Discussion Paper Series

RIEB

Kobe University

DP2013-23

Stochastic Optimal Growth with Risky Labor Supply

Yiyong CAI

Takashi KAMIHIGASHI

John STACHURSKI

August 8, 2013

Research Institute for Economics and Business Administration

Kobe University

2-1 Rokkodai, Nada, Kobe 657-8501 JAPAN

Stochastic Optimal Growth with Risky Labor Supply*

Yiyong Cai,[a,b] Takashi Kamihigashi[c] and John Stachurski[d]

[a]CSIRO Centre for Complex Systems Science,
Commonwealth Scientific and Industrial Research Organisation, Australia

[b]Centre for Applied Macroeconomic Analysis, Australian National University

[c]Research Institute for Economics and Business Administration, Kobe University

[d]Research School of Economics, Australian National University

August 8, 2013

Abstract

Production takes time, and labor supply and profit maximization decisions that relate to current production are typically made before all shocks affecting that production have been realized. In this paper we re-examine the problem of stochastic optimal growth with aggregate risk where the timing of the model conforms to this information structure. We provide a set of conditions under which the economy has a unique, nontrivial and stable stationary distribution. In addition, we verify key optimality properties in the presence of unbounded shocks and rewards, and provide the sample path laws necessary for consistent estimation and simulation.

Keywords: stochastic stability, elastic labor, optimal growth
JEL Classification: C61, C62, O41

1 Introduction

In their 1983 paper "Real Business Cycles," John Long and Charles Plosser not only coined a phrase that every economist would soon come to know, they also helped lay the foun-

*We thank Chris Edmond, Tomoo Kikuchi, Pedro Gomis Porqueras, Kevin Reffett, Cuong Le Van and Shenghao Zhu for many helpful comments. Financial support from CBE Research Grant 62888-CGH2 and ARC Discovery Grant DP120100321 is gratefully acknowledged. *Email addresses:* `yiyong.cai@csiro.au`, `tkamihig@rieb.kobe-u.ac.jp`, `john.stachurski@anu.edu.au`

dations of a new approach to modeling the business cycle. The particular version of the stochastic optimal growth model used by Long and Plosser in their study contains a notable feature: Decisions regarding labor input must be made before current production shocks are realized. Subsequent research has carried on their agenda, but mainly using a different timing, where those making labor input decisions are permitted to observe the realized value of all shocks that will affect current production before choosing labor input (see Stokey and Lucas [30] for an introductory treatment).

In empirical research, however, these production shocks are never directly observable; they are typically computed as residuals given data on output, capital, and labor. The now conventional shock-labor-output timing (the second timing discussed above) assumes that decision makers can observe these residuals before making labor input decisions. On the other hand, the labor-shock-output timing adopted by Long and Plosser is consistent with the view that decision makers have no more information than macroeconomists, and can observe (or calculate) the residuals only after observing output. Although the relative suitability of the two approaches will vary across different modeling applications, it seems hard to deny that the Long-Plossor approach has not received due attention in the literature.[1]

In this paper, our aim is to address fundamental aspects of the stochastic optimal growth model with Long and Plosser's timing, and provide the underlying results necessary for further research. We provide a detailed analysis of optimality and dynamics with general functional forms. (Long and Plosser's model specialized to the case of log utility and Cobb-Douglas production, which results in a linear law of motion for log output. This case is a useful benchmark, but is limited in the dynamics it is able to represent.) Our first significant contribution in this paper is to provide conditions under which a nontrivial stationary distribution for output exists, and for when it is unique and globally stable. These results are valuable because the dynamics of the stochastic optimal growth model with elastic labor and general functional forms are still largely unknown, both for the traditional timing *and* the timing studied here. This is remarkable, given that the dynamics in the inelastic labor case were established so many years ago [6, 23], that all realistic applications of this class of models allow labor to be endogenously supplied, and that almost all estimation and calibration techniques depend at a fundamental level on the existence of unique, nontrivial stationary solutions.[2]

[1] A number of papers working with the Long-Plossor timing can be found in the literature. For a recent example see Balbus, Reffett and Woźny [2, p. 8]. What is lacking, however, is a foundational treatment like the one given in this paper.

[2] That quantitative applications of the stochastic optimal growth model adopt endogenous labor supply is not surprising. Not only does endogenous labor supply add realism, it also permits modelers to address some of the most fundamental questions of macroeconomics. Fluctuations in employment and the comovement of output, investment and labor supply are key phenomena of the business cycle. Because the

We also establish geometric ergodicity of the output process, and laws of large numbers and central limit theorems for functions of output, investment and labor. These properties are fundamental to almost all quantitative analysis, and the sample path limit theorems are essential for simulation and estimation strategies. Finally, as an additional contribution, we provide weak conditions on the primitives of the model under which the Bellman equation holds, optimal policies exist and are unique, continuous and (in the case of savings and consumption) monotone. These conditions permit shocks, utility and the state space to be unbounded. Our approach to the dynamic programming problem is based on the use of weighted-supremum norms.

Regarding these results, it should be noted that the Long-Plosser timing used in this paper has some technical advantages vis-a-vis the standard timing, particularly when proving uniqueness and stability results for stationary equilibria. For example, the different timing leads to different state variables, and this difference between state variables means that, in the Long-Plosser timing, the next-period production shock appears outside rather than inside the optimal policy function. This makes it more feasible to assess the impact of these shocks. Nonetheless, our paper should provide a useful starting point for proving analogous results under the standard timing.

1.1 Related Literature

The stochastic optimal growth model analyzed by Brock and Mirman [6, 7] motivated many subsequent studies aimed at characterizing optimal investment. See, for example, Mirman and Zilcha [24], Razin and Yahav [26], Donaldson and Mehra [11], Brock and Majumdar [5], Stokey et al. [30], Hopenhayn and Prescott [17], Mirman [23], Stachurski [27], Zhang [31], Nishimura and Stachurski [25], and Kamihigashi [18]. In all of these papers, labor is assumed to be inelastically supplied.

The joint behavior of capital and labor in stochastic dynamic recursive economies with market distortions and externalities has been considered in Greenwood and Huffman [15], Coleman [9], and Datta et al. [10] under a set of conditions related to monotonicity of the marginal utilities. Most recently, under more general setting, Bosi and Le Van [3], and Goenka et al. [14] have studied similar problems in deterministic Ramsey models with and without borrowing constraints, respectively. In these papers, the focus is on the existence of competitive equilibria, and the problem of stability is largely untreated.

In the dynamic stochastic general equilibrium (DSGE) literature, models usually are approximated using Taylor expansions or similar techniques (e.g. Kydland and Prescott

efficiency of the labor market is a crucial determinant of the efficiency of the whole economy, labor-output dynamics have important implications for policy makers.

[20], Hansen [16], and Galí [13, Chapter 2]). With this approach, the co-movements of capital investment and labor supply around the steady states or balanced growth path can be studied. However, it is not in general true that stability of the linear approximation implies stability of the original model (see Stachurski [28]). Furthermore, the higher order properties that are eliminated may be critical to understanding actual dynamics (see Durlauf and Quah [12]).

1.2 Structure of the Paper

The rest of the paper is structured as follows. Section 2 sets up the model and studies the social planner's problem. Section 3 gives conditions under which a nontrivial stationary distribution of output exist. Section 4 presents results on stability and uniqueness, and on sample path properties such as the law of large numbers and the central limit theory. Section 5 concludes. All proofs are deferred to the appendices.

2 The Model

In this section we first define the model and solve the social planner's problem. Below we let $\mathbb{R}_+ := [0, \infty)$ and $\mathbb{R}_{++} := (0, \infty)$. For a generic function h, the symbols h'_i, h''_i and h''_{ij} refer to the first-order, second-order, and cross partial derivatives respectively, with i, j indexing the arguments.

2.1 Model and Assumptions

We begin with an elementary description of the basic model suitable for optimization by a social planner. (There are no externalities or distortions in the model, and a discussion of decentralization can be found in Long and Plosser, 1983.) Final output is denoted y_t, and is treated as a state variable. It is observed at the start of period t and can be transformed one-for-one into current physical investment k_t. Investment and labor ℓ_t are choice variables, selected at the start of time t. A shock z_{t+1} is then revealed and production takes place, yielding at the start of next period

$$y_{t+1} = z_{t+1} F(k_t, \ell_t). \tag{1}$$

The convention with subscripts is that a time t subscript indicates that the variable lies in the time t information set and not the $t-1$ information set. In particular, z_{t+1} is not previsible at t. The function F represents the common production technology, and the shock z_{t+1} is aggregate. The value y_{t+1} that we refer to as "output" is more correctly thought of as the sum of current output and capital net of depreciation.[3]

The information structure adopted in (1) differs from the conventional specification $y_t = z_t F(k_t, \ell_t)$, under which the decision on labor input is made with the knowledge of the productivity shock that affects current production. This is because we follow Long and Plosser in assuming that labor supply is "risky;" in other words, the planner does not know the productivity shock when making decisions on labor input.

Our assumptions on the shock process are very standard:

Assumption 2.1. The shocks $\{z_t\}$ follow the Markov process on $\mathbb{R}_+$ given by

$$z_{t+1} = \psi(z_t, e_{t+1}), \quad \{e_t\} \overset{\text{IID}}{\sim} \mu, \quad t = 0, 1, \ldots \tag{2}$$

The IID sequence $\{e_t\}$ is defined on a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ and takes values in a measurable space $(E, \mathscr{E})$ with common distribution μ. The function $\psi \colon \mathbb{R}_+ \times E \to \mathbb{R}_+$ is jointly measurable, and $z \mapsto \psi(z, e)$ is continuous and increasing on $\mathbb{R}_+$ for each $e \in E$.

We take Π to be the associated stochastic kernel (i.e., transition probability function), so that, in particular, $\Pi(z, B) = \mathbb{P}\{\psi(z, e_{t+1}) \in B\}$ for all $z \in \mathbb{R}_+$ and Borel subsets B of $\mathbb{R}_+$.[4]

Assumption 2.2. The production function F is homogeneous of degree one, increasing and concave on $\mathbb{R}^2_+$. On $\mathbb{R}^2_{++}$, it is twice differentiable, strictly positive, and satisfies

1. $F'_i(k, \ell) > 0$, $F''_{ij}(k, \ell) \geq 0$, and $F''_i(k, \ell) < 0$ for $i, j = 1, 2$.
2. $\lim_{k \to 0} F'_1(k, \ell) = \lim_{\ell \to 0} F'_2(k, \ell) = \infty$.

These assumptions are standard. For example, they are satisfied when

$$F(k, \ell) = A(\alpha k^\gamma + \beta \ell^\gamma)^{1/\gamma}, \quad A, \alpha, \beta > 0, \ \gamma \in (0, 1) \tag{3}$$

or

$$F(k, \ell) = A k^\alpha \ell^{1-\alpha}, \quad A > 0, \ \alpha \in (0, 1). \tag{4}$$

[3] To be more explicit, we could take $F(k, \ell) := F_c(k, \ell) + (1 - \delta)k$, where $F_c(k, \ell)$ is current output and δ parametrizes depreciation. It is easy to verify that if F_c satisfies the conditions of our assumptions below then so does F.

[4] Some authors prefer to take the stochastic kernel Π as the primitive. The two approaches are equivalent, in the sense that every stochastic kernel on a completely metrizable topological space can be represented in the form of (2) given suitable choice of ψ, Ω, $\mathscr{F}$ and $\mathbb{P}$. See, for example, Bhattacharya and Majumdar [4, proposition C1.1].

Assumption 2.3. The period utility function $U\colon \mathbb{R}_+^2 \to \mathbb{R}_+$ is additively separable, strictly increasing and concave. In particular, we assume that $U(x,y) := u(x) + v(y)$ with the standard restrictions $h(0) = 0$, $h' > 0$, $h'' < 0$ and $\lim_{x\to 0} h'(x) = \infty$ for $h \in \{u, v\}$.

Some comments on our assumptions are in order. First, the interaction between shock and output in (1) is somewhat specialized, and at times it might be desirable to consider more general formulations such as $y_{t+1} = F(k_t, \ell_t, z_{t+1})$. In fact the fundamental optimality results presented below (e.g., finiteness of the value function, validity of the Bellman equation and existence of optimal policies) go through in this more general setting under mild assumptions. The proofs can be obtained by modifying theorem 2.1 in Kamihigashi [18].

2.2 Optimality

The social planner maximizes the expected discounted sum $\mathbb{E}\sum_{t=0}^{\infty}\beta^t U(c_t, 1-\ell_t)$, where $\beta \in (0,1)$ is the discount factor. Let $\Gamma(y) := [0,y] \times [0,1]$. A Borel measurable function

$$\sigma\colon \mathbb{R}_+^2 \ni (y,z) \mapsto \sigma(y,z) := (k(y,z), \ell(y,z)) \in \mathbb{R}_+ \times [0,1]$$

is called a *feasible policy* if $\sigma(y,z) \in \Gamma(y)$ for all $(y,z) \in \mathbb{R}_+^2$. Let the set of all feasible policies be denoted by Σ. The value $V_\sigma(y,z)$ of a feasible policy $\sigma(y,z) = (k(y,z), \ell(y,z))$ is defined as the expected discounted value of following σ. That is,

$$V_\sigma(y,z) := \mathbb{E}\left\{\sum_{t=0}^{\infty}\beta^t U(y_t - k(y_t,z_t), 1 - \ell(y_t,z_t))\right\} \qquad ((y,z) \in \mathbb{R}_+^2) \tag{5}$$

where $y_{t+1} = z_{t+1}F[k(y_t,z_t), \ell(y_t,z_t)]$ and $(y_0,z_0) = (y,z)$. To ensure that the optimization problem is well defined, we require a restriction over the set of primitives that ensures the value $V_\sigma(y,z)$ is finite and bounded over $\sigma \in \Sigma$. To impose such a restriction, we bound the value of the process

$$\hat{y}_{t+1} = z_{t+1}F(\hat{y}_t, 1), \qquad (\hat{y}_0, z_0) = (y,z) \in \mathbb{R}_+^2, \tag{6}$$

the output process generated when all output is invested and no leisure is consumed.

Assumption 2.4. $\exists\, \delta < 1$ with $m(y,z) := \mathbb{E}\sum_{t=0}^{\infty}\delta^t u(\hat{y}_t) < \infty$ for all $(y,z) \in \mathbb{R}_+^2$.[5]

[5] If, in addition to our other assumptions, $u'(c) \to 0$ as $c \to \infty$, then there exists a $K < \infty$ such that $u(y) \leq y + K$ for all y, and finiteness of $\sum_t \delta^t \mathbb{E}\,\hat{y}_t$ is sufficient for assumption 2.4.

The value function V is defined by $V(y,z) = \sup_{\sigma \in \Sigma} V_\sigma(y,z)$ for all (y,z). A feasible policy σ is called *optimal* if $V_\sigma = V$. If $\sigma(y,z) = (k(y,z), \ell(y,z))$ is optimal, then the corresponding optimal consumption is defined as $c(y,z) := y - k(y,z)$. The *optimal reproduction function* is defined as $g = F \circ \sigma$. That is,

$$g(y,z) := F[k(y,z), \ell(y,z)] := F(\sigma(y,z)).$$

Let w be the function defined by $w(y,z) := m(y,z) + v(1) + 1$, where m is the function defined in assumption 2.4 and v is the utility of leisure (see assumption 2.3). Letting $\| \cdot \|_\infty$ be the usual supremum norm, a real-valued function h on $\mathbb{R}^2_+$ is called w-bounded if the weighted supremum norm $\|h\|_w := \|h/w\|_\infty$ is finite. Let $\mathscr{W}$ be the set of continuous w-bounded functions on $\mathbb{R}^2_+$.

Theorem 2.1. *If assumptions 2.1–2.4 hold, then the following statements are true.*

1. *The value function is finite, continuous and w-bounded. It is the unique function in $\mathscr{W}$ satisfying the Bellman equation*

$$V(y,z) = \max_{k,\ell \in \Gamma(y)} \left\{ U(y-k, 1-\ell) + \beta \int V[z' F(k,\ell), z'] \Pi(z, dz') \right\}. \tag{7}$$

2. *A unique optimal policy $\sigma(y,z) = (k(y,z), \ell(y,z))$ exists. For each $(y,z) \in \mathbb{R}^2_+$, it satisfies*

$$\sigma(y,z) = \operatorname*{arg\,max}_{(k,\ell) \in \Gamma(y)} \left\{ U(y-k, 1-\ell) + \beta \int V[z' F(k,\ell), z'] \Pi(z, dz') \right\}. \tag{8}$$

3. *V is strictly concave, strictly increasing and differentiable in its first argument. When $y, z > 0$ we have $V_1'(y,z) = u' \circ c(y,z) := u'(c(y,z))$.*

4. *For each $z > 0$, the optimal policy $(k(\cdot, z), \ell(\cdot, z))$ is continuous and interior, and $k(\cdot, z)$, $c(\cdot, z)$ and $g(\cdot, z)$ are all monotone increasing.*

5. *If $y, z > 0$, then the optimal policies satisfy*

$$u' \circ c(y,z) = \beta \int u' \circ c[z' g(y,z), z'] z' F_1'(\sigma(y,z)) \Pi(z, dz') \tag{9}$$

$$v'(1 - \ell(y,z)) = \beta \int u' \circ c[z' g(y,z), z'] z' F_2'(\sigma(y,z)) \Pi(z, dz'). \tag{10}$$

Proof. See appendix A. □

3 Existence of Stationary Distributions

In this section we study the existence of nontrivial stationary distributions for the optimal output process, where nontrivial means that probability mass is not concentrated on zero. In all of what follows, the term *distribution* will be synonymous with "Borel probability measure." When studying existence and stability, we will restrict attention to the case of IID shocks:

Assumption 3.1. The shocks $\{z_t\}$ are IID with common distribution μ. The distribution μ satisfies $\mu(\mathbb{R}_{++}) = 1$, $\int z\mu(dz) < \infty$ and $\int (1/z)\mu(dz) < \infty$.

Of these assumptions, the IID assumption is certainly the most restrictive. However, very little is known about the stability properties of stochastic optimal growth models with either correlated shocks *or* elastic labor. Even the question of existence of a nontrivial stationary distribution is problematic (particularly since we permit shocks to be unbounded). Here we focus on our primary interest, which is permitting elastic labor.

The assumptions $\int z\mu(dz) < \infty$ and $\int (1/z)\nu(dz) < \infty$ ensure that the right and left hand tails of μ are not excessively large. They are satisfied if, for example, μ is the lognormal density. While the condition $\int z\mu(dz) < \infty$ is difficult to improve on, the restriction $\int (1/z)\mu(dz) < \infty$ could potentially be weakened along the lines of assumption 3.4 in Kamihigashi [18].

When assumption 3.1 holds and the shocks are IID, the stochastic kernel $\Pi(z, dz')$ is equal to $\mu(dz')$. It is then immediate from the Bellman equation (7) that the value function $V(y, z)$ is constant in z. Likewise, it is immediate from (8) that the optimal policies are also constant in z, and hence the optimal reproduction function g is constant in z. In all of what follows, when treating the IID case, we will simplify notation by omitting this constant second argument. Hence $V(y, z)$ is written as $V(y)$ and treated as a function on $\mathbb{R}_+$, and likewise for σ, k, ℓ, c and g.

As a final comment, we also note that our stability results are all obtained in a setting where shocks are assumed to be independent rather than correlated.

Once an initial level of output y_0 is given, the optimal investment-labor policy determines the optimal output process $\{y_t\}$ via the stochastic recursion

$$y_{t+1} = z_{t+1}F[k(y_t), \ell(y_t)] = z_{t+1}g(y_t) \qquad t = 0, 1, \ldots \tag{11}$$

The dynamics from $y_0 = 0$ are degenerate and trivial. On the other hand, since policies are interior, if y_t is strictly positive, then both $k(y_t)$ and $\ell(y_t)$ are strictly positive. Moreover, $F(k, \ell)$ is assumed to be strictly positive whenever $k, \ell > 0$, and the distribution μ is concentrated on $\mathbb{R}_{++}$. It follows that $y_{t+1} \in \mathbb{R}_{++}$ with probability one. As a result, in all of our stability analysis we shall restrict the state space to $\mathbb{R}_{++}$. One benefit of this approach is that any stationary distribution we obtain is automatically nontrivial.

A distribution π on $\mathbb{R}_{++}$ is called *stationary* for the process (11) if y_{t+1} has distribution π whenever y_t has distribution π. More formally, π is stationary for (11) if

$$\int h(y)\, \pi(dy) = \int \left[\int h[zg(y)] \mu(dz) \right] \pi(dy) \tag{12}$$

for all bounded Borel measurable $h \colon \mathbb{R}_{++} \to \mathbb{R}$. Note that if π is stationary and y_t has distribution π at some time t, then the process from t on is strict sense stationary, with common marginal distribution π. In the current setting, existence of a stationary distribution is a nontrivial problem. The main part of the proof involves showing that optimal output does not diverge to infinity or collapse to zero. Establishing these properties requires additional assumptions on β, μ and F. We will require that one of the following two assumptions holds.

Assumption 3.2. Together, β, F and μ satisfy

$$\int \frac{1}{z} \mu(dz) < \beta \lim_{k \to 0} F_1'(k, 0) \quad \text{and} \quad \lim_{k \to \infty} \frac{F(k,1)}{k} \int z \mu(dz) < 1. \tag{13}$$

For example, the CES production function in (3) satisfies the restrictions in (13) whenever $\mathbb{E}\,(1/z_t) < (\beta A \alpha^{1/\gamma})/\gamma$ and $A\alpha^{1/\gamma} \mathbb{E}\, z_t < 1$. The first inequality in (13) implies sufficient labor and investment to prevent y_t converging to zero. It states that the probability of bad shocks is small relative to the patience of the agent and marginal productivity of capital near zero. The second bound prevents y_t from diverging to infinity.

Assumption 3.3. The function F satisfies

$$F(k, 0) = F(0, \ell) = \lim_{k \to \infty} F_2'(k, \ell) = \lim_{\ell \to \infty} F_2'(k, \ell) = 0 \text{ for all } (k, \ell) \in \mathbb{R}_+^2. \tag{14}$$

For example, the Cobb-Douglas production function in (4) satisfies (14).

Theorem 3.1. *If assumptions 2.2–2.4 and 3.1 all hold, and at least one of assumptions 3.2 and 3.3 holds, then then the optimal output process has at least one nontrivial stationary distribution.*

Proof. See appendix B. □

4 Stability and Ergodicity

Existence of a stationary distribution is necessary for the model to have a stationary equilibrium that can be tested against data. However, such an equilibrium cannot be considered as a prediction of the model unless it is stable, and preferably unique. Here uniqueness means that no more than one distribution on $\mathbb{R}_{++}$ satisfies (12), and stability means that there exists a unique stationary distribution π, and, moreover, the distribution of y_t converges to π weakly as $t \to \infty$, irrespective of the initial condition $y_0 \in \mathbb{R}_{++}$.[6]

The conditions of theorem 3.1 are not sufficient for uniqueness and stability of the stationary distribution. In this section we look at providing weak conditions for uniqueness and stability. Uniqueness and stability cannot be obtained without sufficient mixing in the stochastic process $\{y_t\}$.[7] One approach to mixing is via irreducibility, but this approach requires relatively strict conditions on the distribution of the shock that are not in fact necessary for uniqueness and stability (see section 4 for more discussion). Instead we impose the following assumption on μ, which states that arbitrarily low productivity shocks are possible (although the probability can be extremely small).

Assumption 4.1. If $a \in \mathbb{R}_{++}$, then $\mathbb{P}\{z_t \leq a\} > 0$.

From an econometric and computational perspective, another desirable property of the optimal output process is that the law of large numbers is valid, in the sense that time series averages converge to the corresponding expectations under the stationary distribution: For $h \colon \mathbb{R}_{++} \to \mathbb{R}$,

$$\frac{1}{n} \sum_{t=1}^{n} h(y_t) \stackrel{\text{a.s.}}{\to} \int h \, d\pi \qquad (n \to \infty), \tag{15}$$

where $\stackrel{\text{a.s.}}{\to}$ means almost sure convergence.[8] Many results on laws of large numbers require that the process $\{y_t\}$ in (15) is either weak or strict sense stationary, which in the Markov setting translates to the requirement that the first element y_1 is drawn from a stationary distribution. Such a result is much less useful, since it cannot in general be used for simulation-based computations (because we have no way of drawing from the unknown stationary distribution). The following theorem provides uniqueness, stability and a law of large numbers where the initial condition can be arbitrary.

[6]Recall that a sequence of distributions $\{\pi_n\}$ on $\mathbb{R}_{++}$ converges weakly to π if $\int h d\pi_n$ converges to $\int h d\pi$ for every continuous bounded $h \colon \mathbb{R}_{++} \to \mathbb{R}$.

[7]For intuition, see, for example, Stokey *et al.* [30, p. 380].

[8]As usual, we say that a sequence of random variables $\{X_n\}$ converges to a random variable X *almost surely* if the scalar convergence $X_n(\omega) \to X(\omega)$ holds on the complement of a $\mathbb{P}$-null set.

Theorem 4.1. *Assume the conditions of theorem 3.1. Let $\{y_t\}$ be the optimal output process, with arbitrary initial condition $y_0 \in \mathbb{R}_{++}$. If assumption 4.1 also holds, then*

1. *the stationary distribution π is unique,*
2. *the distribution of y_t converges weakly to π as $t \to \infty$, and*
3. *the convergence in* (15) *is valid whenever h is bounded and either continuous or monotone.*

Proof. See appendix C. □

While theorem 4.1 is a useful result, there are several ways in which it can be strengthened. First, the stability result in theorem 4.1 is defined in terms of weak convergence, which is a weaker and less quantitative measure of divergence than, say, total variation convergence (see below). Second, the theorem provides no information on the rate of convergence. Third, the law of large numbers holds only for a limited class of functions. Fourth, a central limit theorem is lacking. The last two properties are particularly important for estimation and simulation-based analysis. In order to address these limitations, we now consider stronger assumptions:

Assumption 4.2. The distribution μ is absolutely continuous with respect to Lebesgue measure, and its density is strictly positive and continuous on $\mathbb{R}_{++}$. In addition, the moment $\mathbb{E}\left(z^p\right)$ is finite for some $p \geq 2$, and, moreover,

$$\lim_{k \to \infty} \left[\frac{F(k,1)}{k} \right]^p \int z^p \mu(dz) < 1. \tag{16}$$

For example, the lognormal distribution satisfies assumption 4.2 for all $p \in \mathbb{N}$.

Recall that the *total variation distance* between two distributions ϕ and ϕ' is defined as $\|\phi - \phi'\|_{TV} := \sup_B |\phi(B) - \phi'(B)|$, where the supremum is over all Borel sets. The optimal output process $\{y_t\}$ is said to be *geometrically ergodic* if, in addition to the existence of a unique stationary distribution π, there exists an $\alpha \in [0,1)$ such that, for any initial condition y_0, the total variation distance between the distribution of y_t and π is $O(\alpha^t)$.[9] We can now present the following result, which gives both a stronger convergence result than theorem 4.1 and a rate of convergence. The proof is in appendix D.

Theorem 4.2. *Assume the conditions of theorem 3.1. If assumption 4.2 also holds, then $\{y_t\}$ is geometrically ergodic.*

[9] In other words, if π_t is the distribution of y_t when $y_0 = y$, then $\|\pi_t - \pi\|_{TV} \leq \alpha^t M(y)$ for some function M that is everywhere finite on $\mathbb{R}_{++}$.

An initial sample path result was presented in theorem 4.1. With the current assumptions we can obtain stronger results. To simplify the presentation in what follows, we introduce some additional notation. If $h\colon \mathbb{R}_{++} \to \mathbb{R}$ is such that $\int h d\pi$ exists, then we define the asymptotic variance

$$\gamma_h^2 := \lim_{n\to\infty} \frac{1}{n} \mathbb{E}_{\pi} \left\{ \sum_{t=1}^{n} \left[h(y_t) - \int h d\pi \right] \right\}^2,$$

where $\mathbb{E}_{\pi}$ denotes the conditional expectation given $y_0 \sim \pi$.

Theorem 4.3. *Let $\{y_t\}$ be the optimal output process starting from arbitrary $y_0 \in \mathbb{R}_{++}$, and let h be any measurable real-valued function defined on $\mathbb{R}_{++}$. If the conditions of theorem 4.2 hold, then γ_h^2 is finite and the following implications are true:*

$$|h(y)| \leq y^p \quad \Longrightarrow \quad \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} h(y_t) = \int h d\pi \quad a.s. \tag{17}$$

$$h(y)^2 \leq y^p \text{ and } \gamma_h^2 = 0 \quad \Longrightarrow \quad \lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{t=1}^{n} h(y_t) = 0 \quad a.s. \tag{18}$$

$$h(y)^2 \leq y^p \text{ and } \gamma_h^2 > 0 \quad \Longrightarrow \quad \frac{1}{\sqrt{n}} \sum_{t=1}^{n} \left[h(y_t) - \int h d\pi \right] \xrightarrow{d} N(0, \gamma_h^2). \tag{19}$$

Proof. See appendix D. □

An immediate consequence of theorem 4.3 is that all moments of the optimal output process up to order p satisfy the law of large numbers, while all moments up to order q satisfy the central limit theorem, where q is the largest integer such that $2q \leq p$. Moreover, since $0 < k(y)^p < y^p$, $0 < c(y)^p < y^p$ and $0 < \ell(y)^p < 1$, the investment, consumption and labor processes also satisfy the limit theorems.

5 Concluding Remarks

In this paper we addressed foundational aspects of the stochastic optimal growth model with Long and Plosser's (1983) timing. We provided conditions under which a nontrivial stationary distribution for output exists, and for when it is unique and globally stable. We showed how geometric ergodicity of the output process, laws of large numbers and central limit theorems can be obtained for functions of output, investment and labor. These amount to the first set of detailed stability results for a stochastic optimal growth model with elastic labor supply. In addition, we used a weighted supremum norm approach to provide weak conditions on the primitives of the model under which the Bellman equation holds, optimal policies exist and are unique, continuous and, in some cases, monotone. These conditions allow shocks, utility and the state space to be unbounded.

Appendix

Before commencing proofs we collect all notation that will be used in the remainder of the paper. First, for given function h, the subscript $_+$ and superscript $^-$ denote the lower right and upper left partial derivatives, respectively. For any partially ordered topological space T, the symbol $\mathscr{B}(T)$ always denotes the Borel σ-algebra of subsets of T, while $\mathscr{P}(T)$ is the set of distributions (i.e., Borel probability measures) on T. In addition, we let $C(T)$ be the set of all bounded continuous functions from T to $\mathbb{R}$, $M(T)$ be the set of all bounded, measurable functions from T to $\mathbb{R}$, and $iM(T)$ be increasing functions in $M(T)$.

A *stochastic kernel* P on T is function from T to $\mathscr{P}(T)$ such that $x \mapsto P(x, B)$ is Borel measurable for each $B \in \mathscr{B}(T)$.[10] The value $P(x, B)$ represents the probability of output transitioning from current state x into set B in one unit of time. Given a stochastic kernel P, we define two linear operators. The first is the so-called *left Markov operator*, which maps $\mathscr{P}(T)$ into itself via

$$(\phi P)(B) = \int P(x, B)\phi(dx) \qquad (\phi \in \mathscr{P}(T),\ B \in \mathscr{B}(T)).$$

The second is the so-called *right Markov operator*, which maps $M(T)$ into itself via

$$(Ph)(x) = \int h(y)P(x, dy) \qquad (h \in M(T),\ x \in T).$$

For the right Markov operator, every constant function on T is a fixed point. (We will also use the notation $(Ph)(x) = \int h(y)P(x, dy)$ for unbounded nonnegative h, but with the understanding that the integral in this expression may not be finite-valued.) The interpretation of the t-th iterate $(P^t h)(x)$ is the expectation of $h(X_t)$ given that $X_0 = x$ and the process $\{X_t\}$ updates by P. The interpretation of $(\phi P^t)(B)$ is the probability that $X_t \in B$ given that X_0 has distribution ϕ. Following standard usage we typically write $P^t(x, B)$ for $(\delta_x P^t)(B)$, where δ_x is the probability measure concentrated on x. This expression gives the probability of the state moving from x now into set B in t steps.

P is said to have the *Feller* property if $Ph \in C(T)$ whenever $h \in C(T)$, and *increasing* if $Ph \in iM(T)$ whenever $h \in iM(T)$. A distribution $\pi \in \mathscr{P}(T)$ is called *stationary* for P iff $\pi = \pi P$. A stochastic kernel P is called *globally stable* if there exists one and only one

[10] See Meyn and Tweedie [22] for more discussion on stochastic kernels and Markov processes.

stationary distribution π in $\mathscr{B}(T)$, and, moreover, the sequence $\{\phi P^t\}_{t\geq 0}$ converges to π in distribution as $t \to \infty$ for any $\phi \in \mathscr{P}(T)$. Letting $\leq$ be the partial order on T, we say that P is *downward reaching* [19] if, for each $a \in T$ and $x \in T$, there exists a $t \in \mathbb{N}$ with $P^t(x, \{y \in T : y \leq a\}) > 0$. In other words, given any initial state x and any other point a in the state space, the process falls below a at some point in time with positive probability.

A sequence of distributions $\{\phi_n\}$ in $\mathscr{P}(T)$ is called *tight* if, for all $\varepsilon > 0$, there exists a compact set $K \subset T$ such that $\sup_{n\in\mathbb{N}} \phi_n(T \setminus K) \leq \varepsilon$. A stochastic kernel P is called *bounded in probability* if $\{\delta_x P^t\}_{t\geq 0}$ is tight for every $x \in T$. A nonnegative measurable function W on T is called *coercive* if there exists an increasing sequence of compact sets $\{C_n\}$ such that $T = \cup_n C_n$ and $\lim_{n\to\infty} \inf_{y\notin C_n} W(y) = \infty$. When $T = \mathbb{R}_{++}$, coerciveness of W is equivalent to

$$\lim_{y\to 0} W(y) = \lim_{y\to\infty} W(y) = \infty. \tag{20}$$

Let $\phi \in \mathscr{P}$. A stochastic kernel P on T is called *ϕ-irreducible* if, for all $x \in T$ and $B \in \mathscr{B}(T)$ with $\phi(B) > 0$, there exists a t with $P^t(x, B) > 0$. A set $C \in \mathscr{B}(T)$ is called a *$\mathscr{C}$-set* for P if there exists a non-trivial measure ϕ on $\mathscr{B}(T)$ such that

$$x \in C \Longrightarrow \{P(x, B) \geq \phi(B),\ \forall B \in \mathscr{B}(T)\}. \tag{21}$$

P is called *strongly aperiodic* if (21) holds for some $C \in \mathscr{B}(T)$ with positive ϕ-measure. Letting $\{X_t\}$ be a Markov chain with stochastic kernel P, a set A is called *Harris recurrent* if $\mathbb{P}\{X_t \in A \text{ infinitely often}\} = 1$ whenever $X_0 \in A$. P is called *Harris recurrent* if it is ϕ-irreducible for some $\phi \in \mathscr{P}$, and every set with positive ϕ-measure is Harris recurrent. P is called *positive Harris recurrent* if it is Harris recurrent and has a stationary distribution.

Appendix A: Proof of Theorem 2.1

Throughout this section, we maintain assumptions 2.1–2.4. The weight function w, the function space $\mathscr{W}$ and the norm $\|\cdot\|_w$ are as defined in section 2.2. The maximal output process $\{\hat{y}_t\}$ is as given in (6). We let R be the corresponding Markov operator on $\mathbb{R}^2_+$. That is, given integrable $h\colon \mathbb{R}^2_+ \to \mathbb{R}$,

$$(Rh)(y, z) = \int h[z' F(y, 1), z'] \Pi(z, dz') \qquad (y, z) \in \mathbb{R}^2_+.$$

Lemma A-1. *The weight function w is finite, continuous and increasing on $\mathbb{R}^2_+$.*

Proof. Recall that $w(y, z) = m(y, z) + v(1) + 1$, and $m(y, z)$ is finite by assumption 2.4. Hence it remains only to check that m is continuous and increasing. Abusing notation slightly we let $u(y,z) := u(y)$, in which case m can be expressed using the Markov operator R as

$$m(y,z) = \sum_{t=0}^{\infty} \delta^t (R^t u)(y,z) \qquad (y,z) \in \mathbb{R}^2_+.$$

It follows from assumptions 2.1 and 2.2 that R is increasing, and, since u is itself increasing, $R^t u$ is increasing for all t. Hence m is increasing.

Regarding continuity, it also follows from assumptions 2.1 and 2.2 that R is Feller, and, since u is continuous, $R^t u$ is continuous for all t. Hence the partial sum $\sum_{t=0}^{k} \delta^t (R^t u)$ is continuous for all k. Continuity of the infinite sum m can be justified via the dominated convergence theorem in a straightforward way.[11] □

Proposition A-1. *The value function is finite, continuous and w-bounded. It is the unique function in $\mathscr{W}$ satisfying the Bellman equation* (7).

Proof. To verify the statement in proposition A-1, it suffices to check the conditions of theorem 12.2.22 of Stachurski (2009). The only nontrivial condition in our setting is to show that for the weight function $w = m + v(1) + 1$ we have

1. $\sup_{k,\ell \in \Gamma(y)} U(y-k, 1-\ell) \leq w(y,z)$ for all $(y,z) \in \mathbb{R}^2_+$.
2. There exists a $\rho < 1/\beta$ such that

$$\sup_{k,\ell \in \Gamma(y)} \int w[z' F(k,\ell), z'] \Pi(z, dz') \leq \rho w(y,z) \quad \text{for all} \quad (y,z) \in \mathbb{R}^2_+.$$

3. The function $(y,z,k,\ell) \mapsto \int w[z' F(k,\ell), z'] \Pi(z, dz')$ is continuous on the set of feasible state-action pairs.

Part 1 is trivial, because $\sup_{k,\ell \in \Gamma(y)} U(y-k, 1-\ell) = u(y) + v(1) \leq m(y,z) + v(1) \leq w(y,z)$. Regarding part 2, observe that

$$\sup_{k,\ell \in \Gamma(y)} \int w[z' F(k,\ell), z'] \Pi(z, dz') = \int w[z' F(y,1), z'] \Pi(z, dz') = Rw(y,z).$$

and that

$$Rm = R \sum_{t=0}^{\infty} \delta^t R^t u = \sum_{t=0}^{\infty} \delta^t R^{t+1} u = \frac{1}{\delta} \sum_{t=0}^{\infty} \delta^{t+1} R^{t+1} u \leq \frac{1}{\delta} m.$$

[11] Take $(y_n, z_n) \to (y,z) \in \mathbb{R}^2_+$. Let $\bar{y} := \sup_n y_n$ and $\bar{z} := \sup_n z_n$. Let $a_n(t) := \delta^t (R^t u)(y_n, z_n)$, $a(t) := \delta^t (R^t u)(y,z)$ and $\bar{a}(t) := \delta^t (R^t u)(\bar{y}, \bar{z})$. By continuity of $R^t u$ we have $a_n(t) \to a(t)$ for all t. Moreover, $a_n(t) \leq \bar{a}(t)$ for all t by monotonicity of $R^t u$, and $\sum_t \bar{a}_t = m(\bar{y}, \bar{z}) < \infty$. Hence the dominated convergence theorem applies, and $\sum_t a_n(t) \to \sum_t a(t)$ as $n \to \infty$. That is, $m(y_n, z_n) \to m(y,z)$, and m is continuous.

As a consequence, we have

$$Rw = R(m + v(1) + 1) = Rm + v(1) + 1 \leq \frac{1}{\delta} m + v(1) + 1 \leq \frac{1}{\delta}(m + v(1) + 1) = \frac{1}{\delta} w.$$

Combining these inequalities shows that part 2 is valid with $\rho := 1/\delta$.[12]

Regarding part 3, take arbitrary sequence of feasible state-action pairs (y_n, z_n, k_n, ℓ_n) converging to (y, z, k, ℓ) as $n \to \infty$. By continuity of w (lemma A-1) and the primitives F and ψ, we have

$$w[\psi(z_n, e)F(k_n, \ell_n), \psi(z_n, e)] \to w[\psi(z, e)F(k, \ell), \psi(z, e)] \qquad (n \to \infty).$$

To extend this to convergence of the integrals (and hence verify part 3), we need only show that $w[\psi(z_n, e)F(k_n, \ell_n), \psi(z_n, e)]$ is dominated by an μ-integrable function of e. In view of the fact that w, F and $z \mapsto \psi(z, e)$ are all increasing, a suitable dominating function is given by $D(e) := w[\psi(\bar{z}, e)F(\bar{y}, 1), \psi(\bar{z}, e)]$ where $\bar{y} := \sup_n y_n$ and $\bar{z} := \sup_n z_n$. The function D is integrable because $\int D(e)\mu(de) = (Rw)(\bar{y}, \bar{z})$. As proved in part 2, the last term is dominated by $(1/\delta)w(\bar{y}, \bar{z})$, which is finite. □

Proposition A-2. *V is strictly increasing and strictly concave in its first argument. For each z, the optimal investment and labor policies $k(\cdot, z)$ and $\ell(\cdot, z))$ are single-valued and continuous.*

Proof. The arguments are standard. A detailed proof can be found in the working paper version [8]. □

Corollary A-1. *The optimal consumption function $c(\cdot, z)$ and reproduction function $g(\cdot, z)$ are both continuous for each $z \in \mathbb{R}_+$.*

Proposition A-3. *If $c(y, z) > 0$, then $V_1'(y, z)$ exists and $V_1'(y, z) = u' \circ c(y, z)$.*

Proof. A proof based on Mirman and Zilcha [24, lemma 1] is in the working paper version [8]. □

Proposition A-4. *The optimal investment policy $k(\cdot, z)$ is increasing for each $z \in \mathbb{R}_+$.*

Proof. Fix $z \in \mathbb{R}_+$. Choose $y_1 < y_2$, and let (k_1, ℓ_1) and (k_2, ℓ_2) be the corresponding optimal choices. Suppose for a contradiction that $k_1 > k_2$. Define $c_1 := y_1 - k_1$, $c_2 := y_2 - k_2$, and $\widehat{c} := k_1 - k_2 > 0$. It follows that

$$c_2 - \widehat{c} = y_2 - k_1 > y_1 - k_1 = c_1 \geq 0.$$

[12] Recall from assumption 2.4 that $\beta < \delta$. Hence $\rho = 1/\delta < 1/\beta$ as required.

Note that $c_1 + \widehat{c} + k_2 = y_1$, and hence $(c_1 + \widehat{c}, \ell_2)$ is a feasible choice from y_1. By optimality, we have

$$U(c_1, 1 - \ell_1) + \beta \int V[z' F(k_1, \ell_1), z'] \Pi(z, dz') \\ \geq U(c_1 + \widehat{c}, 1 - \ell_2) + \beta \int V[z' F(k_2, \ell_2), z'] \Pi(z, dz'). \quad \text{(A-1)}$$

In addition, $c_2 - \widehat{c} + k_1 = y_2$, and hence $(c_2 - \widehat{c}, \ell_1)$ is feasible at y_2. It follows that

$$U(c_2 - \widehat{c}, 1 - \ell_1) + \beta \int V[z' F(k_1, \ell_1), z'] \Pi(z, dz') \\ \leq U(c_2, 1 - \ell_2) + \beta \int V[z' F(k_2, \ell_2), z'] \Pi(z, dz'). \quad \text{(A-2)}$$

Subtracting (A-2) from (A-1) yields

$$U(c_2, 1 - \ell_2) - U(c_1 + \widehat{c}, 1 - \ell_2) \geq U(c_2 - \widehat{c}, 1 - \ell_1) - U(c_1, 1 - \ell_1)$$

By separability of U, this is equivalent to

$$u(c_2) - u(c_1 + \widehat{c}) \geq u(c_2 - \widehat{c}) - u(c_1).$$

Because $c_1 + \widehat{c} > c_1$, we obtain a contradiction to the strict concavity of u. □

Proposition A-5. *If $y, z > 0$, then the optimal action $(k(y, z), \ell(y, z))$ is interior.*

Proof. The proof is long but relatively standard. All details can be found in the working paper version [8]. □

Proposition A-6. *If $y, z > 0$, then the optimal policy satisfies* (9) *and* (10).

Proof. Fix $y, z > 0$. By proposition A-5, both $k(y, z)$ and $\ell(y, z)$ are interior. Then the result follows from proposition A-3 and the first order conditions of optimality. □

Proposition A-7. *The optimal reproduction function $g(y, z)$ is increasing in y for each $z \in \mathbb{R}_{++}$.*

Proof. Fix $z \in \mathbb{R}_{++}$. Choose $0 < y_1 < y_2$. Let $k_i := k(y_i, z)$ and $\ell_i := \ell(y_i, z)$ for $i = 1, 2$. Suppose for a contradiction that $F(k_1, \ell_1) > F(k_2, \ell_2)$. By proposition A-4 we have $k_1 \leq k_2$. Since F is increasing, it must be that $0 < \ell_2 < \ell_1 < 1$. As F is concave, u is strictly concave and V is strictly concave in its first argument, proposition A-6 now gives

$$\begin{aligned} v'(1 - \ell_2) &= \beta \int V'[z' F(k_2, \ell_2), z'] F_2'(k_2, \ell_2) z' \Pi(z, dz') \\ &> \beta \int V'[z' F(k_1, \ell_1), z'] F_2'(k_2, \ell_2) z' \Pi(z, dz') \\ &\geq \beta \int V'[z' F(k_1, \ell_1), z'] F_2'(k_1, \ell_2) z' \Pi(z, dz') \\ &> \beta \int V'[z' F(k_1, \ell_1), z'] F_2'(k_1, \ell_1) z' \Pi(z, dz') = v'(1 - \ell_1) \end{aligned}$$

This is a contradiction, since u_2' is strictly decreasing. Hence $F(k_1, \ell_1) \leq F(k_2, \ell_2)$. $\square$

Proposition A-8. *The optimal consumption policy $c(\cdot, z)$ is strictly increasing for each $z > 0$.*

Proof. Fix $z > 0$. If $y = 0$, then clearly $c(y, z) = 0$. For $y > 0$, by proposition A-5, $k(y, z)$ is interior, and thus $c(y, z) > 0$. By propositions A-2 and A-3, we have $V_1'(y, z) = u' \circ c(y, z)$, or, equivalently, $c = (u')^{-1} \circ V_1'(\cdot, z)$. Since $V_1'(\cdot, z)$ and $(u')^{-1}$ are strictly decreasing, $c(\cdot, z)$ strictly increasing. $\square$

All of the results in theorem 2.1 have now been verified.

Appendix B: Proof of Theorem 3.1

Throughout this section we maintain the assumptions of theorem 3.1. In particular, we suppose that assumptions 2.2–2.4 and 3.1 all hold, and at least one of assumptions 3.2 and 3.3 holds. Recall from the discussion in section 3 that in this IID setting, the value and policy functions are constant in z, and hence we simplify notation by omitting this constant second argument. Hence $V(y, z)$ is written as $V(y)$ and treated as a function on $\mathbb{R}_+$, and likewise for σ, k, ℓ, c and g. For example, the Euler equations from theorem 2.1 become

$$u' \circ c(y) = \beta \int u' \circ c[zg(y)]\, zF_1'(k(y), \ell(y))\mu(dz), \quad \text{and} \tag{B-1}$$

$$v'(1 - \ell(y)) = \beta \int u' \circ c[zg(y)]\, zF_2'(k(y), \ell(y))\mu(dz). \tag{B-2}$$

We let Q be the stochastic kernel on $\mathbb{R}_{++}$ representing optimal output dynamics. Thus,

$$Q(y, B) = \mathbb{P}\{\omega \in \Omega : z_t(\omega)g(y) \in B\} = \mu\{z \in \mathbb{R}_{++} : zg(y) \in B\}. \tag{B-3}$$

We now introduce some preliminary results that are needed for the proof of theorem 3.1.

Proposition B-1. *The stochastic kernel Q is increasing.*

Proof. Fix $h \in iM(\mathbb{R}_{++})$, fix $z \in \mathbb{R}_{++}$ and $y, y' \in \mathbb{R}_{++}$ with $y \leq y'$. Since g is increasing, it follows from the monotonicity of the integral that $Qh(y) = \int h[zg(y)]\mu(dz) \leq \int h[zg(y')]\nu(dz) = Qh(y')$. Hence $Qh \in iM(\mathbb{R}_{++})$ as required. $\square$

Lemma B-1. *If F satisfies assumption 3.3, then* $\lim_{y \to 0}\ F_1'(k(y), \ell(y)) = \infty$.

Proof. Under assumption 3.3 we have $F(0, \ell) = 0$ for all ℓ, and hence, from the Bellman equation, $V(0) = \max_{\ell \in [0,1]} \{v(1-\ell) + \beta V(0)\}$. It follows immediately that $\ell(0) = 0$. Moreover, since ℓ is single-valued and continuous, we must then have $\lim_{y \to 0} \ell(y) = 0$. We now claim that $\lim_{y \to 0} F_2'(k(y), \ell(y)) = 0$. To see this, recall the Euler equation (B-2). As $y \to 0$, both $k(y), \ell(y) \to 0$. The left-hand-side of (B-2) converges to the constant $v'(1)$, while on the right-hand-side, for any given $z \in \mathbb{R}_{++}$, we have $u' \circ c[zg(y)] \to u'(c(0)) = \infty$. Thus, for (B-2) to hold, it must be that $F_2'(k(y), \ell(y)) \to 0$. Since F is homogeneous of degree one, its derivatives are homogeneous of degree zero, and hence we can write this convergence as $F_2'(1, \ell(y)/k(y)) \to 0$ as $y \to 0$. This implies that $\ell(y)/k(y) \to \infty$, or, conversely, that $k(y)/\ell(y) \to 0$. Using homogeneity of degree zero again, combined with assumption 2.2, we obtain

$$\lim_{y \to 0} F_1'(k(y), \ell(y)) = \lim_{y \to 0} F_1'(k(y)/\ell(y), 1) = \infty,$$

as was to be shown. $\square$

Lemma B-2. *Let $W_1 := \sqrt{u' \circ c}$. If at least one of assumptions 3.2 and 3.3 hold, then there exists a $\delta > 0$ and an $a_1 \in (0,1)$ such that*

$$(QW_1)(y) \leq a_1 W_1(y), \quad \forall y < \delta. \tag{B-4}$$

Proof. The argument is an extension of techniques used in [25]. We begin the proof with the claim that

$$(QW_1)(y) \leq W_1(y) \left[\int \frac{1}{\beta z F_1'(k(y), \ell(y))} \mu(dz) \right]^{1/2}. \tag{B-5}$$

To see this, note that

$$(QW_1)(y) = \int \{u' \circ c[zg(y)]\}^{1/2} \mu(dz) = \int \left[\beta\, u' \circ c[zg(y)] \frac{zF_1'(k(y)\ell(y))}{\beta z F_1'(k(y)\ell(y))} \right]^{1/2} \mu(dz).$$

Applying the Cauchy-Schwartz inequality yields

$$(QW_1)(y) \leq \left[\beta \int u' \circ c[zg(y)] z F_1'(k(y)\ell(y)) \mu(dz) \right]^{1/2} \left[\int \frac{1}{\beta z F_1'(k(y), \ell(y))} \mu(dz) \right]^{1/2}.$$

An application of the Euler equation (B-1) and the definition of W_1 produces (B-5).

In light of (B-5), to complete the proof of lemma B-2, we need only show that

$$\lim_{y \to 0} \left[\int \frac{1}{\beta z F_1'(k(y), \ell(y))} \mu(dz) \right]^{1/2} < 1. \tag{B-6}$$

Consider first the case where F satisfies assumption 3.2. From our restrictions on the derivatives of F in assumption 2.2, we have

$$\left[\int \frac{1}{\beta z F_1'(k(y), \ell(y))} \mu(dz)\right]^{1/2} \leq \left[\int \frac{1}{\beta z F_1'(y, 0)} \mu(dz)\right]^{1/2}.$$

Equation (B-6) is now immediate from (13).

Alternatively, suppose that assumption 3.3 holds. Then, by proposition B-1, we have

$$\lim_{y \to 0} \int \frac{1}{\beta z F_1'(k(y), \ell(y))} \mu(dz) = 0.$$

Once again, (B-6) is established, and the proof of lemma B-2 is done.[13] □

Lemma B-3. *If the conditions of lemma B-2 hold, then, given δ as in (B-4), there exists a $b_1 < \infty$ such that $(QW_1)(y) \leq b_1$ for all $y \geq \delta$.*

Proof. Since the optimal consumption function c is increasing on $\mathbb{R}_{++}$, the composition $u' \circ c$ is decreasing, and hence $-W_1$ is increasing. Since Q is increasing (proposition B-1), it follows that $Q(-W_1)$ is also increasing, and hence $(Q(-W_1))(\delta) \leq (Q(-W_1))(y)$ for all $y \geq \delta$. Using the fact that Q is a linear operator, we obtain $(QW_1)(y) \leq (QW_1)(\delta)$ for all $y \geq \delta$, and hence the claim in the lemma is satisfied with $b_1 := (QW_1)(\delta)$. □

Lemma B-4. *If W_2 is the identity on $\mathbb{R}_{++}$, then there exists positive constants $a_2 < 1$ and $b_2 < \infty$ such that $QW_2 \leq a_2 W_2 + b_2$ on $\mathbb{R}_{++}$.*

Proof. Since F is increasing in both arguments we have

$$(QW_2)(y) = \int z\mu(dz) g(y) \leq \int z\mu(dz) F(y, 1). \tag{B-7}$$

In view of (16), there exists a constant $a_2 \in (0, 1)$ and a $\bar{y} > 0$ such that

$$y > \bar{y} \implies \int z F(y, 1)\mu(dz) < a_2 y. \tag{B-8}$$

On the other hand,

$$y \leq \bar{y} \implies \int z F(y, 1)\mu(dz) \leq \int z F(\bar{y}, 1)\mu(dz) =: b_2. \tag{B-9}$$

Combining (B-7), (B-8) and (B-9) yields $(QW_2)(y) \leq a_2 y + b_2$ for all $y > 0$. This amounts to the claim in lemma B-4.

[13] We are using the finiteness of $\mathbb{E}\,(1/z_t)$ imposed in assumption 3.1.

Turing to the case of assumption 3.3, let $\gamma \in (0, 1)$ be such that $a_2 := \gamma \int z\mu(dz) < 1$. As $\lim_{k\to\infty} F'(k, 1) = 0$, there exists a $d < \infty$ such that $F(y, 1) \leq \gamma y$ for all $y > d$. This leads to the estimate $(QW_2)(y) \leq a_2 y$ for all $y > d$. On the other hand, we clearly have

$$(QW_2)(y) \leq \int zF(d, 1)\mu(dz) =: b_2, \quad \forall y \leq d.$$

Hence $(QW_2)(y) \leq a_2 W_2(y) + b_2$ for all $y > 0$. The proof of lemma B-4 is now complete. □

Proposition B-2. *The stochastic kernel Q is bounded in probability.*

Proof. Let W_1 and W_2 be as in lemmas B-2–B-4, and let $W(y) := W_1(y) + W_2(y) := \sqrt{u_1' \circ c} + y$. By (20), the function W is coercive on $\mathbb{R}_{++}$. Observe that by linearity of the operator Q and lemmas B-2–B-4, we have $QW = QW_1 + QW_2 \leq a_1 W_1 + a_2 W_2 + b_1 + b_2$ pointwise on $\mathbb{R}_{++}$. Letting $\lambda := \max\{a_1, a_2\}$ and $B := b_1 + b_2$ yields $QW \leq \lambda W + B$ with $\lambda < 1$. Using this result, we will show that $\limsup_{t\to\infty}(Q^t W)(y)$ is finite for any $y \in \mathbb{R}_{++}$. To see that this is so, observe that repeatedly applying Q to $QW \leq \lambda W + B$ yields

$$Q^t W \leq \lambda^t W + \frac{B(1 - \lambda^t)}{1 - \lambda}.$$

Taking the limit gives finiteness of $\limsup_{t\to\infty}(Q^t W)(y)$ for any $y \in \mathbb{R}_{++}$. As per Meyn and Tweedie [22, lemma D.5.3], this implies that Q is bounded in probability. □

Proposition B-3. *The stochastic kernel Q is Feller.*

Proof. Fix $h \in C(\mathbb{R}_{++})$. Boundedness of Qh is obvious. To obtain continuity, recall that g is continuous. Fix $y \in \mathbb{R}_{++}$ and $y_n \to y$. By the dominated convergence theorem, $(Qh)(y_n) = \int h[g(y_n, z)]\mu(dz) \to \int h[g(y, z)]\mu(dz) = (Qh)(y)$ as $n \to \infty$. Hence Q is Feller as claimed. □

Proof of theorem 3.1. The Feller property (proposition B-3) combined with boundedness in probability (proposition B-2) implies existence of a stationary distribution by [22, proposition 12.1.3]. □

Appendix C: Proof of Theorem 4.1

In this section we maintain the assumptions of theorem 4.1. In particular, we suppose that assumptions 2.2–2.4 and 3.1 all hold, at least one of assumptions 3.2 and 3.3 holds, and assumption 4.1 holds.

Proof of theorem 4.1. We have shown that Q is increasing, bounded in probability and Feller (propositions B-1–B-3). Since assumption 4.1 is also taken to be valid, the kernel Q is clearly downward reaching. It follows that the stationary distribution π of Q is unique and globally stable [19, theorem 3.2 and corollary 4.1]. This shows parts 1–2 of theorem 4.1. Part 3 follows from [19, theorem 3.3 and 3.4]. $\square$

Appendix D: Proofs of Theorem 4.2–4.3

In this section we maintain the assumptions of theorem 4.2. In particular, we suppose that assumptions 2.2–2.4 and 3.1 all hold, at least one of assumptions 3.2 and 3.3 holds, and assumption 4.2 holds. The results in this section extend those in Nishimura and Stachurski [25, theorems 3.1–3.2].

Lemma D-1. *The stochastic kernel Q is irreducible with respect to the lognormal density on $\mathbb{R}_{++}$.*

Proof. Take any $B \in \mathscr{B}(\mathbb{R}_{++})$ with positive measure under the lognormal density, and fix any $y_0 \in \mathbb{R}_{++}$. It is easy to check that the set $B/g(y_0)$ has positive Lebesgue measure, and hence positive μ-measure (by the strict positivity of μ in assumption 4.2). In addition, we have $\mathbb{P}\{y_1 \in B \mid y_0\} = \mu\{z \in \mathbb{R}_{++} : zg(y_0) \in B\} = \mu(B/g(y_0))$. The final term is strictly positive and equal to $Q(y_0, B)$. The claim in the lemma follows. $\square$

Lemma D-2. *Every compact subset of $\mathbb{R}_{++}$ is a $\mathscr{C}$-set, and Q is strongly aperiodic.*

Proof. Regarding the first claim, it follows from the definition that measurable subsets of $\mathscr{C}$-sets are themselves $\mathscr{C}$-sets. Hence it suffices to show that the interval $C_n := [1/n, n]$ is a $\mathscr{C}$-set for every $n \in \mathbb{N}$. Pick any $n \in \mathbb{N}$. By interiority and monotonicity of g, $0 < g(1/n) \leq g(y) \leq g(n) < \infty$ for all $y \in C_n$. Since μ is continuous and strictly positive, it follows that

$$\inf_{C_n \times C_n} \mu\left(\frac{y'}{g(y)}\right)\frac{1}{g(y)} =: r > 0.$$

Let λ be Lebesgue measure. Choosing ϕ to be the measure defined by $\phi(B) = r \cdot \lambda(B \cap C_n)$ and picking any $y \in C_n$, we have

$$Q(y, B) = \int_B \mu\left(\frac{y'}{g(y)}\right)\frac{1}{g(y)}dy' \geq \int_{B \cap C_n} \mu\left(\frac{y'}{g(y)}\right)\frac{1}{g(y)}dy' \geq r\lambda(B \cap C_n) = \phi(B).$$

In other words, (21) is valid.

Regarding the claim that Q is strongly aperiodic, we need to show that $\phi(C_n) > 0$, where ϕ and C_n are as defined immediately above. Since $\phi(C_n) = r\lambda(C_n) = r(n - 1/n)$, this follows immediately from positivity of r. $\square$

Lemma D-3. *If p is as in assumption 4.2, and $W_3(y) := y^p$, then there exists positive constants $a_3 < 1$ and $b_3 < \infty$ such that $QW_3 \leq a_3 W_3 + b_3$ on $\mathbb{R}_{++}$.*

Proof. Since F is increasing in both arguments we have

$$(QW_3)(y) = \int z^p \mu(dz) g(y)^p \leq \int z^p \mu(dz) F(y,1)^p. \tag{D-1}$$

In view of (16), there exists a constant $a_3 \in (0,1)$ and a $\bar{y} > 0$ such that

$$y > \bar{y} \implies \int z^p F(y,1)^p \mu(dz) < a_3 y^p. \tag{D-2}$$

On the other hand,

$$y \leq \bar{y} \implies \int z^p F(y,1)^p \mu(dz) \leq \int z^p F(\bar{y},1)^p \mu(dz) =: b_3. \tag{D-3}$$

Combining (D-1), (D-2) and (D-3) yields $QW_3 \leq a_3 W_3 + b_3$. □

Lemma D-4. *There exists a $\mathscr{C}$-set $C \in \mathscr{B}$, a $\theta < 1$, an $L < \infty$ and $\hat{W} \colon \mathbb{R}_{++} \to [1,\infty)$ such that the pointwise inequality $Q\hat{W} \leq \theta \hat{W} + L\mathbb{1}_C$ holds everywhere on $\mathbb{R}_{++}$.*

Proof. Recall from the proof of proposition B-2 that if $W := W_1 + W_2$, then there exist positive constants $\lambda < 1$ and $B < \infty$ such that $QW \leq \lambda W + B$ pointwise on $\mathbb{R}_{++}$. An identical argument to the proof given there (but replacing W_2 with W_3, a_2 with a_3 and b_2 with b_3) shows that if instead we take $W := W_1 + W_3$, then once again there exist positive constants $\lambda < 1$ and $B < \infty$ such that $QW \leq \lambda W + B$ pointwise on $\mathbb{R}_{++}$. Using this choice of W, let $\hat{W} := W + 1$, let θ be any number in $(\lambda, 1)$ and let $L := B + 1$. Choose C to be a compact set such that $\hat{W}(y) \geq (B+1)/(\theta - \lambda)$ whenever $y \notin C$. (Existence of C follows from the definition of W.) We claim that for this choice of $\hat{W}$, θ, L and C, the conditions of lemma D-4 are satisfied. To see that this is in fact the case, observe first that C is a $\mathscr{C}$-set by lemma D-2. Second, if $y \in C$, then we have

$$Q\hat{W}(y) = QW(y) + 1 \leq \lambda W(y) + B + 1 \leq \theta W(y) + L\mathbb{1}_C(y) \leq \theta \hat{W}(y) + L\mathbb{1}_C(y).$$

In other words, if $y \in C$, then the inequality in the statement of the lemma is verified.

It remains only to check that the same inequality holds when $y \notin C$. To see that this is so, recall that for such a y we have $\hat{W}(y) \geq (B+1)/(\theta - \lambda)$, and hence

$$Q\hat{W}(y) \leq \left[\lambda + \frac{B+1}{\hat{W}(y)}\right] \hat{W}(y) \leq \theta \hat{W}(y) = \theta \hat{W}(y) + L\mathbb{1}_C(y).$$

We have now verified that $Q\hat{W} \leq \theta \hat{W} + L\mathbb{1}_C$ holds everywhere on $\mathbb{R}_{++}$. □

Proof of theorem 4.2. Let $\hat{W}$ be the function $W_1 + W_3 + 1$ defined in lemma D-4. If follows from lemmas D-1–D-4 and [22, theorem 16.1.2] that the stochastic kernel Q is so-called *$\hat{W}$-uniformly ergodic* (see [22, chapter 16] for the definition). Geometric ergodicity is an immediate consequence [22, theorem 16.0.1]. □

Proof of theorem 4.3. In the proof of theorem 4.2, it was shown that that Q is $\hat{W}$-uniformly ergodic. The claims in theorem 4.3 now follow from [22, theorem 17.0.1]. □

References

[1] Aliprantis, C.D. and K.C. Border (2006): *Infinite Dimensional Analysis: A Hitchhiker's Guide*, 3rd Edition, Springer, Berlin.

[2] Balbus, L., K. Reffett and L. Woźny (2012): "Stationary Markovian Equilibrium in Altruistic Stochastic OLG Models with Limited Commitment," *Journal of Mathematical Economics*, in press.

[3] Bosi, S. and C. Le Van (2011): "On the Existence of a Ramsey Equilibrium with Endogenous Labor Supply and Borrowing Constraints," DEPOCEN Working Paper Series No. 2011/24.

[4] Bhattacharya, R. and M. Majumdar (2007): *Random Dynamical Systems: Theory and Applications*, Cambridge University Press.

[5] Brock, W.A. and M. Majumdar (1978): "Global Asymptotic Stability Results for Multisector Models of Optimal Growth under Uncertainty When Future Utilities Are Discounted," *Journal of Economic Theory*, 18, 225–243.

[6] Brock, W.A. and L. J. Mirman (1972): "Optimal Economic Growth and Uncertainty: the Discounted Case," *Journal of Economic Theory*, 4, 479–513.

[7] Brock, W.A. and L. J. Mirman (1973): "Optimal Economic Growth and Uncertainty: the No Discounting Case," *International Economic Review*, 14, 560–573.

[8] Cai, Y., T. Kamihigashi and J. Stachurski (2013): "Stochastic Optimal Growth with Risky Labor Supply," RSE Discussion Paper, Australian National University.

[9] Coleman, W. J. (1997): "Equilibria in Distorted Infinite-Horizon Economies with Capital and Labor," *Journal of Economic Theory*, 72, 446–461.

[10] Datta, M., L.J. Mirman and K.L. Reffett (2002): "Existence and Uniqueness of Equilibrium in Distorted Dynamic Economies with Capital and Labor," *Journal of Economic Theory*, 103, 377–410.

[11] Donaldson, J.B. and R. Mehra (1983): "Stochastic Growth with Correlated Production Shocks," *Journal of Economic Theory,* 29, 282–312.

[12] Durlauf, S.N. and D. Quan (1999): "The New Empirics of Economic Growth," NBER Working Papers No. 6422.

[13] Galí, J. (2008): *Monetary Policy, Inflation, and the Business Cycle,* Princeton University Press, Princeton and Oxford.

[14] Goenka, A., C. Le Van and Nguyen (2011): "Existence of Competitive Equilibrium in an Optimal Growth Model with Heterogeneous Agents and Endogenous Leisure," DEPOCEN Working Paper Series No. 2011/21.

[15] Greenwood, J. and G.W. Huffman (1995): "On the Existence of Nonoptimal Equilibria in Dynamic Stochastic Economies," *Journal of Economic Theory,* 65, 611–623.

[16] Hansen, G.D. (1985): "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics,* 16, 309–327.

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

[18] Kamihigashi, T. (2007): "Stochastic Optimal Growth with Bounded or Unbounded Utiltity and Bounded or Unbounded Shocks," *Journal of Mathematical Economics,* 43, 477–500.

[19] Kamihigashi, T. and J. Stachurski (2011): "Existence, Stability and Computation of Stationary Distributions: An Extension of the Hopenhayn-Prescott Theorem," RIEBA Discussion paper DP2011-32, Kobe University.

[20] Kydland, F.E. and E.C. Prescott (1982): "Time to Build and Aggregate Fluctuations," *Econometrica,* 50, 1345–1370.

[21] John B. Long, Jr. and Charles I. Plosser, 1983. "Real Business Cycles" *Journal of Political Economy,* 91(1), 39-69.

[22] Meyn, S. and R. L. Tweedie (2009): *Markov Chains and Stochastic Stability,* 2nd Edition, Cambridge University Press, Cambridge.

[23] Mirman, L. J. (1973): "On the Existence of Steady State Measure for One Sector Growth Models with Uncertain Technology," *International Economic Review,* 13, 271–286.

[24] Mirman, L. J. and I. Zilcha (1975): "On Optimal Growth under Uncertainty," *Journal of Economic Theory,* 11, 329–339.

[25] Nishimura, K. and J. Stachurski (2005): "Stability of Stochastic Optimal Growth Models: A New Approach," *Journal of Economic Theory,* 122, 100–118.

[26] Razin, A. and J.A. Yahav (1979): "On Stochastic Models of Economic Growth," *International Economic Review*, 20, 599–604.

[27] Stachurski, J. (2002): "Stochastic Optimal Growth with Unbounded Shock," *Journal of Economic Theory*, 106, 40–65.

[28] Stachurski, J. (2007): "Log-Linearization of Stochastic Economic Models," *Journal of Difference Equations and Applications*, 13, 217–222.

[29] Stachurski, J. (2009): *Economic Dynamics: Theory and Computation*, MIT Press, Cambridge, MA.

[30] Stokey, N.L. and R.E. Lucas, Jr., with E.C. Prescott (1989): *Recursive Methods in Economic Dynamics*, Harvard University Press, Cambridge, MA.

[31] Zhang, Y. (2007): "Stochastic Optimal Growth with a Non-Compact State Space," *Journal of Mathematical Economics*, 43, 115–129.