \infty$. From Mirman and Zilcha (1975), for $\tau = \infty$, the optimal policy $c^S(y) \equiv \rho^S_\infty(y;\theta^*)$ satisfies the Euler equation

$$u'(c^S(y)) = \delta \int_{r\in\mathcal{H}} f_1(y - c^S(y), r) \cdot u'(c^S(\hat{y}^S(r)))\phi(r|\theta^*)\mathrm{d}r, \tag{5}$$

where $\hat{y}^S(r) \equiv f(y - c^S(y), r)$.

2.3 Learning Model

Having presented the benchmark models, we now described our dynamic model with learning. In the **learning environment**, the planner faces uncertainty about future production shocks as well as uncertainty about the true distribution of production shocks, i.e., the parameter θ^* is unknown to the planner. While the planner does not know θ^*, observing past production shocks provides information about the true distribution, which is used to update beliefs via Bayesian methods. Given today's prior beliefs about θ^* expressed as a prior p.d.f. ξ on Θ and the observation r,[8] tomorrow's posterior beliefs are

$$\hat{\xi}(\theta|r) = \frac{\phi(r|\theta)\xi(\theta)}{\int_{x\in\Theta}\phi(r|x)\xi(x)\mathrm{d}x}, \tag{6}$$

$\theta \in \Theta$, by Bayes' Theorem. Since there is a one-to-one relationship between output and the shock (i.e., $f_2 > 0$), observing the production shock is equivalent to observing output. In other words, there is no active learning (or experimentation). Under active learning (or experimentation), the planner's decision has an effect on the information, i.e., posterior beliefs depend on the decision.[9] While we consider passive learning in which the planner's decision has no effect on the information used to learn about the unknown parameter, the presence of passive learning in dynamic models increases risk in future

[8] That is, given prior beliefs ξ, the probability that $\theta^* \in S$ is $\int_{\theta\in S}\xi(\theta)\mathrm{d}\theta$ for any $S \subset \Theta$.
[9] See Mirman, Samuelson, and Urbano (1993) for a model with active learning.

payoffs, which alters behavior.

The planner makes consumption decision, while learning about θ^*. That is, endowed with initial stock and beliefs, consumption is chosen. The production shock r is then realized and the output, in the subsequent period, is determined from (1). Information is gleaned from observing r, which, from (6), affects beliefs about θ^*. For $\tau = 1, 2, \ldots$, the τ-period-horizon value function in a learning (L) environment is

$$V_\tau^L(y;\xi) = \max_{c\in[0,y]} \left\{ u(c) + \delta \int_{r\in\mathcal{H}} V_{\tau-1}^L(f(y-c,r);\hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \tag{7}$$

where $\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta$ is the expected p.d.f. of the production shock given prior beliefs.

In addition to anticipating the effect of the consumption decision on future output, the planner anticipates learning. In a dynamic and learning context, rational expectations imply that the information contained in the future production shock is anticipated. The anticipation of learning is integrated into (7) by anticipating the updated beliefs from ξ to $\hat{\xi}(\cdot|r)$ using (6). While learning is passive, the evolution of beliefs must be taken into account in dynamic programming. Bayesian dynamics complicates the maximization problem because the planner makes consumption and investment decisions, anticipating updating beliefs every period. That is, the continuation value function $V_{\tau-1}^L(f(y-c,r),\hat{\xi}(\cdot|r))$ in (7) encompasses beliefs that have been updated many times, and in the infinite-horizon case *infinitely* many times.

The maximum is obtained at a unique point $c = \rho_\tau^L(y;\xi)$ since the maximand is strictly concave. Moreover, there cannot be any corner solutions, i.e., $\rho_\tau^L(y;\xi) \in (0,y)$ since $u'(0) = \infty$. Thus, for the infinite horizon, all programs from any initial point never exhaust the stock and are infinite. Focusing on the infinite horizon, Lemma 2.1 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 to derive the Euler equation under learning in Proposition 2.2. The proof follows closely the proof of Lemma 1 in Mirman and Zilcha (1975) because learning is passive and a change in today's consumption with no corresponding change in investment has no effect on the type of information that the agent either anticipates acquiring or actually acquires from observing future production shocks.

Lemma 2.1. *$\partial V_{\infty}^{L}(y;\xi)/\partial y$ exists for all y and ξ and*

$$\frac{\partial V_{\infty}^{L}(y;\xi)}{\partial y} = u'(\rho_{\infty}^{L}(y;\xi)). \tag{8}$$

Proposition 2.2 states the Euler equation in the learning environment.

Proposition 2.2. *For all y and ξ, the optimal policy $\rho_{\infty}^{L}(y;\xi)$ satisfies the Euler equation*

$$u'(\rho_{\infty}^{L}(y;\xi)) = \delta \int_{r\in\mathcal{H}} f_1(y-\rho_{\infty}^{L}(y;\xi), r)\cdot u'(\rho_{\infty}^{L}(\hat{y}^{L}(r);\hat{\xi}(\cdot|r))) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r, \tag{9}$$

where $\hat{y}^{L}(r) \equiv f(y-\rho_{\infty}^{L}(y;\xi), r)$.

Proof. The first-order condition corresponding to (7) is

$$u'(c) = \delta \int_{r\in\mathcal{H}} f_1(y-c, r) \frac{\partial V_{\tau}^{L}(\hat{y}, \hat{\xi}(\cdot|r))}{\partial \hat{y}} \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \tag{10}$$

evaluated at $c = \rho_{\infty}^{L}(y;\xi)$. Since (8) holds for all y and ξ, $\frac{\partial V_{\infty}^{L}(\hat{y},\hat{\xi}(\cdot|r))}{\partial \hat{y}} = u'(\rho_{\infty}^{L}(y;\xi))$ yielding (9). □

Observe that the changes between the Euler equations for the benchmark models (defined by (3) and (5)) and the Euler equation under learning (defined by (9)) are subtle but important. In particular, if we compare the stochastic and learning cases, learning does more than changing the distribution of the production shock from the true distribution used in (5) to the believed distribution used in (9). It also alters the marginal utility evaluated at tomorrow's consumption through the randomness of future beliefs, i.e., the term $u'(\rho_{\infty}^{L}(\hat{y}^{L}(r);\hat{\xi}(\cdot|r)))$ in (9). Indeed, anticipating learning through the posterior beliefs embedded in the Euler equation implies that the dynamics in output and beliefs are entwined through the production shock as shown in (9).

In order to study the effect of learning on optimal growth, we make further assumptions on the utility and production functions but retain general distributions of the production shock and 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. In the remainder of the paper, we study the class of optimal stochastic growth models with iso-elasticity utility function and linear production function under multiplicative uncertainty. It turns out that for this class of models an implicit solution can be characterized and is valid for a wide range of priors, even those that are outside of families of distributions that are closed under sampling.

Assumption 2.3. *The utility function is iso-elastic:* $u(c) = c^{\alpha}, \alpha \in (0, 1)$.

Assumption 2.4. *The production function is linear:* $f(k, r) = rk, r > 0$.

Given Assumptions 2.3 and 2.4, the Euler equation can be used to derive the infinite-horizon optimal policy function for any environment. Beginning with the benchmark models, from (3) and (5), optimal policy functions are linear in y, i.e., $\rho^D_\infty(y; \bar{r}) = \omega^D_\infty(\bar{r})y$ and $\rho^S_\infty(y; \theta^*) = \omega^S_\infty(\theta^*)y$, where

$$\omega^D_\infty(\bar{r}) = 1 - \delta^{\frac{1}{1-\alpha}} \bar{r}^{\frac{\alpha}{1-\alpha}}, \tag{11}$$

$$\omega^S_\infty(\theta^*) = 1 - \delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} r^{\alpha} \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right)^{\frac{1}{1-\alpha}}. \tag{12}$$

For the learning case, plugging the linear solution $\rho^L_\infty(y; \xi) = \omega^L_\infty(\xi)y$ into (9) yields an implicit solution for $\omega^L_\infty(\xi)$:

$$\frac{\omega^L_\infty(\xi)^{\alpha-1}}{(1 - \omega^L_\infty(\xi))^{\alpha-1}} = \delta \int_{r\in\mathcal{H}} r^{\alpha} \omega^L_\infty(\hat{\xi}(\cdot|r))^{\alpha-1} \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r. \tag{13}$$

However, the Euler equation method hides all the intricacies of learning and prevents a thorough analysis of the effect of learning on dynamic programming. In particular, it is not clear whether (13) is consistent with the limit of the finite programs. To study the effect of learning on behavior, we proceed as follows. Section 3 presents optimal behavior for the benchmark models. The step-by-step analysis is necessary to understand how learning alters dynamic programming, which is discussed in Section 4.

3 Benchmark Models

3.1 Deterministic Model

In a deterministic environment, given Assumptions 2.3 and 2.4, (2) is rewritten as

$$V_\tau^D(y;\bar{r}) = \max_{c\in(0,y)} \left\{c^\alpha + \delta V_{\tau-1}^D(\bar{r}(y-c);\bar{r})\right\}. \tag{14}$$

Using (14) and the fact that $V_0^D(y;\bar{r}) = y^\alpha$,[10] the one-period-horizon value function is

$$V_1^D(y;\bar{r}) = \max_{c\in(0,y)} \left\{c^\alpha + \delta V_0^D(\bar{r}(y-c);\bar{r})\right\}, \tag{15}$$

$$= \max_{c\in(0,y)} \left\{c^\alpha + \delta\bar{r}^\alpha(y-c)^\alpha\right\}, \tag{16}$$

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

$$\rho_1^D(y;\bar{r}) = \frac{y}{1+\delta^{\frac{1}{1-\alpha}}\bar{r}^{\frac{\alpha}{1-\alpha}}}. \tag{17}$$

Plugging (17) back into (16) yields

$$V_1^D(y;\bar{r}) = \left(1+\delta^{\frac{1}{1-\alpha}}\bar{r}^{\frac{\alpha}{1-\alpha}}\right)^{1-\alpha} y^\alpha. \tag{18}$$

Given that, from (18), 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 $V_{\tau-1}^D(y;\bar{r}) = \kappa_{\tau-1}^D y^\alpha$ with constant parameter $\kappa_{\tau-1}^D > 0$. For

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

$\tau = 2, 3, \ldots$, the τ-period-horizon value function is

$$V_\tau^D(y;\bar{r}) = \max_{c\in(0,y)} \left\{c^\alpha + \delta V_{\tau-1}^D(\bar{r}(y-c);\bar{r})\right\}, \tag{19}$$

$$= \max_{c\in(0,y)} \left\{c^\alpha + \delta \kappa_{\tau-1}^D \bar{r}^\alpha (y-c)^\alpha\right\}, \tag{20}$$

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

$$\rho_\tau^D(y;\bar{r}) = \frac{y}{1 + \left(\kappa_{\tau-1}^D\right)^{\frac{1}{1-\alpha}} \delta^{\frac{1}{1-\alpha}} \bar{r}^{\frac{\alpha}{1-\alpha}}}. \tag{21}$$

Plugging (21) back into (20) yields

$$V_\tau^D(y;\bar{r}) = \left(1 + \left(\kappa_{\tau-1}^D\right)^{\frac{1}{1-\alpha}} \delta^{\frac{1}{1-\alpha}} \bar{r}^{\frac{\alpha}{1-\alpha}}\right)^{1-\alpha} y^\alpha, \tag{22}$$

$$\equiv \kappa_\tau^D y^\alpha, \tag{23}$$

so that

$$\kappa_\tau^D = \left(1 + \left(\kappa_{\tau-1}^D\right)^{\frac{1}{1-\alpha}} \delta^{\frac{1}{1-\alpha}} \bar{r}^{\frac{\alpha}{1-\alpha}}\right)^{1-\alpha} \tag{24}$$

with, from (18), initial condition

$$\kappa_1^D = \left(1 + \delta^{\frac{1}{1-\alpha}} \bar{r}^{\frac{\alpha}{1-\alpha}}\right)^{1-\alpha}. \tag{25}$$

Proposition 3.1 provides the optimal policy function for any finite horizon.

Proposition 3.1. *In a deterministic environment, for $\tau = 0, 1, ...$,*

$$\rho_\tau^D(y;\bar{r}) = \frac{y}{\sum_{t=0}^{\tau} \delta^{\frac{\tau}{1-\alpha}} \bar{r}^{\frac{\alpha\tau}{1-\alpha}}}. \tag{26}$$

Proof. Solving (24) and imposing the initial condition (25) yields

$$\kappa_\tau^D = \left(\sum_{t=0}^{\tau} \delta^{\frac{\tau}{1-\alpha}} \bar{r}^{\frac{\alpha\tau}{1-\alpha}}\right)^{1-\alpha}. \tag{27}$$

Plugging (27) back into (21) yields (26). ☐

Proposition 3.2 provides the optimal policy function for an infinite horizon.

Proposition 3.2. *Suppose that* $\bar{r}^{\frac{\alpha}{1-\alpha}} \in (0,1)$*. Then, from* (26), $\lim_{\tau\to\infty} \rho^D_\tau(y;\bar{r}) \equiv \rho^D_\infty(y;\bar{r})$ *exists and*

$$\rho^D_\infty(y;\bar{r}) = (1-\delta^{\frac{1}{1-\alpha}}\bar{r}^{\frac{\alpha}{1-\alpha}})y. \tag{28}$$

3.2 Stochastic Model

In a stochastic environment, given Assumptions 2.3 and 2.4, (4) is rewritten as

$$V^S_\tau(y;\theta^*) = \max_{c\in(0,y)} \left\{ c^\alpha + \delta \int_{r\in\mathcal{H}} V^S_{\tau-1}(r(y-c);\theta^*)\phi(r|\theta^*)\mathrm{d}r \right\}, \tag{29}$$

Using (29) and the fact that $V^S_0(y;\theta^*) = y^\alpha$, the one-period-horizon value function is

$$\begin{aligned} V^S_1(y;\theta^*) &= \max_{c\in(0,y)} \left\{ c^\alpha + \delta \int_{r\in\mathcal{H}} V^S_0(r(y-c);\theta^*)\phi(r|\theta^*)\mathrm{d}r \right\}, && (30)\\ &= \max_{c\in(0,y)} \left\{ c^\alpha + \delta \left(\int_{r\in\mathcal{H}} r^\alpha \phi(r|\theta^*)\mathrm{d}r \right)(y-c)^\alpha \right\} && (31) \end{aligned}$$

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

$$\rho^S_1(y;\theta^*) = \frac{y}{1+\delta^{\frac{1}{1-\alpha}}\left(\int_{r\in\mathcal{H}} r^\alpha\phi(r|\theta^*)\mathrm{d}r\right)^{\frac{1}{1-\alpha}}}. \tag{32}$$

Plugging (32) back into (31) yields

$$V^S_1(y;\theta^*) = \left(1+\delta^{\frac{1}{1-\alpha}}\left(\int_{r\in\mathcal{H}} r^\alpha\phi(r|\theta^*)\mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha} y^\alpha. \tag{33}$$

Given that, from (33), 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 $V^S_{\tau-1}(y;\theta^*) = \kappa^S_{\tau-1}y^\alpha$ with constant parameter $\kappa^S_{\tau-1} > 0$. For

$\tau = 2, 3, \ldots$, the τ-period-horizon value function is

$$V_\tau^S(y;\theta^*) = \max_{c\in(0,y)} \left\{ c^\alpha + \delta \int_{r\in\mathcal{H}} V_{\tau-1}^S(r(y-c);\theta^*)\phi(r|\theta^*)\mathrm{d}r \right\}, \tag{34}$$

$$= \max_{c\in(0,y)} \left\{ c^\alpha + \delta\kappa_{\tau-1}^S \left(\int_{r\in\mathcal{H}} r^\alpha \phi(r|\theta^*)\mathrm{d}r \right) (y-c)^\alpha \right\}, \tag{35}$$

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

$$\rho_\tau^S(y;\theta^*) = \frac{y}{1 + \left(\kappa_{\tau-1}^S\right)^{\frac{1}{1-\alpha}} \delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} r^\alpha\phi(r|\theta^*)\mathrm{d}r\right)^{\frac{1}{1-\alpha}}}. \tag{36}$$

Plugging (36) back into (35) yields

$$V_\tau^S(y;\theta^*) = \left(1 + \left(\kappa_{\tau-1}^S\right)^{\frac{1}{1-\alpha}} \delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} r^\alpha\phi(r|\theta^*)\mathrm{d}r \right)^{\frac{1}{1-\alpha}} \right)^{1-\alpha} y^\alpha, \tag{37}$$

$$\equiv \kappa_\tau^S y^\alpha, \tag{38}$$

so that

$$\kappa_\tau^S = \left(1 + \left(\kappa_{\tau-1}^S\right)^{\frac{1}{1-\alpha}} \delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} r^\alpha\phi(r|\theta^*)\mathrm{d}r \right)^{\frac{1}{1-\alpha}} \right)^{1-\alpha}, \tag{39}$$

with, from (33), initial condition

$$\kappa_1^S = \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} r^\alpha\phi(r|\theta^*)\mathrm{d}r \right)^{\frac{1}{1-\alpha}} \right)^{1-\alpha}. \tag{40}$$

Proposition 3.3 provides the optimal policy function for any finite horizon.

Proposition 3.3. *Suppose that $\int_{r\in\mathcal{H}} r^\alpha\phi(r|\theta^*)\mathrm{d}r < \infty$. In a stochastic environment, for $\tau = 0, 1, \ldots,$*

$$\rho_\tau^S(y;\theta^*) = \frac{y}{\sum_{t=0}^{\tau} \delta^{\frac{\tau}{1-\alpha}} \left(\int_{r\in\mathcal{H}} r^\alpha\phi(r|\theta^*)\mathrm{d}r\right)^{\frac{\tau}{1-\alpha}}}. \tag{41}$$

Proof. Solving (39) and imposing the initial condition (40) yields

$$\kappa_\tau^S = \left(\sum\nolimits_{t=0}^{\tau} \delta^{\frac{\tau}{1-\alpha}} \left(\int_{r\in\mathcal{H}} r^\alpha \phi(r|\theta^*)\mathrm{d}r\right)^{\frac{\tau}{1-\alpha}}\right)^{1-\alpha} \tag{42}$$

Plugging (42) back into (36) yields (41). □

Proposition 3.4 provides the optimal policy function for an infinite horizon.

Proposition 3.4. *Suppose that* $\int_{r\in\mathcal{H}} r^\alpha\phi(r|\theta^*)\mathrm{d}r \in (0,1)$*. Then, from* (41)*,* $\lim_{\tau\to\infty}\rho_\tau^S(y;\theta^*) \equiv \rho_\infty^S(y;\theta^*)$ *exists and*

$$\rho_\infty^S(y;\theta^*) = \left(1 - \delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} r^\alpha \phi(r|\theta^*)\mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right) y. \tag{43}$$

Before proceeding with the learning environment, we compare optimal policy functions between deterministic and stochastic environments. First, from (21) and (36) (or (28) and (43)), there is some sort of certainty equivalence between deterministic and stochastic environments. That is, replacing $\bar{r}^\alpha$ by $\int_{r\in\mathcal{H}} r^\alpha\phi(r|\theta^*)\mathrm{d}r$ in (21) yields (36). Second, while adding uncertainty does not alter the functional form of the policy, risk does have an effect on the optimal amount consumed. To see this, suppose that $\bar{r} = \int_{r\in\mathcal{H}} r\phi(r|\theta^*)\mathrm{d}r$. Then, from (21) and (36) and using the fact that $\alpha \in (0,1)$, uncertainty makes future payoffs riskier, which increases present consumption, i.e., for $\tau = 1, 2, \ldots, \infty$, $\rho_\tau^D(y;\bar{r}) < \rho_\tau^S(y;\theta^*)|_{\int_{r\in\mathcal{H}} r\phi(r|\theta^*)\mathrm{d}r=\bar{r}}$.

4 Learning Model

Having fully characterized the optimal behavior under deterministic and stochastic environments, we turn to the learning model. By considering finite horizons, we show explicitly how learning alters the maximization problem.

In a learning environment, given Assumptions 2.3 and 2.4, (7) is rewritten as

$$V_{\tau}^{L}(y;\xi) = \max_{c\in(0,y)} \left\{ c^{\alpha} + \delta \int_{r\in\mathcal{H}} V_{\tau-1}^{L}(r(y-c);\hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \tag{44}$$

Using (44) and the fact that $V_0^L(y;\xi) = y^{\alpha}$, the one-period-horizon value function is the one-period-horizon value function is

$$V_{1}^{L}(y;\xi) = \max_{c\in(0,y)} \left\{ c^{\alpha} + \delta \int_{r\in\mathcal{H}} V_{0}^{L}(r(y-c);\hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \tag{45}$$

$$= \max_{c\in(0,y)} \left\{ c^{\alpha} + \delta \left(\int_{r\in\mathcal{H}} r^{\alpha} \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right) (y-c)^{\alpha} \right\}, \tag{46}$$

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

$$\rho_1^L(y;\xi) = \frac{y}{1 + \delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} r^{\alpha} \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right)^{\frac{1}{1-\alpha}}} \tag{47}$$

Plugging (47) back into (46) yields

$$V_1^L(y;\xi) = \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} r^{\alpha} \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right)^{\frac{1}{1-\alpha}} \right)^{1-\alpha} y^{\alpha}. \tag{48}$$

For a one-period horizon, the presence of learning does not alter the functional form of the policy function. Indeed, replacing $\bar{r}$ in (17) or $\int_{r\in\mathcal{H}} r^{\alpha}\phi(r|\theta^*)\mathrm{d}r$ in (32) by

$$\int_{r\in\mathcal{H}} r^{\alpha} \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \tag{49}$$

yields (47). However, for higher horizon, learning considerably alters the maximization problem of the planner. To see this, using (48), the two-period-horizon value function is

$$
\begin{aligned}
&V_2^L(y;\xi)\\
&= \max_{c\in(0,y)} \left\{ c^\alpha + \delta \int_{r\in\mathcal{H}} V_1^L(r(y-c);\hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \right\}, \qquad (50)\\
&= \max_{c\in(0,y)} \Bigg\{ c^\alpha\\
&\quad + \delta \int_{r\in\mathcal{H}} \left[\left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_{r'\in\mathcal{H}} r'^\alpha \left[\int_{\theta'\in\Theta} \phi(r'|\theta')\hat{\xi}(\theta'|r)\mathrm{d}\theta' \right] \mathrm{d}r' \right)^{\frac{1}{1-\alpha}} \right)^{1-\alpha} r^\alpha (y-c)^\alpha \right]\\
&\quad \cdot \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r, \Bigg\} \qquad (51)
\end{aligned}
$$

where $\hat{\xi}(\theta|r)$ is defined by (6).

Expression (51) has multiple integrals for the production shock. The outer integral with dummy r reflects the uncertainty faced by the planner today (i.e., in period 1) about today's production shock which is revealed tomorrow. The uncertainty emanating from today's yet-to-be-realized production shock has an effect on the stock tomorrow (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 planner takes tomorrow (i.e., in period 2 or second-to-last period) for tomorrow's production shock affecting stochastically production in after tomorrow (i.e., in period 3 or last period). To see this from another point of view, (51) can be simplified to

$$
V_2^L(y;\xi) = \max_{c\in(0,y)} \left\{ c^\alpha + \delta \int_{r\in\mathcal{H}} \left(\omega_1^L(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha (y-c)^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \tag{52}
$$

where, from (47),

$$
\omega_1^L(\hat{\xi}(\cdot|r)) = \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right)^{\frac{1}{1-\alpha}} \right)^{-1} \tag{53}
$$

is the optimal consumption rate for the one-period horizon. From (52), today's shock affects tomorrow's payoff via the stock and via tomorrow's optimal behavior in a one-period horizon.

The first-order condition corresponding to (52) is

$$c^{\alpha-1} - \delta \int_{r\in\mathcal{H}} \left(\omega_1^L(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha (y-c)^{\alpha-1} \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r = 0, \quad (54)$$

yielding the two-period-horizon policy

$$\rho_2^L(y;\xi) = \frac{y}{1+\delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} \left(\omega_1^L(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right)^{\frac{1}{1-\alpha}}}. \quad (55)$$

Note that (55), retains the linearity in y, and, thus, the two-period-horizon value function also retains the linearity in y^α.

Given that, from (48), 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 $V_{\tau-1}^L(y;\xi) = \left(\omega_{\tau-1}^L(\xi)\right)^{\alpha-1} y^\alpha$, where, unlike the deterministic and stochastic environments, $\omega_{\tau-1}^L(\xi) \in (0,1)$ is not a constant, and depends on beliefs that evolve over time. For $\tau = 2, 3, \ldots$, the τ-period-horizon value function is

$$V_\tau^L(y;\xi) = \max_{c\in(0,y)} \left\{c^\alpha + \delta \int_{r\in\mathcal{H}} V_{\tau-1}^L(r(y-c), \hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right\}, \quad (56)$$

$$= \max_{c\in(0,y)} \left\{c^\alpha + \delta \int_{r\in\mathcal{H}} \left(\omega_{\tau-1}^L(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha (y-c)^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right\}. \quad (57)$$

The first-order condition

$$c^{\alpha-1} = \delta(y-c)^{\alpha-1} \int_{r\in\mathcal{H}} \left(\omega_{\tau-1}^L(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r \quad (58)$$

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

$$\omega_\tau^L(\xi)^{\alpha-1} = \delta(1-\omega_\tau^L(\xi))^{\alpha-1} \int_{r\in\mathcal{H}} \left(\omega_{\tau-1}^L(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r. \tag{59}$$

Plugging $\rho_\tau^L(y;\xi) = \omega_\tau^L(\xi)y$ back into (57) yields

$$\begin{aligned} V_\tau^L(y;\xi) = &\Bigg(\omega_\tau^L(\xi)^\alpha + \delta(1-\omega_\tau^L(\xi))^\alpha \int_{r\in\mathcal{H}} \left(\omega_{\tau-1}^L(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \\ &\cdot \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\Bigg) y^\alpha, &(60)\\ \equiv &\left(\omega_\tau^L(\xi)\right)^{\alpha-1} y^\alpha, &(61)\end{aligned}$$

so that the optimal consumption rate for τ-period horizon is consistent with the functional form of the continuation value function and implicitly characterized by

$$\begin{aligned} \omega_\tau^L(\xi)^{\alpha-1} = &\,\omega_\tau^L(\xi)^\alpha + \delta(1-\omega_\tau^L(\xi))^\alpha \int_{r\in\mathcal{H}} \left(\omega_{\tau-1}^L(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \\ &\cdot \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r, &(62)\\ \omega_\tau^L(\xi) = &\frac{1}{1+\delta^{\frac{1}{1-\alpha}}\left(\int_{r\in\mathcal{H}} \left(\omega_{\tau-1}^L(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right)^{\frac{1}{1-\alpha}}}, &(63)\end{aligned}$$

with, from (47), initial condition

$$\omega_1^L(\xi) = \left(1+\delta^{\frac{1}{1-\alpha}}\left(\int_{r\in\mathcal{H}} r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{-1}. \tag{64}$$

Proposition 4.1 provides the optimal policy function for any finite horizon.

Proposition 4.1. *Suppose that $\int_{r\in\mathcal{H}} r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r < \infty$. In a learning environment, for $\tau = 0, 1, \ldots,$ $\rho_\tau^L(y;\xi) = \omega_\tau^L(\xi)y$, where $\omega_\tau^L(\xi)$ is* *recursively defined by*

$$\omega_\tau^L(\xi)^{\alpha-1} = \left(1+\delta^{\frac{1}{1-\alpha}}\left(\int_{r\in\mathcal{H}}\left(\omega_{\tau-1}^L(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha}, \tag{65}$$

with initial condition

$$\omega_1^L(\xi) = \left(1+\delta^{\frac{1}{1-\alpha}}\left(\int_{r\in\mathcal{H}} r^\alpha \left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{-1}. \tag{66}$$

Proof. Rearranging (63) yields (65). □

Proposition 4.2 provides the optimal policy function for an infinite horizon.

Proposition 4.2. *Suppose that* $\int_{r\in\mathcal{H}} r^\alpha \left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r \in (0,1)$*. Then, from* (65), $\lim_{\tau\to\infty}\rho_\tau^L(y;\xi) \equiv \rho_\infty^L(y;\xi) = \omega_\infty^L(\xi)y$ *exists and* $\omega_\infty^L(\xi)\in(0,1)$ *is implicitly defined by*

$$\omega_\infty^L(\xi)^{\alpha-1} = \left(1+\delta^{\frac{1}{1-\alpha}}\left(\int_{r\in\mathcal{H}}\left(\omega_\infty^L(\hat{\xi}(\cdot|r))\right)^{\alpha-1} r^\alpha \left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha}. \tag{67}$$

Proof. Let $\kappa_\tau(\xi)\equiv\omega_\tau^L(\xi)^{\alpha-1}$ so that (65) is rewritten as

$$\kappa_\tau(\xi) = \left(1+\delta^{\frac{1}{1-\alpha}}\left(\int_{r\in\mathcal{H}}\kappa_{\tau-1}(\hat{\xi}(\cdot|r)) r^\alpha \left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha}. \tag{68}$$

1. Monotonicity of $\kappa_\tau(\xi)$. From (47) and (55), we know that

$$\kappa_0(\xi) = 1 < \kappa_1(\xi) = \left(1+\delta^{\frac{1}{1-\alpha}}\left(\int_{r\in\mathcal{H}} r^\alpha \left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha}. \tag{69}$$

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

$$\kappa_{\tau+1}(\xi) = \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} \kappa_\tau(\hat{\xi}(\cdot|r)) r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha} \tag{70}$$

$$> \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} \kappa_{\tau-1}(\hat{\xi}(\cdot|r)) r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha} \tag{71}$$

$$= \kappa_\tau(\xi). \tag{72}$$

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

$$M = \left(\frac{1}{1 - \delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right)^{\frac{1}{1-\alpha}}}\right)^{1-\alpha} > 1 \tag{73}$$

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

$$\kappa_{\tau+1}(\xi) = \left(1 + \delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} \kappa_\tau(\hat{\xi}(\cdot|r)) r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha} \tag{74}$$

$$< \left(1 + \delta^{\frac{1}{1-\alpha}} M^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right)^{1-\alpha}, \tag{75}$$

$$= M. \tag{76}$$

where the last equality comes from (73).

3. Since $\lim_{\tau\to\infty} \kappa_\tau(\xi)$ exists, so does $\lim_{\tau\to\infty} \omega_\tau^L(\xi)$. Since $\kappa_\tau(\xi) \equiv \omega_\tau^L(\xi)^{1-\alpha}$, taking the limits on both sides of (68) yields (67).

□

Having studied the effect of learning on the planner's maximization problem and characterized optimal behavior in a learning model, we now discuss the effect of learning on optimal behavior. Specifically, we compare optimal behavior under stochastic and learning environments (in the context of an iso-elastic utility and linear production).

Proposition 4.3. *Suppose that beliefs about the random production shock are unbiased, i.e.,*

$$\int_{r\in\mathcal{H}} r\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r=\int_{r\in\mathcal{H}} r\phi(r|\theta^*)\mathrm{d}r. \tag{77}$$

Then, learning increases present consumption, i.e., $\rho^L_\infty(y;\xi)>\rho^S_\infty(y;\theta^*)$.

Proof. First, from (77) and the fact that $\alpha\in(0,1)$, it follows that

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

$$1-\delta^{\frac{1}{1-\alpha}}\left(\int_{r\in\mathcal{H}} r^\alpha\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right)^{\frac{1}{1-\alpha}}>1-\delta^{\frac{1}{1-\alpha}}\left(\int_{r\in\mathcal{H}} r^\alpha\phi(r|\theta^*)\mathrm{d}r\right)^{\frac{1}{1-\alpha}}. \tag{78}$$

Second, from Proposition 4.2, $\rho^L_\infty(y;\xi)=\omega^L_\infty(\xi)y$ such that from the proof of Proposition 4.2,

$$\omega^L_\infty(\xi)=\frac{1}{\kappa^L_\infty(\xi)^{\frac{1}{1-\alpha}}} \tag{79}$$

where from (74), (75), and (76), $\kappa^L_\infty(\xi)<M$, M defined by (73). Since $\kappa^L_\infty(\xi)\equiv\omega^L_\infty(\xi)^{\alpha-1}$, it follows that

$$\omega^L_\infty(\xi)>\frac{1}{M^{\frac{1}{1-\alpha}}}. \tag{80}$$

Plugging (73) into (80) yields

$$\omega^L_\infty(\xi)>1-\delta^{\frac{1}{1-\alpha}}\left(\int_{r\in\mathcal{H}} r^\alpha\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r\right)^{\frac{1}{1-\alpha}}. \tag{81}$$

Combining inequalities (78) and (81) with (43) implies that learning increases consumption, i.e.,

$$\rho_\infty^L(y;\xi) = \omega_\infty^L(\xi)y \tag{82}$$

$$> \left(1 - \delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} r^\alpha \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right) y, \tag{83}$$

$$> \left(1 - \delta^{\frac{1}{1-\alpha}} \left(\int_{r\in\mathcal{H}} r^\alpha \phi(r|\theta^*)\mathrm{d}r\right)^{\frac{1}{1-\alpha}}\right) y \tag{84}$$

$$= \rho_\infty^S(y;\theta^*). \tag{85}$$

□

5 Discussion: Learning in Mirman-Zilcha

We now compare the effect of learning on the maximization problem between our model (with iso-elastic utility and linear production) and the Mirman-Zilcha model (with log utility and Cobb-Douglas production). While the planner's policy function in the Mirman-Zilcha model is derived in KMS, there is no explanation about how learning alters the maximization problem in that context. Here, we show that the combination of a log utility and Cobb-Douglas production removes some (but not all) of the effect of learning.

To consider the Mirman-Zilcha model, suppose that $u(c) = \ln c$ and $\hat{y} = (y-c)^r$, $r \in (0,1)$. Then, the τ-period-horizon value function is

$$W_\tau^L(y;\xi) = \max_{c\in(0,y)} \left\{\ln c + \delta \int_0^1 W_{\tau-1}^L((y-c)^r; \hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] \mathrm{d}r\right\}, \tag{86}$$

with $W_0^L(y;\xi) = \ln y$. The one-period-horizon value function is

$$W_1^L(y;\xi) = \max_{c\in(0,y)} \left\{ \ln c + \delta \int_0^1 W^0((y-c)^r, \hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \tag{87}$$

$$= \max_{c\in(0,y)} \left\{ \ln c + \delta \int_0^1 r \ln(y-c) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \tag{88}$$

$$= \max_{c\in(0,y)} \left\{ \ln c + \delta \ln(y-c) \int_0^1 r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \tag{89}$$

$$= \max_{c\in(0,y)} \left\{ \ln c + \delta \ln(y-c) \int_{\theta\in\Theta} \left[\int_0^1 r\phi(r|\theta)\mathrm{d}r \right] \xi(\theta)\mathrm{d}\theta \right\}, \tag{90}$$

$$= \max_{c\in(0,y)} \left\{ \ln c + \delta \left(\int_{\theta\in\Theta} \mu(\theta)\xi(\theta)\mathrm{d}\theta \right) \ln(y-c) \right\}, \tag{91}$$

where $\mu(\theta) \equiv \int_0^1 r\phi(r|\theta)\mathrm{d}r$. The first-order condition $\frac{1}{c} - \frac{\delta \int_{\theta\in\Theta} \mu(\theta)\xi(\theta)\mathrm{d}\theta}{y-c} = 0$ yields

$$\rho_1^L(y;\xi) = \frac{y}{1 + \delta \left(\int_{\theta\in\Theta} \mu(\theta)\xi(\theta)\mathrm{d}\theta \right)}, \tag{92}$$

$$\equiv \omega_1^L(\xi) y. \tag{93}$$

Plugging (93) into (91) yields

$$W_1^L(y;\xi) = \ln \rho_1^L(y;\xi) + \delta \left(\int_{\theta\in\Theta} \mu(\theta)\xi(\theta)\mathrm{d}\theta \right) \ln(y - \rho_1^L(y;\xi)), \tag{94}$$

$$= \ln \omega_1^L(\xi) y + \delta \left(\int_{\theta\in\Theta} \mu(\theta)\xi(\theta)\mathrm{d}\theta \right) \ln(y - \omega_1^L(\xi) y), \tag{95}$$

$$\begin{aligned} &= \left(1 + \delta \int_{\theta\in\Theta} \mu(\theta)\xi(\theta)\mathrm{d}\theta \right) \ln y \\ &\quad + \ln \omega_1^L(\xi) + \delta \left(\int_{\theta\in\Theta} \mu(\theta)\xi(\theta)\mathrm{d}\theta \right) \ln(1 - \omega_1^L(\xi)). \end{aligned} \tag{96}$$

Using (96), the two-period-horizon value function is

$$
\begin{aligned}
&W^2(y;\xi)\\
&= \max_{c\in(0,y)} \left\{ \ln c + \delta \int_0^1 W^1((y-c)^r, \hat{\xi}(\cdot|r)) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \qquad (97)\\
&= \max_{c\in(0,y)} \left\{ \ln c + \delta \int_0^1 \left(1 + \delta \int_{\theta'\in\Theta} \mu(\theta')\hat{\xi}(\theta'|r)\mathrm{d}\theta' \right) \ln(y-c)^r \left[\int_{\theta'\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right.\\
&\quad + \delta \int_0^1 \left(\ln \omega_1^L(\hat{\xi}(\cdot|r)) + \delta \left(\int_{\theta'\in\Theta} \mu(\theta')\hat{\xi}(\theta'|r)\mathrm{d}\theta' \right) \ln(1-\omega_1^L(\hat{\xi}(\cdot|r))) \right)\\
&\quad \left. \cdot \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \qquad (98)\\
&= \max_{c\in(0,y)} \left\{ \ln c + \delta \left(\int_0^1 \left(1 + \delta \int_{\theta'\in\Theta} \mu(\theta')\hat{\xi}(\theta'|r)\mathrm{d}\theta' \right) r \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right) \ln(y-c) \right.\\
&\quad + \delta \int_0^1 \left(\ln \omega_1^L(\hat{\xi}(\cdot|r)) + \delta \left(\int_{\theta'\in\Theta} \mu(\theta')\hat{\xi}(\theta'|r)\mathrm{d}\theta' \right) \ln(1-\omega_1^L(\hat{\xi}(\cdot|r))) \right)\\
&\quad \left. \cdot \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \qquad (99)
\end{aligned}
$$

where $\omega_1^L(\hat{\xi}(\cdot|r))$ is the optimal consumption rate for a one-period horizon evaluated at the posterior beliefs.

Consider expression (99) more closely:

$$
\begin{aligned}
&W_2^L(y;\xi)\\
&= \max_{c\in(0,y)} \left\{ \ln c + \delta \underbrace{\left(\int_0^1 \left(1 + \delta \int_{\theta'\in\Theta} \mu(\theta')\hat{\xi}(\theta'|r)\mathrm{d}\theta' \right) r \ln(y-c) \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right)}_{=\mathcal{A}} \right.\\
&\quad + \delta \int_0^1 \underbrace{\left(\ln \omega_1^L(\hat{\xi}(\cdot|r)) + \delta \left(\int_{\theta'\in\Theta} \mu(\theta')\hat{\xi}(\theta'|r)\mathrm{d}\theta' \right) \ln(1-\omega_1(\hat{\xi}(\cdot|r))) \right)}_{=\mathcal{B}}\\
&\quad \left. \cdot \left[\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta \right] \mathrm{d}r \right\}, \qquad (100)
\end{aligned}
$$

where the planner anticipates the effect of today's production shock using the expected p.d.f. of $\tilde{r}$ given prior beliefs, i.e., $\int_{\theta\in\Theta} \phi(r|\theta)\xi(\theta)\mathrm{d}\theta$. From

(100), the effect of learning is two-fold. First, the term $\mathcal{A}$ refers to the anticipation of the planner about the stochastic effect of today's production shock r on tomorrow's stock (i.e., $r\ln(y-c)$), as well as the stochastic effect of today's production shock on tomorrow's expectations about the last period production shock, i.e., $\int_{\theta'\in\Theta}\mu(\theta')\hat{\xi}(\theta'|r)\mathrm{d}\theta'$. In the Mirman-Zilcha model, the term $\mathcal{A}$ can be further simplified to

$$\mathcal{A} = \int_0^1 \left(1 + \delta \int_{\theta'\in\Theta} \mu(\theta') \frac{\phi(r|\theta')\xi(\theta')}{\int_{\theta''\in\Theta}\phi(r|\theta'')\xi(\theta'')\mathrm{d}\theta''}\mathrm{d}\theta'\right) r\ln(y-c) \left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right]\mathrm{d}r, \tag{101}$$

$$= \int_0^1 \left(\left[\int_{\theta\in\Theta}\phi(r|\theta)\xi(\theta)\mathrm{d}\theta\right] + \delta\int_{\theta'\in\Theta}\mu(\theta')\phi(r|\theta')\xi(\theta')\mathrm{d}\theta'\right) r\ln(y-c)\mathrm{d}r, \tag{102}$$

$$= \left(\left[\int_{\theta\in\Theta}\left(\int_0^1 r\phi(r|\theta)\mathrm{d}r\right)\xi(\theta)\mathrm{d}\theta\right] + \delta\int_{\theta'\in\Theta}\mu(\theta')\left(\int_0^1\phi(r|\theta')r\mathrm{d}r\right)\xi(\theta')\mathrm{d}\theta'\right)\ln(y-c), \tag{103}$$

$$= \left(\left[\int_{\theta\in\Theta}\mu(\theta)\xi(\theta)\mathrm{d}d\theta\right] + \delta\int_{\theta'\in\Theta}\mu(\theta')^2\xi(\theta')\mathrm{d}\theta'\right)\ln(y-c), \tag{104}$$

which implies that the effect of the posterior beliefs enters through integrals of functions of of the conditional mean production shock, i.e., $\mu(\theta), \theta\in\Theta$. Note that from (52), for the case of iso-elastic utility and linear production with multiplicative uncertainty, it is impossible to obtain this type of conditional certainty equivalence.

The term B reflects the planner's (stochastic) anticipation of the effect of observing the production shock on the future optimal decision once he reaches a one-period horizon program through the optimal consumption rate $\omega_1^L(\hat{\xi}(\cdot|r))$. This is the indirect effect of anticipation of learning on beliefs through future optimal decisions, i.e., what the planner will do once he gets to a one-period horizon program, observes r, and updates beliefs. From (100), only the term $\mathcal{A}$ matters for optimization. In other words, optimal behavior in shorter horizon or subsequent periods (embedded in the term $\mathcal{B}$ in (100)) do not matter. However, from (52), in the case of iso-elastic utility and linear production, both the direct and the indirect effects remain. In particular, posterior beliefs cannot be simplified and optimal behavior in the future remains affected by learning.

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.W. Beck and V. Wieland. Learning and Control in a Changing Economic Environment. *J. Econ. Dynam. Control*, 26(9):1359–1377, 2002.

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.

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

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, 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.