ONLINE APPENDIX FOR EXISTENCE OF RECURSIVE CONSTRAINED OPTIMA IN THE HETEROGENEOUS AGENT NEOCLASSICAL GROWTH MODEL

Akshay Shanker[1]

Australian National University (ANU)

APPENDIX C: MATHEMATICAL PRELIMINARIES

C.1. *Correspondences*

Let (X, τ) and (Y, τ') be topological vector spaces. A correspondence from a space X to Y is a set valued function denoted by $\Gamma\colon X \twoheadrightarrow Y$. The image of a subset A of X under the correspondence Γ will be the set

$$\Gamma(A) := \{y \in Y | y \in \Gamma(x) \text{ for some } x \in A\}$$

A correspondence will be called **compact valued** if $\Gamma(x)$ is compact for $x \in X$. We can also define the graph of a correspondence as $\operatorname{Gr}\Gamma := \{(x, y) | y \in \Gamma(x)\}$. A correspondence has a (sequentially) **closed graph** if $\operatorname{Gr}\Gamma$ is (sequentially) closed in $X \times Y$.

The correspondence Γ is **upper hemi-continuous** if for every x and neighbourhood U of $\Gamma(x)$, there is a neighbourhood V of x such that $z \in V$ implies $\Gamma(z) \subset U$. A correspondence is **lower hemi-continuous** if at each x, for every open set U such that $\Gamma(x) \cap U \neq \emptyset$ there is a neighbourhood V of x such that for any $z \in V$ we have $\Gamma(z) \cap U \neq \emptyset$.

Upper hemicontinuous correspondences need not be compact valued or have closed graph. Closed graph correspondences also need not be upper hemi-continuous. However,

LEMMA C.1 *If $\Gamma\colon X \twoheadrightarrow Y$ is upper hemicontinuous and compact valued, then for $C \subset X$ such that C is compact, $\Gamma(C)$ is compact.*

See Lemma 17.8 by Aliprantis and Border (2006)) for a proof.

C.2. *Semicontinuity*

A function $f\colon X \to \mathbb{R} \cup \{-\infty, +\infty\}$ is sequentially **upper semi-continuous** if the upper contour sets

$$UC_f(\epsilon) := \{x \in X \,|\, f(x) \geq \epsilon\} \tag{C.1}$$

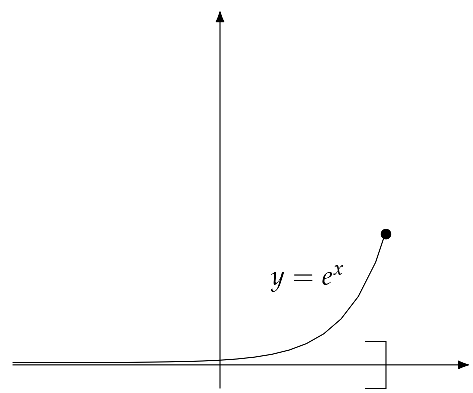
are sequentially closed for all $\epsilon \in \mathbb{R}$. When writing $UC_f(\epsilon)$, I will omit the term ϵ if the choice of ϵ is clear.

[1]Please downloand the latest version of this appendix at `https://github.com/akshayshanker/Existence_of_Social_Optimia_Aiyagari`

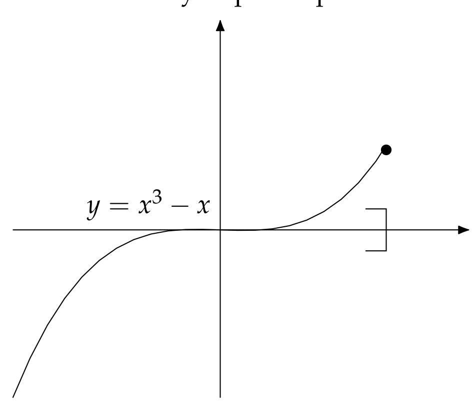
Equivalently, f is sequentially upper semi-continuous if for $x_n \to x$ with $x \in X$, we have $\limsup_{n\to\infty} f(x_n) \leq f(x)$ (see Proposition 4.5 by Barbu and Precupanu (2012) or 2.10 by Aliprantis and Border (2006).)

Let D be a subset of $\mathbb{R} \cup \{-\infty, +\infty\}$. A function $f\colon X \to D$ is **sup-compact** if the sets $UC_f(\epsilon)$ are sequentially compact for all $\epsilon \in \mathbb{R}$.

If X is not compact and D is bounded below, then f cannot be sup-compact; we will need a weaker condition for bounded functions that captures the same idea as sup-compactness. Mild sup-compactness generalises sup-compactness by only requiring upper contour sets away from the infimum to be sequentially compact. Suppose $D \subset \mathbb{R}$, then the function f will be called **mildly sup-compact** if the sets $UC_f(\epsilon)$ are sequentially compact for all $\epsilon > \inf f$.



Mildly Sup-Compact



Sup-Compact

FIGURE 1.— Not all upper-contour sets of the the mildly sup-compact function are compact since it is bounded below, but the function still has a maximum.

C.3. *Probability and Conditional Expectation*

Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space. We work with the following definition of conditional expectation.

DEFINITION C.1 Let $\mathcal{H} \subset \Sigma$ be a sub σ-algebra of Σ and let $x \colon \Omega \to \mathbb{R}^n$ be a random variable. The **conditional expectation** of x given $\mathcal{H}$ is any $\mathcal{H}$-measurable random variable y which satisfies

$$\int_B y \,\mathrm{d}\mathbb{P} = \int_B x \, d\mathbb{P}, \qquad B \in \mathcal{H}$$

If y is a conditional expectation of x given $\mathcal{H}$, we write $y = \mathbb{E}(x|\mathcal{H})$.

Recall the definition of a σ-algebra generated by a family of functions $\{y_i\}_{i\in F}$:

$$\sigma(\{y_i\}_{i\in F}) \colon = \sigma(\cup_{i\in F}\sigma(y_i))$$

Recall also that if $\mathcal{A}$ and $\mathcal{B}$ are independent sub σ-algebras of Σ and if a sub-σ-algebra $\mathcal{C}$ satisfies $\mathcal{C} \subset \mathcal{A}$, then $\mathcal{C}$ and $\mathcal{B}$ will be independent.

The following facts are standard.

FACT C.1 *Let $\mathcal{G}$ and $\mathcal{H}$ be sub σ-algebras of Σ and let $x \colon \Omega \to \mathbb{R}$ be a random variable. If $\mathcal{H}$ and $\sigma(\mathcal{G}, \sigma(x))$ are independent, then $\mathbb{E}(x|\sigma(\mathcal{G}, \mathcal{H})) = \mathbb{E}(x|\mathcal{G})$.*

PROOF: See section 9.7 in Williams (1991). *Q.E.D.*

FACT C.2 *If $\mathcal{A}, \mathcal{B}$ and $\mathcal{C}$ are σ-algebras, then $\sigma(\mathcal{A} \cup \mathcal{B}) \cup \mathcal{C} \subset \sigma(\mathcal{A} \cup \mathcal{B} \cup \mathcal{C})$*

PROOF: Recall that for collections of sets $\mathcal{A}$ and $\mathcal{B}$, with $\mathcal{A} \subset \mathcal{B}$, we have $\sigma(\mathcal{A}) \subset \sigma(\mathcal{B})$.

Now pick any B such that $B \in \sigma(\mathcal{A} \cup \mathcal{B}) \cup \mathcal{C}$. We have either $B \in \sigma(\mathcal{A} \cup \mathcal{B})$ or $B \in \mathcal{C}$. If $B \in \sigma(\mathcal{A} \cup \mathcal{B})$, then since $\sigma(\mathcal{A} \cup \mathcal{B}) \subset \sigma(\mathcal{A} \cup \mathcal{B} \cup \mathcal{C})$, $B \in \sigma(\mathcal{A} \cup \mathcal{B} \cup \mathcal{C})$. Alternatively, if $B \in \mathcal{C}$, then because $\mathcal{C} \subset \mathcal{A} \cup \mathcal{B} \cup \mathcal{C} \subset \sigma(\mathcal{A} \cup \mathcal{B} \cup \mathcal{C})$, we can conclude that $B \in \sigma(\mathcal{A} \cup \mathcal{B} \cup \mathcal{C})$. *Q.E.D.*

FACT C.3 ***(Doob-Dynkin)*** *Let $(\Omega, \Sigma, \mathbb{P})$ be a probability space. Let $f \colon \Omega \to \mathbb{R}^k$ and $g \colon \Omega \to \mathbb{R}^n$. The generated σ-algebras satisfy $\sigma(f) \subset \sigma(g)$ if and only if there exists a measurable function $h \colon \mathbb{R}^n \to \mathbb{R}^k$ such that $f = h \circ g$.*

PROOF: See Lemma 1.13 by Kallenberg (1997). *Q.E.D.*

FACT C.4 ***(Jensen's Inequality)***

Let x be a random variable on a probability space $(\Omega, \Sigma, \mathbb{P})$. If $c \colon \mathbb{R} \to \mathbb{R}$ is convex, and $\mathbb{E}|c(x)| < \infty$, then

$$\mathbb{E}(c(x)|\mathcal{H}) \geq c(\mathbb{E}(x|\mathcal{H}))$$

almost everywhere, for any sub σ - algebra $\mathcal{H}$

For a proof, see section 9.7 in Williams (1991).

FACT C.5 ***(Reverse Fatou's Lemma)*** *If (f_n) is a sequence of real valued functions defined on a measure space (Ω, Σ, μ), and each f_n satisfies $f_n \leq g$ and $f_n \geq 0$, where g is an integrable function on (Ω, Σ, μ), then*

$$\limsup_{n\to\infty} \int f_n d\mu \leq \int \limsup_{n\to\infty} f_n d\mu$$

For a proof, see 5.4 in Williams (1991).

C.4. *Constructing Probability Space for the Sequential Planner*

Here I discuss how to construct a probability space $(\Omega, \mathscr{F}, \mathbb{P})$ described in section 2.3. Let $\tilde{\mu}_0$ be the marginal distribution of x_0^i and let $(\Omega, \mathscr{F}, \mathbb{P})$ denote the product of the measure spaces $(A, \mathscr{B}(A), \tilde{\mu}_0)$, $(E, \mathscr{B}(E), \psi)$ and $([0,1], \mathscr{B}([0,1]), \lambda)^{\mathbb{N}}$. The spaces A, E and $[0,1]$ are second countable, hence, following Lemma 1.2 by Kallenberg (1997), $\mathscr{F} = \mathscr{B}(\bar{Z})$, where $\bar{Z} = A \times E \times [0,1]^{\mathbb{N}}$. Moreover, the product topology on $\bar{Z}$ will have a countable basis (see Kelley (1975) p103). Since each open set in $\bar{Z}$ will be a countable union of basic open sets, the (countable) basic open sets generate the Boreal sigma algebra $\mathscr{B}(\bar{Z})$.

Let $(\pi_t)_{t=0}^{\infty}$ be the projection maps on Ω and define x_0 on Ω as $x_0(\omega) = \pi_0(\omega)$ for $\omega \in \Omega$, $e_0 = \pi_1(\omega)$ and $\eta_t = \pi_{t+1}(\omega)$ for $t \geq 1$. The sequence $\{x_0, e_0, \eta_1, \eta_2, \dots\}$ will be independent and the marginal distribution of x_0 and e_0 will be $\tilde{\mu}_0$ and ψ (for details on construction of the product measure $\mathbb{P}$, see section IV, Theorem 4.7 and section IV, 5 by Çinlar (2011) or Theorem B, section 38 by Halmos (1974)).

For each t, define $e_{t+1} = G(e_t, \eta_{t+1})$ where $G(x, \cdot)$ is the inverse of $y \mapsto Q(x, (-\infty, y))$. The process $(e_t)_{t=0}^{\infty}$ will be Markov with kernel Q (for details, see Borovkov (2013), section 17.1).

APPENDIX D: PROOFS

Proofs for Section 2.1

Consider the setting described in section 2.1.

PROOF OF CLAIM 2.1: We show the joint (empirical) distribution of $i \mapsto \{x_t^i, e_t^i\}$ is μ_t. By definition, μ_t is the distribution of $\{x_t^i, e_t^i\}$, holding i fixed. Recall since $(\mu_t)_{t=0}^{\infty}$ satisfies (2) for a sequence of policy functions $(h_t)_{t=0}^{\infty}$, x_t^i will be $\{x_0^i, e_0^i, \dots, e_{t-1}^i\}$ measurable and there exists a measurable function g such that $x_t^i = g(x_0^i, e_0^i, \dots e_{t-1}^i)$. As such, for $B_A \times B_E \in \mathscr{B}(S)$,

$$\begin{aligned}\mu_t(B_A \times B_E) = \bar{\mathbb{P}}\{x_t^i, e_t^i \in B\} &= \bar{\mathbb{P}}\{x_0^i, e_0^i, \dots, e_{t-1}^i, e_t^i \in g^{-1}(B_A) \times B_E\} \\ &= P\big(g^{-1}(B_A) \times B_E \times E^{\mathbb{N}}\big)\end{aligned}$$

By Assumption 2.2,

$$\begin{aligned}\zeta\{i \in I \,|\, x_t^i, e_t^i \in B_A \times B_E\} &= \zeta\{i \in I \,|\, x_0^i, e_0^i, \dots, e_{t-1}^i, e_t^i \in g^{-1}(B_A) \times B_E\} \\ &= P\big(g^{-1}(B_A) \times B_E \times E^{\mathbb{N}}\big)\end{aligned}$$

Conclude $\mu_t(B_A \times B_E) = \zeta\{x_t^i, e_t^i \in B_A \times B_E\}$ with probability one for any $B_A \times B_E \in \mathscr{B}(S)$, implying μ_0 is the joint (empirical) distribution of $i \mapsto \{x_t^i, e_t^i\}$. The integration at Equation (3) is then valid by Theorem 2.4, ch. II by Çinlar (2011).

Q.E.D.

The following claim defines the distributions of $(x_t^i)_{t=0}^{\infty}$ and $(e_t^i)_{t=0}^{\infty}$ under a sequence of policy functions $(h_t)_{t=0}^{\infty}$.

CLAIM D.1 *Let Assumption 2.1 hold. If $(x_t^i)_{t=0}^{\infty}$ is defined by the recursion in Equation* (1) *of the main paper for each $i \in [0,1]$, then for each $t \in \mathbb{N}$ and $i \in [0,1]$, $\{x_{t+1}^i, e_{t+1}^i\} \sim \mu_{t+1}$ and $\{x_t^i, e_t^i\} \sim \mu_t$ where μ_{t+1} and μ_t satisfy the recursion* (2) *of the main paper.*

PROOF: We use proof by induction and first show the claim holds for $t = 0$. By Assumption 2.1, let μ_0 be given as the joint distribution of x_0^i and e_0^i. We then have $\{x_0^i, e_0^i\} \sim \mu_0$ for each i.

Let $\mathscr{F}_1^i$ be the σ-algebra generated by $\{x_0^i, e_0^i\}$. The joint distribution of x_1^i and e_1^i will be

$$\begin{aligned}
\mu_1(B_A \times B_E) :&= \bar{\mathbb{P}}\left\{x_1^i \in B_A, e_1^i \in B_E\right\} \\
&= \mathbb{E}\left\{\mathbb{E}\left\{\mathbb{1}_{B_A}(x_1^i) \times \mathbb{1}_{B_E}(e_1^i)|\mathscr{F}_1^i\right\}\right\} \\
&= \mathbb{E}\left\{\mathbb{1}_{B_A}(x_1^i) \times \mathbb{E}\left\{\mathbb{1}_{B_E}(e_1^i)|\mathscr{F}_1^i\right\}\right\} \\
&= \mathbb{E}\left\{\mathbb{1}_{B_A}(x_1^i) \times \bar{\mathbb{P}}\left\{e_1^i \in B_E|e_0^i\right\}\right\} \\
&= \int\int \mathbb{1}_{B_A}\left\{h_0(x,e)\right\} Q(e, B_E)\mu_0(dx, de)
\end{aligned}$$

The first equality is given by the standard definition of expectations along with the Tower Property.[1] In the second equality, we 'pull out' the term $\mathbb{1}_{B_A}(x_1^i)$ from the conditional expectation since x_1^i is $\mathscr{F}_1^i$ measurable.[2] The third equality follows from the definition of conditional probability[3] and since x_0^i is independent of $\sigma(e_0^i, e_1^i)$ by 3. of Assumption 2.1.[4] The fourth line follows from properties of the Markov kernel[5] and because μ_0 is the marginal distribution of $\{x_0^i, e_0^i\}$. Thus we have shown μ_1 and μ_0 satisfy the recursion (2) and hence the claim holds for $t = 0$.

Now make the inductive assumption that the claim holds for arbitrary t, that is, $\{x_{t+1}^i, e_{t+1}^i\} \sim \mu_{t+1}$ and $\{x_t^i, e_t^i\} \sim \mu_t$ where μ_t and μ_{t+1} satisfy the recursion (2) for each i. Let $\mathscr{F}_{t+2}^i$ be the σ-algebra

[1] See Williams (1991) 9.7 i) and note the trivial σ-algebra is a subset of $\mathscr{F}_1^i$.
[2] See Williams (1991), 9.7 j).
[3] See Williams (1991), 9.9.
[4] See Williams (1991), 9.7 k).
[5] See Equation (5.8), IV in Çinlar (2011).

generated by $\{x_0^i, e_0^i, e_1^i \ldots, e_{t+1}^i\}$. To see the claim holds for $t+1$,

$$
\begin{aligned}
\mu_{t+2}(B_A \times B_E) :&= \bar{\mathbb{P}}\left\{x_{t+2}^i \in B_A, e_{t+2}^i \in B_E\right\} \\
&= \mathbb{E}\left\{\mathbb{E}\left\{\mathbb{1}_{B_A}(x_{t+2}^i) \times \mathbb{1}_{B_E}(e_{t+2}^i) \mid \mathscr{F}_{t+2}^i\right\}\right\} \\
&= \mathbb{E}\left\{\mathbb{1}_{B_A}(x_{t+2}^i) \times \mathbb{E}\left\{\mathbb{1}_{B_E}(e_{t+2}^i) \mid \mathscr{F}_{t+2}^i\right\}\right\} \\
&= \mathbb{E}\left\{\mathbb{1}_{B_A}(x_{t+2}^i) \times \bar{\mathbb{P}}\left\{e_{t+2}^i \in B_E | e_{t+1}^i\right\}\right\} \\
&= \int\int \mathbb{1}_{B_A}\left\{h_{t+1}(x,e)\right\} Q(e, B_E)\mu_{t+1}(dx, de)
\end{aligned}
$$

where each step is analogous to the steps described above. We have thus shown the claim holds for $t+1$. As such, conclude the claim will hold for each t by the Principle of Induction.

Q.E.D.

CLAIM D.2 *If Assumption 2.1 holds and $r(\mu_t) < \infty$ for each t, then x_t^i has finite variance for each t.*

PROOF: We use a proof by induction. First we confirm that if x_t^i has finite variance for some t, then x_{t+1}^i will have finite variance. Since x_{t+1}^i satisfies Equation (7) in the main paper, the following holds

$$
\int (x_{t+1}^i)^2 \, \mathrm{d}\bar{\mathbb{P}} \leq \int \left[(1 + r(\mu_t))x_t^i + w(\mu_t)e_t^i\right]^2 \, \mathrm{d}\bar{\mathbb{P}} < \infty
$$

Since x_0^i has finite variance by assumption, x_1^i will have finite variance. Moreover, if x_t^i for any $t \in \mathbb{N}$ has finite variance, then x_{t+1}^i will have finite variance. By the principle of induction x_t^i will have finite variance for all $t \in \mathbb{N}$.

Q.E.D.

Proofs for Section 2.4

For the next lemma, consider the setting of section 2.3 of the main paper.

LEMMA D.1 *Fix any $t \in \mathbb{N}$. If $(x_t)_{t=0}^{\infty}$ and $(y_t)_{t=0}^{\infty}$ are random variables adapted to $(\mathscr{F}_t)_{t=0}^{\infty}$, then for any $j \geq t$,*

$$
\mathbb{E}(y_{j+1} | \sigma(x_t, e_t, \ldots, e_{j+1})) = \mathbb{E}(y_{j+1} | \sigma(x_t, e_t, \ldots, e_j))
$$

PROOF: Observe $\mathbb{E}(y_{j+1} | \sigma(x_t, e_t, \ldots, e_j))$ will be $\sigma(x_t, e_t, \ldots, e_{j+1})$ measurable since

$$
\sigma(x_t, e_t, \ldots, e_j)) \subset \sigma(x_t, e_t, \ldots, e_{j+1})
$$

Thus, we will prove the lemma by using the definition of conditional expectations at Section C.3 of the online appendix, and show

$$
\int_B \mathbb{E}(y_{j+1} | \sigma(x_t, e_t, \ldots, e_j)) \, \mathrm{d}\mathbb{P} = \int_B y_{j+1} \, \mathrm{d}\mathbb{P} \tag{D.2}
$$

for all $B \in \sigma(x_t, e_t, \ldots, e_{j+1})$.

Begin by verifying

$$\mathbb{E}(y_{j+1}|\sigma(x_t, e_t, \ldots, e_j)) = \mathbb{E}(y_{j+1}|\sigma(x_t, e_t, \ldots, e_j, \eta_{j+1})) \tag{D.3}$$

By construction of the Markov process, $e_{i+1} = G(e_i, \eta_{i+1})$ for each $i \geq 1$, where $\sigma(\eta_{i+1})$ is independent of $\sigma(x_0, e_0, \eta_1, \ldots, \eta_i)$ and $G \colon E \times \Omega \to E$ is measurable. As such, each e_i is a function of the shocks e_0 and $\eta_1, \ldots, \eta_i$; applying the Doob-Dynkin Lemma (Fact C.3), we have

$$\sigma(x_t, e_t, \ldots, e_j) \subset \sigma(x_0, e_0, \eta_1, \ldots, \eta_j)$$

It follows that since $\sigma(\eta_{j+1})$ and $\sigma(x_0, e_0, \eta_1, \ldots, \eta_j)$ are independent, $\sigma(\eta_{j+1})$ and $\sigma(x_t, e_t, \ldots, e_j)$ will also be independent.

Now, use Fact C.2 in the online appendix to write

$$\sigma(x_t) \cup \sigma(x_t, e_t, \ldots, e_j) \subset \sigma(x_t, e_t, \ldots, e_j)$$

as such,

$$\sigma(\sigma(x_t), \sigma(x_t, e_t, \ldots, e_j)) \subset \sigma(x_t, e_t, \ldots, e_j)$$

Thus $\sigma(\eta_{j+1})$ and $\sigma(\sigma(x_t), \sigma(x_t, e_t, \ldots, e_j))$ will be independent, since we showed above that $\sigma(\eta_{j+1})$ and $\sigma(x_t, e_t, \ldots, e_j)$ are independent. By Fact C.1 in the online appendix, Equation (D.3) follows.

Next, by the Doob-Dynkin Lemma, $\sigma(e_{j+1}) \subset \sigma(e_j, \eta_{j+1})$. As such, we can write the following inclusions

$$\begin{aligned}
\sigma(x_t, e_t, \ldots, e_{j+1}) &\subset \sigma\left(\sigma(x_t) \cup \sigma(e_t) \cup \cdots \cup \sigma(e_j) \cup \sigma(e_j, \eta_{j+1})\right) \\
&\subset \sigma\left(\sigma(x_t) \cup \sigma(e_t) \cup \cdots \cup \sigma(e_j) \cup \sigma(e_j) \cup \sigma(\eta_{j+1})\right) \\
&= \sigma\left(\sigma(x_t) \cup \sigma(e_t) \cup \cdots \cup \sigma(e_j) \cup \sigma(\eta_{j+1})\right)
\end{aligned} \tag{D.4}$$

where the second inclusion follows from Fact C.2.

To complete the proof by showing (D.2), let B satisfy $B \in \sigma(x_t, e_t, \ldots, e_{j+1})$. Recall Equation (D.3) and write

$$\begin{aligned}
&\int_B \mathbb{E}(y_{j+1}|\sigma(x_t, e_t, \ldots, e_j))\, \mathrm{d}\mathbb{P} \\
&\qquad = \int_B \mathbb{E}(y_{j+1}|\sigma(x_t, e_t, \ldots, e_j, \eta_{j+1})\, \mathrm{d}\mathbb{P} = \int_B y_{j+1}\, \mathrm{d}\mathbb{P}
\end{aligned}$$

where the final equality comes from the definition of conditional expectation and since, by (D.4), B will satisfy $B \in \sigma(x_t, e_t, \ldots, e_j, \eta_{j+1})$.

Q.E.D.

CLAIM D.3 *Let $(y_t)_{t=0}^{\infty}$ be a solution to the sequential problem. If $(x_t)_{t=0}^{\infty}$ is a sequence of random variables defined by* (20) *in the main text, then*

$$x_t \colon = \mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1})) = \mathbb{E}(y_t|\sigma(x_t, e_t)), \qquad \forall t \in \mathbb{N}$$

PROOF: We will prove the claim by using the definition of conditional expectation from Section C.3 showing $\mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1}))$ is $\sigma(x_t, e_t)$ measurable and satisfies

$$\int_B \mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1}))\, \mathrm{d}\mathbb{P} = \int_B y_t\, \mathrm{d}\mathbb{P}$$

for $B \in \sigma(x_t, e_t)$.

To show $\mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1}))$ is $\sigma(x_t, e_t)$ measurable, observe $\mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1}))$ can be written as a function of $\{x_t, e_t\}$ as follows:

$$\{x_t, e_t\} = \{\mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1})), e_t\} \mapsto \mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1}))$$

and thus measurability follows from the Doob-Dynkin Lemma (Fact C.3).

Next, by Lemma D.1, we have

$$\mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1})) = \mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1}, e_t))$$

Moreover, $\sigma(x_t, e_t) \subset \sigma(x_{t-1}, e_{t-1}, e_t)$ by the Doob-Dynkin Lemma since x_t is $\sigma(x_{t-1}, e_{t-1})$ measurable by definition of x_t. Now take any B satisfying $B \in \sigma(x_t, e_t)$. Since $B \in \sigma(x_{t-1}, e_{t-1}, e_t)$, we can write

$$\int_B \mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1}))\, \mathrm{d}\mathbb{P} = \int_B \mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1}, e_t))\, \mathrm{d}\mathbb{P} = \int_B y_t\, \mathrm{d}\mathbb{P}$$

as was to be shown to prove the claim.

Q.E.D.

PROOF OF PROPOSITION 2.1: We show the sequence $(x_t)_{t=0}^{\infty}$ is feasible and achieves the sequential planner's value function.

Before proceeding, note the following holds due to the Tower Property of conditional expectations.

$$\int y_t\, \mathrm{d}\mathbb{P} = \int \mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1}))\, \mathrm{d}\mathbb{P} = \int x_t\, \mathrm{d}\mathbb{P}, \qquad t \in \mathbb{N} \tag{D.5}$$

To show feasibility of $(x_t)_{t=0}^{\infty}$, we verify $x_{t+1} \in \Gamma_t(x_t)$ for each t, where Γ_t is defined by (16) in the main text. First, $x_{t+1} = h_t(x_t, e_t)$ for a measurable function h_t for each t with x_0 given. Thus x_t can be written as a measurable function of $x_0, \ldots, e_{t-1}$, implying $x_t \in m\mathscr{F}_t$ for each t. Moreover, by (D.5), $\int x_t\, \mathrm{d}\mathbb{P} = \int y_t\, \mathrm{d}\mathbb{P}$ and since $\int y_t\, \mathrm{d}\mathbb{P} \in [0, \bar{K}]$ for each t, we have $\int x_t\, \mathrm{d}\mathbb{P} \in [0, \bar{K}]$. And by positivity of conditional expectation,[6] since $y_t \geq 0$, $x_t = \mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1})) \geq 0$.

Next, we check x_{t+1} satisfies the budget constraints in the definition of $\Gamma_t(x_t)$, given by (16) in the main text, for each t. Set any $t \in \mathbb{N}$. There are two cases to consider: first $\int x_t\, \mathrm{d}\mathbb{P} > 0$ and second $\int x_t\, \mathrm{d}\mathbb{P} = 0$. Suppose $\int x_t\, \mathrm{d}\mathbb{P} > 0$. By (D.5), we have $\int y_t\, \mathrm{d}\mathbb{P} = \int x_t\, \mathrm{d}\mathbb{P} > 0$. Since $(y_t)_{t=0}^{\infty}$

[6] See 9.7 d) by Williams (1991).

is a solution to the sequential planner's problem, we have $y_{t+1} \in \Gamma_t(y_t)$ and thus, $y_{t+1} \leq (1 + \tilde{r}(y_t))y_t + \tilde{w}(y_t)e_t$. To show $x_{t+1} \leq (1 + \tilde{r}(x_t))x_t + \tilde{w}(x_t)e_t$, consider,

$$\begin{aligned}
x_{t+1} &= \mathbb{E}(y_{t+1}|\sigma(x_t, e_t)) \\
&\leq \mathbb{E}((1 + \tilde{r}(y_t))y_t + \tilde{w}(y_t)e_t|\sigma(x_t, e_t)) \\
&= (1 + \tilde{r}(x_t))\mathbb{E}(y_t|\sigma(x_t, e_t)) + \tilde{w}(x_t)e_t \\
&= (1 + \tilde{r}(x_t))\mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1})) + \tilde{w}(x_t)e_t \\
&= (1 + \tilde{r}(x_t))x_t + \tilde{w}(x_t)e_t
\end{aligned}$$

where, noting (D.5), the third line follows from

$$\tilde{r}(y_t) = F_1\left(\int y_t \, d\mathbb{P}, L\right) = F_1\left(\int x_t \, d\mathbb{P}, L\right) = \tilde{r}(x_t)$$

A similar argument shows $\tilde{w}(y_t) = \tilde{w}(x_t)$. The fourth line follows from Claim D.3 and the final line follows from the definition of x_t by (20) in the main text.

On the other hand, suppose $\int x_t \, d\mathbb{P} = 0$. We have $\int y_t = \int x_t \, d\mathbb{P} = 0$ by (D.5). As such, since $y_{t+1} \in \Gamma_t(y_t)$ and noting the definition of Γ_t at (16) in the main text, $\int x_{t+1} \, d\mathbb{P} = \int y_{t+1} \, d\mathbb{P} = 0$. Since $x_{t+1} \geq 0$, x_{t+1} satisfies $x_{t+1} = 0$.[7]

To re-cap, for each t, x_t is $\mathscr{F}_t$ measurable and satisfies $\int x_t \leq \bar{K}$ and $x_t \geq 0$. Hence $x_t \in \mathrm{S}_t$. Moreover, x_{t+1} satisfies the budget constraints in the definition of Γ_t for each t. Thus $x_{t+1} \in \Gamma_t(x_t)$ for each t.

Next, we check $\rho_t(x_t, x_{t+1}) \geq \rho_t(y_t, y_{t+1})$ for each t. Select any t and consider the case $\int x_t \, d\mathbb{P} > 0$. We have

$$\begin{aligned}
\rho_t(x_t, x_{t+1}) &= \int v\left((1 + \tilde{r}(x_t))x_t + \tilde{w}(x_t)e_t - x_{t+1})\right) d\mathbb{P} \\
&= \int v\big((1 + \tilde{r}(y_t))\mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1})) + \tilde{w}(y_t)e_t \\
&\quad - \mathbb{E}(y_{t+1}|\sigma(x_t, e_t))\big) \, d\mathbb{P} \\
&= \int v\left(\mathbb{E}\left[(1 + \tilde{r}(y_t))y_t + \tilde{w}(y_t)e_t - y_{t+1}|\sigma(x_t, e_t)\right]\right) d\mathbb{P} \\
&\geq \int \mathbb{E}\left(v((1 + \tilde{r}(y_t))y_t + \tilde{w}(y_t)e_t - y_{t+1})|\sigma(x_t, e_t)\right) d\mathbb{P} \\
&= \rho_t(y_t, y_{t+1})
\end{aligned}$$

where the second line is due to the definition x_t and x_{t+1}. The third line follows from Claim D.3, the fourth line follows from Jensen's inequality (Fact C.4 in the Mathematical Preliminaries) and the final line is due to the Tower Property.

If $\int x_t \, d\mathbb{P} = 0$, then $\rho_t(x_t, x_{t+1}) = 0$ by definition of ρ_t. Since $\int y_t \, d\mathbb{P} = 0$ by (D.5), $\rho_t(y_t, y_{t+1}) = 0$. Conclude

$$J(x_0) = \sum_{t=0}^{\infty} \beta^t \rho_t(y_t, y_{t+1}) \leq \sum_{t=0}^{\infty} \beta^t \rho_t(x_t, x_{t+1}) \tag{D.6}$$

[7] See Theorem 1.1.20 by Tao (2010).

Since $J(x_0)$ achieved the supremum of all pay-offs from feasible sequences and $(x_t)_{t=0}^{\infty}$ satisfies $x_{t+1} \in \Gamma_t(x_t)$ for each t, $J(x_0) = \sum_{t=0}^{\infty} \beta^t \rho_t(x_t, x_{t+1})$, allowing us to conclude $(x_t)_{t=0}^{\infty}$ is a solution to the sequential problem.

Q.E.D.

PROOF OF THEOREM 2.2: Consider the setting of sections 2.1 - 2.3 and note μ_0 is the initial state of the economy for the recursive planner. Let $(y_t)_{t=0}^{\infty}$ be a solution to the sequential problem (Definition 2.2 in the main paper). Since $y_0 = x_0$ and $\{x_0, e_0, e_1, \dots\}$ have a joint distribution P, $\{y_0, e_0\} \sim \mu_0$.

Construct $(x_t)_{t=0}^{\infty}$ according to (20) in the main paper; recall $(x_t)_{t=0}^{\infty}$ satisfies $x_{t+1} = h_t(x_t, e_t)$ for a sequence of measurable functions $(h_t)_{t=0}^{\infty}$ with $h_t \colon S \to A$ for each t. Define a sequence of Borel probability measures $(\mu_t)_{t=0}^{\infty}$ using (21) in the main paper and note by definition that μ_t will the distribution of $\{x_t, e_t\}$ for each t.

We show $(\mu_t, h_t)_{t=0}^{\infty}$ solves the recursive problem in two steps. First we verify $(\mu_t, h_t)_{t=0}^{\infty}$ is feasible for the recursive problem (part 1) and then verify the sum of discounted pay-offs from $(\mu_t, h_t)_{t=0}^{\infty}$ dominates the sum of discounted pay-offs from any other feasible sequence of distributions and policy functions (part 2).

Part 1: Show $(\mu_t, h_t)_{t=0}^{\infty}$ satisfies feasibility for the recursive problem

This part shows $(\mu_t, h_t)_{t=0}^{\infty}$ satisfies (12) in the main paper, that is, $h_t \in \Lambda(\mu_t)$ and $\mu_{t+1} = \Phi(\mu_t, h_t)$ for each t.

Fix any $t \in \mathbb{N}$. To confirm $h_t \in \Lambda(\mu_t)$, we consider two cases: when $\int\int x\mu_t(dx, de) > 0$ and when $\int\int x\mu_t(dx, de) = 0$. First suppose $\int\int x\mu_t(dx, de) > 0$, we show

$$\mu_t\{a, e \in S \,|\, h_t(a, e) \notin [0, (1 + r(\mu_t))a + w(\mu_t)e]\} = 0$$

The condition says the policy function h_t satisfies agents' budget constraints μ_t - almost everywhere. Using the definition of μ_t by Equation (21) of the main paper,

$$\begin{aligned}
&\mu_t\{a, e \in S \,|\, h_t(a, e) \notin [0, (1 + r(\mu_t))a + w(\mu_t)e]\} \\
&= \mathbb{P}\{\omega \in \Omega \,|\, h_t(x_t(\omega), e_t(\omega)) \\
&\quad \notin [0, (1 + \tilde{r}(x_t))x_t(\omega) + \tilde{w}(x_t)e_t(\omega)]\} \\
&= \mathbb{P}\{\omega \in \Omega \,|\, x_{t+1}(\omega) \\
&\quad \notin [0, (1 + \tilde{r}(x_t))x_t(\omega) + \tilde{w}(x_t)e_t(\omega)]\} \\
&= 0
\end{aligned} \tag{D.7}$$

The first equality uses the following observation, which holds because μ_t is the joint distribution of $\{x_t, e_t\}$:

$$\int\int x\mu_t(dx, de) = \int x_t \, \mathrm{d}\mathbb{P} > 0 \tag{D.8}$$

whence,

$$r(\mu_t) = F_1\left(\int\int x\mu_t(dx, de), L\right) = F_1\left(\int x_t \, \mathrm{d}\mathbb{P}, L\right) = \tilde{r}(x_t) \tag{D.9}$$

An identical arguments shows $\tilde{w}(x_t) = w(\mu_t)$. The second equality in (D.7) follows from $x_{t+1} = h_t(x_t, e_t)$. The final equality is true because $\int x_t \, \mathrm{d}\mathbb{P} > 0$ and because the sequence $(x_t)_{t=0}^{\infty}$ satisfies $x_{t+1} \in \Gamma_t(x_t)$ for each t. Thus, $0 \leq h_t(a,e) \leq (1 + r(\mu_t))a + w(\mu_t)e$ for μ_t almost everywhere, implying $h_t \in \Lambda(\mu_t)$ if $\int \int x\mu_t(dx, de) > 0$.

Now suppose $\int \int x\mu_t(dx, de) = 0$. Observe

$$\int x \int \mu_t(dx, de) = \int x_t \, \mathrm{d}\mathbb{P} = 0 \tag{D.10}$$

Since $(x_t)_{t=0}^{\infty}$ satisfies $x_{t+1} \in \Gamma_t(x_t)$ for each t, $x_{t+1} = 0$. Whence,

$$\mu_t \{h_t(a, e) \neq 0\} = \mathbb{P}\{h_t(x_t, e_t) \neq 0\} = \mathbb{P}\{x_{t+1} \neq 0\} = 0$$

Thus $h_t = 0$ for μ_t almost everywhere and $h_t \in \Lambda(\mu_t)$ if $\int \int x\mu_t(dx, de) = 0$.

Now we show $\mu_{t+1} = \Phi(\mu_t, h_t)$ for each t. Let $B \in \mathscr{B}(S)$, where $B = B_A \times B_E$ for $B_A \in \mathscr{B}(A)$ and $B_E \in \mathscr{B}(E)$. Use the definition of μ_{t+1} to write

$$\begin{aligned}
\mu_{t+1}(B_A \times B_E) :&= \mathbb{P}\{x_{t+1} \in B_A, e_{t+1} \in B_E\} \\
&= \mathbb{E}\left\{\mathbb{E}\left\{\mathbb{1}_{B_A}(x_{t+1}) \times \mathbb{1}_{B_E}(e_{t+1}) \mid \mathscr{F}_{t+1}\right\}\right\} \\
&= \mathbb{E}\left\{\mathbb{1}_{B_A}(x_{t+1}) \times \mathbb{E}\left\{\mathbb{1}_{B_E}(e_{t+1}) \mid \mathscr{F}_{t+1}\right\}\right\} \\
&= \mathbb{E}\left\{\mathbb{1}_{B_A}(x_{t+1}) \times \mathbb{E}\left\{\mathbb{1}_{B_E}(e_{t+1}) | \sigma(x_0, e_0, \dots, e_t)\right\}\right\} \\
&= \mathbb{E}\left\{\mathbb{1}_{B_A}(x_{t+1}) \times \mathbb{E}\left\{\mathbb{1}_{B_E}(e_{t+1}) | \sigma(e_0, \dots, e_t)\right\}\right\} \\
&= \mathbb{E}\left\{\mathbb{1}_{B_A}(x_{t+1}) \times \mathbb{P}\left\{e_{t+1} \in B_E \,\middle|\, \sigma(e_0, \dots, e_t)\right\}\right\} \\
&= \int \int \mathbb{1}_{B_A}\{h_t(x, e)\} \, Q(e, B_E)\mu_t(dx, de)
\end{aligned}$$

The first equality is given by the standard definition of conditional probability[8] and the Tower Property.[9] In the second equality, we 'pull out' the term $\mathbb{1}_{B_A}(x_{t+1})$ from the conditional expectation since x_{t+1} is $\mathscr{F}_{t+1}$ measurable.[10] The fourth inequality is due to independence of x_0 and $\sigma(e_0, \dots, e_t)$ by 3) of Assumption 2.1.[11] The fifth equality is due to the definition of conditional probability.[12] The final line follows from the definition of a Markov kernel[13] and because μ_t is the distribution of $\{x_t, e_t\}$.

Part 2: Show $(\mu_t, h_t)_{t=0}^{\infty}$ achieves the value function for the recursive problem

We have shown $(\mu_t, h_t)_{t=0}^{\infty}$ satisfies feasibility. Our next task is to show

$$\sum_{t=0}^{\infty} \beta^t u(\mu_t, h_t) \geq \sum_{t=0}^{\infty} \beta^t u(\tilde{\mu}_t, \tilde{h}_t)$$

[8] Williams (1991), 9.9.
[9] Williams (1991), 9.7, i) and note the trivial σ-algebra is a subset of $\mathscr{F}_{t+1}$.
[10] Williams (1991), 9.7, j).
[11] Williams (1991), 9.7, k).
[12] Williams (1991), 9.9.
[13] Equation (5.8), ch. IV by Çinlar (2011).

holds for any other sequence $(\tilde{\mu}_t, \tilde{h}_t)_{t=0}^{\infty}$ feasible for the recursive problem.

Let $(\tilde{\mu}_t, \tilde{h}_t)_{t=0}^{\infty}$ be any other sequence of Borel probability measures on S and measurable policy functions $\tilde{h}_t \colon S \to A$ satisfying $\tilde{\mu}_0 = \mu_0$ and (12) in the main paper. Construct a sequence of A valued random variables $(\tilde{x}_t)_{t=0}^{\infty}$ by letting $\tilde{x}_{t+1} = \tilde{h}_t(\tilde{x}_t, e_t)$ for each $t > 0$ and with $\tilde{x}_0 = x_0$ given. The sequence of random variables $(\tilde{x}_t)_{t=0}^{\infty}$ will be defined on the probability space $(\Omega, \mathscr{F}, \mathbb{P})$. By Claim D.4 below, $\tilde{\mu}_t$ is the distribution of $\{\tilde{x}_t, e_t\}$ for each t. Moreover, using an analogous argument to Claim D.2 above, each $\tilde{x}_t$ will have finite variance and hence $\tilde{x}_t \in L^2(\Omega, \mathbb{P})$ for each t.

Our strategy is to show $(\tilde{x}_t)_{t=0}^{\infty}$ is feasible for the sequential problem and show $u(\tilde{\mu}_t, \tilde{h}_t) = \rho_t(\tilde{x}_t, \tilde{x}_{t+1})$ and $u(\mu_t, h_t) = \rho_t(x_t, x_{t+1})$ for each t. The proof will then be complete since, noting $(x_t)_{t=0}^{\infty}$ is a solution for the sequential problem (Proposition 2.1), $u(\mu_t, h_t) = \rho_t(x_t, x_{t+1}) \geq \rho_t(\tilde{x}_t, \tilde{x}_{t+1}) = u(\tilde{\mu}_t, \tilde{h}_t)$ for each t.

To check $(\tilde{x}_t)_{t=0}^{\infty}$ satisfies $\tilde{x}_{t+1} \in \Gamma_t(\tilde{x}_t)$ for each t, we check the conditions stated at (16) in the main paper for each t. First, we confirm $(\tilde{x}_t)_{t=0}^{\infty}$ is adapted to the filtration $(\mathscr{F}_t)_{t=0}^{\infty}$. Proceed by induction, let $t = 1$ and consider:

$$\tilde{x}_1 = \tilde{h}_1(x_0, e_0)$$

Since $\tilde{h}_1$ is measurable, by the Doob-Dynkin Lemma (see Fact C.3 above), $\tilde{x}_1$ will be $\sigma(x_0, e_0)$ measurable. Now suppose $\tilde{x}_t$ is $\sigma(x_0, e_0, \dots, e_{t-1})$ measurable. Consider

$$\tilde{x}_{t+1} = \tilde{h}_t(\tilde{x}_t, e_t) = \tilde{h}_t(g(x_0, e_0, \dots, e_{t-1}), e_t)$$

for some measurable function $g \colon A \times E^t \to A$. Once again, since $\tilde{h}_t$ is Borel measurable, using the Doob-Dynkin Lemma, $\tilde{x}_{t+1}$ is $\sigma(x_0, e_0, \dots, e_t)$ measurable. By the Principle of Induction, $(\tilde{x}_t)_{t=0}^{\infty}$ is adapted to the filtration $(\mathscr{F}_t)_{t=0}^{\infty}$.

To confirm $\int \tilde{x}_t \, \mathrm{d}\mathbb{P} \in [0, \bar{K}]$ for each t, since $\tilde{\mu}_t \in \mathbb{M}$ and $\tilde{\mu}_t$ is the distribution of $\{\tilde{x}_t, e_t\}$, we have

$$\int \tilde{x}_t \, \mathrm{d}\mathbb{P} = \int \int x \tilde{\mu}_t(dx, de) \in [0, \bar{K}] \tag{D.11}$$

Now we turn to show the sequence $(\tilde{x}_t)_{t=0}^{\infty}$ satisfies the agent budget constraints for each t. Fix any $t \in \mathbb{N}$ and suppose $\int \tilde{x}_t \, \mathrm{d}\mathbb{P} > 0$. We have

$$\begin{aligned}
\mathbb{P}\{\tilde{x}_{t+1} \notin [0, (1 + \tilde{r}(\tilde{x}_t))\tilde{x}_t + \tilde{w}(\tilde{x}_t)e_t]\} &= \mathbb{P}\{\tilde{h}_t(\tilde{x}_t, e_t) \\
&\qquad \notin [0, (1 + \tilde{r}(\tilde{x}_t))\tilde{x}_t + \tilde{w}(\tilde{x}_t)e_t]\} \\
&= \tilde{\mu}_t\{\tilde{h}_t(x, e) \\
&\qquad \notin [0, (1 + r(\tilde{\mu}_t))x + w(\tilde{\mu}_t)e]\} \\
&= 0
\end{aligned}$$

The final equality holds because $\tilde{\mu}_t$ is the distribution of $\{\tilde{x}_t, e_t\}$ and because $\tilde{h}_t$ satisfies the feasibility condition shown at (9) in the main paper for $\tilde{\mu}_t$ - almost everywhere.

On the other hand, suppose $\int \tilde{x}_t \, \mathrm{d}\mathbb{P} = 0$. We have $\int \tilde{x}_t \, \mathrm{d}\mathbb{P} = \int \int x \tilde{\mu}_t(dx, de) = 0$. Since $(\tilde{\mu}_t, \tilde{h}_t)_{t=0}^{\infty}$ satisfies $\tilde{h}_t \in \Lambda(\tilde{\mu}_t)$ for each t, $\tilde{h}_t(x, e) = 0$ for $\tilde{\mu}_t$ almost everywhere. Whence,

$$\mathbb{P}\left\{\tilde{x}_{t+1} \neq 0\right\} = \mathbb{P}\left\{\tilde{h}_t(x_t, e_t) \neq 0\right\} = \tilde{\mu}_t\left\{\tilde{h}_t(x, e) \neq 0\right\} = 0$$

The first equality holds because we defined $\tilde{x}_{t+1} = \tilde{h}_t(\tilde{x}_t, e_t)$. The second inequality holds because $\tilde{\mu}_t$ is the joint distribution of $\{\tilde{x}_t, e_t\}$. As such, for each t, $\tilde{x}_{t+1}$ satisfies all the conditions stated in the definition of the feasibility correspondence, (16) in the main paper, for $\tilde{x}_{t+1} \in \Gamma_t(\tilde{x}_t)$.

To complete the proof, for each t,

$$\begin{aligned} u(\mu_t, h_t) &= \int \nu((1+r(\mu_t))x + w(\mu_t)e - h_t(x,e))\mu_t(dx,de) \\ &= \int \nu((1+\tilde{r}(x_t))x_t + \tilde{w}(x_t)e_t - h_t(x_t,e_t))\,\mathrm{d}\mathbb{P} \\ &= \int \nu((1+\tilde{r}(x_t))x_t + \tilde{w}(x_t)e_t - x_{t+1})\,\mathrm{d}\mathbb{P} \\ &= \rho_t(x_t, x_{t+1}) \end{aligned} \tag{D.12}$$

And similarly, $u(\tilde{\mu}_t, \tilde{h}_t) = \rho_t(\tilde{x}_t, \tilde{x}_{t+1})$ for each t. As such, conclude

$$\sum_{t=0}^{\infty} \beta^t u(\mu_t, h_t) = \sum_{t=0}^{\infty} \beta^t \rho_t(x_t, x_{t+1}) \geq \sum_{t=0}^{\infty} \beta^t \rho_t(\tilde{x}_t, \tilde{x}_{t+1}) = \sum_{t=0}^{\infty} \beta^t u(\tilde{\mu}_t, \tilde{h}_t) \tag{D.13}$$

where the inequality follows since $(x_t)_{t=0}^{\infty}$ is a solution to the sequential problem and its discounted sum of pay-offs dominate the discounted sum of pay-offs from $(\tilde{x}_t)_{t=0}^{\infty}$.

Finally, since any arbitrary feasible sequence $(\tilde{\mu}_t, \tilde{h}_t)_{t=0}^{\infty}$, with $\tilde{\mu}_0 = \mu_0$, satisfies (D.13), we have $V(\mu_0) = \sum_{t=0}^{\infty} \beta^t u(\mu_t, h_t)$. Moreover, since $(x_t)_{t=0}^{\infty}$ solves the sequential planner's problem, the first equality of (D.13) implies $V(\mu_0) = J(x_0)$.

Q.E.D.

For the following claim, consider the setting and notation in the proof for Theorem (2.2), part 2.

CLAIM D.4 *If $\tilde{x}_{t+1} = \tilde{h}_t(\tilde{x}_t, e_t)$ for each $t > 0$ and $\tilde{x}_0 = x_0$, then $\{\tilde{x}_t, e_t\} \sim \tilde{\mu}_t$ for each $t \in \mathbb{N}$.*

PROOF: We use a proof by induction. Let $t = 0$, since $\tilde{x}_0 = x_0$ and $\tilde{\mu}_0 = \mu_0$, by the construction of $\{x_0, e_0, e_1, \dots\}$ in section 2.3, we have $\{\tilde{x}_0, e_0\} \sim \tilde{\mu}_0$. Now make the inductive assumption $\{\tilde{x}_t, e_t\} \sim \tilde{\mu}_t$. We show $\{\tilde{x}_{t+1}, e_{t+1}\} \sim \tilde{\mu}_{t+1}$. Let $B_A \times B_E \in \mathscr{B}(S)$ and observe

$$\begin{aligned} \mathbb{P}\{\tilde{x}_{t+1} \in B_A, e_{t+1} \in B_E\} &= \mathbb{E}\{\mathbb{1}_{B_A}(\tilde{x}_{t+1}) \times \mathbb{1}_{B_E}(e_{t+1})\} \\ &= \mathbb{E}\{\mathbb{E}\{\mathbb{1}_{B_A}(\tilde{x}_{t+1}) \times \mathbb{1}_{B_E}(e_{t+1})|\mathscr{F}_{t+1}\}\} \\ &= \mathbb{E}\{\mathbb{1}_{B_A}(\tilde{x}_{t+1}) \times \mathbb{E}\{\mathbb{1}_{B_E}(e_{t+1})|\mathscr{F}_{t+1}\}\} \\ &= \mathbb{E}\{\mathbb{1}_{B_A}(\tilde{x}_{t+1}) \times \mathbb{E}\{\mathbb{1}_{B_E}(e_{t+1})|e_t\}\} \\ &= \mathbb{E}\{\mathbb{1}_{B_A}(\tilde{h}_t(\tilde{x}_t, e_t)) \times Q(e_t, B_E)\} \\ &= \int\int \mathbb{1}_{B_A}(\tilde{h}_t(x,e)) \times Q(e, B_E)\tilde{\mu}_t(dx,de) \\ &= \Phi(\tilde{\mu}_t, \tilde{h}_t)(B_A, B_E) = \tilde{\mu}_{t+1}(B_A \times B_E) \end{aligned}$$

where the first line follows from the standard definition of expectations. The second equality follows from the Tower property of conditional expectations.[14] The third line is true because $\mathbb{1}_{B_A}(\tilde{x}_{t+1})$

[14] See Williams (1991) 9.7 i) and note the trivial σ-algebra is a subset of $\mathscr{F}_{t+1}$.

is $\mathscr{F}_{t+1}$ measurable.[15] The fourth line follows from properties of the Markov kernel[16] and the fact that x_0 is independent of $\{e_0, \ldots, e_t\}$.[17] The fifth line follows from our inductive assumption $\{\tilde{x}_t, e_t\} \sim \mu_t$ and the final line is from the definition the operator Φ given by Equation (10) of the main paper. Thus for any $B_A \times B_E \in \mathscr{B}(S)$, we have $\mathbb{P}\{\tilde{x}_{t+1} \in B_A, e_{t+1} \in B_E\} = \tilde{\mu}_{t+1}(B_A \times B_E)$, implying $\{\tilde{x}_{t+1} \times e_{t+1}\} \sim \tilde{\mu}_{t+1}$.

To conclude, since $\{\tilde{x}_0, e_0\} \sim \mu_0$ and $\{\tilde{x}_{t+1}, e_{t+1}\} \sim \tilde{\mu}_{t+1}$ if $\{\tilde{x}_t, e_t\} \sim \tilde{\mu}_t$, by the Principle of Induction, $\{\tilde{x}_t, e_t\} \sim \tilde{\mu}_t$ for each $t \in \mathbb{N}$.

Q.E.D.

Proofs for section 4

PROOF OF PROPOSITION 4.1: Let x satisfy $x \in \mathbb{S}_0$. By Assumption 2.4, aggregate capital is bounded above by $\bar{K}$. As such, for any $(x_t)_{t=0}^{\infty} \in \mathcal{G}(x)$ and for any t, we can use Jensen's inequality (fact C.4 in the Mathematical Preliminaries) to arrive at

$$\rho_t(x_t, x_{t+1}) \leq \nu\left((1 + \tilde{r}(x_t))\mathbb{E}x_t + \tilde{w}(x_t)L\right) \leq \nu(F(\bar{K}, L) + (1-\delta)\bar{K})$$

where the second inequality follows from homogeneity of degree one of the production function.

Let $m_t \colon = \nu(F(\bar{K}, L) + (1-\delta)\bar{K})$ for all t. As such, for any $(x_t)_{t=0}^{\infty} \in \mathcal{G}(x)$,

$$\rho_t(x_t, x_{t+1}) \leq m_t$$

for all t. Since m_t is a constant, the sequence $(m_t)_{t=0}^{\infty}$ will satisfy $\sum_{t=0}^{\infty} \beta^t m_t < \infty$.

Q.E.D.

PROOF OF CLAIM A.1: The functional G will be upper semicontinuous with respect to norm convergence in $L^1(\Omega, \mu)$ and concave (Proposition 6.3.1 by Borwein and Vanderwerff (2010)). Since, by Holder's inequality, norm convergence in $L^2(\Omega, \mu)$ implies norm convergence in $L^1(\Omega, \mu)$, G will be upper semicontinuous with respect to norm convergence in $L^2(\Omega, \mu)$. As such, the upper-contour sets of G will be convex (Fact 2.1.7, Borwein and Vanderwerff (2010)) and norm-closed. Since norm closed convex sets in $L^2(\Omega, \mu)$ are also weak sequentially closed (see Theorem 2.5.16 and discussion by Megginson (1998) or Theorem 8.13 by Alt and Nürnberg (2016)), the upper contour sets of G will be weak sequentially closed. By the definition of upper semi-continuity in the online appendix, we conclude G will be weak sequentially upper semicontinuous.

Q.E.D.

PROOF OF PROPOSITION 4.3: Recall the definition of $\mathbb{S}_t$:

$$\mathbb{S}_t \colon = \left\{ x \in m\mathscr{F}_t \,\middle|\, 0 \leq x \,, \int x \, \mathrm{d}\mathbb{P} \leq \bar{K} \right\}$$

Consider a sequence $(x^n)_{n=0}^{\infty}$ with $x^n \in \mathbb{S}_t$ for each n and $x^n \to x^\star$ with $x^\star \in L^2(\Omega, \mathbb{P})$. We show $x^\star \in \mathbb{S}_t$.

[15]See Williams (1991), 9.9.
[16]See section 5 in Çinlar (2011).
[17]See Williams (1991) 9.7 k).

Let $L^2(\Omega, \mathscr{F}_t, \mathbb{P})$ denote the space of $\mathscr{F}_t$ measurable functions on Ω square integrable with respect to $\mathbb{P}$. Observe $x^n \in m\mathscr{F}_t$, thus $x^n \in L^2(\Omega, \mathscr{F}_t, \mathbb{P})$ for each n. Since $L^2(\Omega, \mathscr{F}_t, \mathbb{P}) \subset L^2(\Omega, \mathbb{P})$ and $L^2(\Omega, \mathscr{F}_t, \mathbb{P})$ is a Banach space (Proposition 1.3.7 by Tao (2010)), $L^2(\Omega, \mathscr{F}_t, \mathbb{P})$ will be a norm closed sub-space of $L^2(\Omega, \mathbb{P})$. Moreover, $L^2(\Omega, \mathscr{F}_t, \mathbb{P})$ is convex, thus by Mazur's Lemma, $L^2(\Omega, \mathscr{F}_t, \mathbb{P})$ will be a weak sequentially closed sub-space of $L^2(\Omega, \mathbb{P})$ (see Theorem 2.5.16 and discussion by Megginson (1998) or Theorem 8.13 by Alt and Nürnberg (2016)). As such, $x^\star \in m\mathscr{F}_t$.

Next, the space $\{x \in L^2(\Omega, \mathbb{P}) \,|\, x \geq 0\}$ is norm closed and convex, thus $\{x \in L^2(\Omega, \mathbb{P}) \,|\, x \geq 0\}$ will be weak sequentially closed, implying $x^\star \geq 0$. To conclude, suppose by contradiction $x^\star \notin \mathbb{S}_t$, then $\int x^\star \,\mathrm{d}\mathbb{P} > \bar{K}$. However, since x^n converges weakly, $\int x^n \,\mathrm{d}\mathbb{P} \to \int x^\star \,\mathrm{d}\mathbb{P} > \bar{K}$, yielding a contradiction since $x^n \notin \mathbb{S}_t$ for some n. We conclude $x^\star \in \mathbb{S}_t$ and $\mathbb{S}_t$ is weak sequentially closed.

Q.E.D.

For the following claim, consider the setting of Lemma A.4 in the main paper.

CLAIM D.5 *If $(x_i)_{i=0}^\infty \in \mathcal{G}(x)$, then for each $i > 0$,*

(D.14) $$\mathbb{E}x_i \leq \hat{F}^{i-j}(\mathbb{E}x_j), \qquad \forall j \leq i$$

PROOF: Let $(x_i)_{i=0}^\infty \in \mathcal{G}(x)$ and recall (A.40) from the main text:

(D.15) $$\mathbb{E}x_j \leq \hat{F}(\mathbb{E}x_{j-1}), \qquad j > 0$$

Recall $\hat{F}$ is a strictly increasing function.

We proceed by induction, if $i = 1$, then for $j = 0$,

$$\mathbb{E}x_1 \leq \hat{F}(\mathbb{E}x_0) = \hat{F}^{i-j}(\mathbb{E}x_0)$$

by (D.15).

Next, make the inductive assumption and suppose the claim holds for arbitrary i. We show the claim will for for $i+1$. The inequality (D.14) trivially holds for $j = i+1$. For any $j \leq i$,

$$\begin{aligned}\mathbb{E}x_{i+1} &\leq \hat{F}(\mathbb{E}x_i)\\ &\leq \hat{F}(\hat{F}^{i-j}(\mathbb{E}x_j))\\ &= \hat{F}^{i+1-j}(\mathbb{E}x_j)\end{aligned}$$

where the first inequality is due to (D.15). The second inequality follows from the inductive assumption, because (D.14) holds for i. Thus the claim holds for $i+1$, and by the principle of induction, the claim holds for each $i > 0$.

Q.E.D.

Proofs for section 5

PROOF OF CLAIM 5.1: Define the set C, with $C \subset \mathbb{S}_1$, as follows:

$$C\colon = \{x \in \mathbb{S}_1 \,|\, 0 \leq x \leq (1 + \tilde{r}(x_0))x_0 + \tilde{w}(x_0)e_0\} = \Gamma_0(x_0)$$

The set C will be weakly closed since by Proposition 4.4 in the main paper, Γ_0 has a closed graph, hence $\Gamma_0(x_0)$ is weakly closed. To see C is also weakly compact, note any $x \in C$ satisfies

$$\|x\| \leq \|(1+\tilde{r}(x_0))x_0 + \tilde{w}(x_0)e_0\| :\, = \bar{M}$$

where $\bar{M} < \infty$. Since x_0 and e_0 both have finite norm and $\tilde{r}(x_0)$ is finite, any $x \in C$ will satisfy $\|x\| \leq \bar{M}$. By Alaoglu's Theorem (see Example 6.3 in Mas-colell and Zame (1991)), C will be weakly compact since it is a weakly closed subset of a closed ball in $L^2(, \mathbb{P})$.

Consider the set

$$\Gamma_1(C) :\, = \cup_{x \in C} \Gamma_1(x) = \{y \mid y \in \Gamma_1(x), x \in \Gamma_0(x_0)\}$$

We will construct a norm unbounded sequence in $\Gamma_1(C)$. First define a sequence $(x_n)_{n=0}^{\infty}$ as

$$x_n(x_0, e_0) :\, = x_0^n(1 - x_0^n)$$

Note $x_n \in C$ for each n since $x_0^n(1-x_0^n) \leq 1 \leq w(x_0)e_0$. Next, define

$$y^n :\, = \frac{1}{2}(1+\tilde{r}(x_n))x_n$$

Since $y_n \leq (1+\tilde{r}(x_n))x_n$ and $y_n \geq 0$, we have $y_n \in \Gamma_1(x_n)$, thus $y_n \in \Gamma_1(C)$.

Recall $F_1(K, L) = K^{\alpha-1}$ and use the definition of L^2 norm to write

$$\begin{aligned}
\|y_n\| &= \left(\int (y^n)^2 \,\mathrm{d}\mathbb{P}\right)^{\frac{1}{2}} = \frac{1}{2}\left(\int [(1+\tilde{r}(x^n))x^n]^2 \,\mathrm{d}\mathbb{P}\right)^{\frac{1}{2}} \\
&= \frac{1}{2}(\mathbb{E}(x_n)^{\alpha-1} + 1 - \delta)\left(\int x_0^{2n}(1-x_0^n)^2 \,\mathrm{d}\mathbb{P}\right)^{\frac{1}{2}} \\
&= \frac{1}{2}\left(\left(\int x_0^n(1-x_0^n)\,\mathrm{d}\mathbb{P}\right)^{\alpha-1} + 1 - \delta\right)\left(\int x_0^{2n}(1-x_0^n)^2 \,\mathrm{d}\mathbb{P}\right)^{\frac{1}{2}} \\
&= \frac{1}{2}\left(\left(\int_0^1 a^n(1-a^n)\,\mathrm{d}a\right)^{\alpha-1} + 1 - \delta\right)\left(\int_0^1 a^{2n}(1-a^n)^2 \,\mathrm{d}a\right)^{\frac{1}{2}} \\
&= \frac{1}{2}\left(\left(\frac{n}{1+3n+2n^2}\right)^{\alpha-1} + 1 - \delta\right)\left(\frac{2n^2}{1+9n+26n^2+24n^3}\right)^{\frac{1}{2}} \\
&= \frac{1}{2}\frac{n\sqrt{2}\left[1-\delta+n^{\alpha-1}((1+n)(1+2n))^{1-\alpha}\right]}{\sqrt{(1+2n)(1+3n)(1+4n)}}
\end{aligned}$$

The first and second lines follow from the definition of interest rates and aggregate capital given by Equation (14) in the main paper, along with the definition of y_n. The fourth line follows from the assumption that x_0 is uniformly distributed on the interval $[0, 1]$. The fifth and sixth line are from algebra, solving out the definite integrals.

To conclude the proof, take the limit of the final line as n converges to infinity to conclude $\|y^n\| \to \infty$. Since weakly compact sets must be norm bounded, the claim follows.

Q.E.D.

APPENDIX E: THE BELLMAN PRINCIPLE OF OPTIMALITY

This section presents the Bellman Principle of Optimality for a primitive form problem on arbitrary state-spaces. The results are not unique, however, the purpose here is to clarify the existence of optimal policy functions on arbitrary spaces. The primitive form problem consists of:

1. a state-space $\mathbb{X}$
2. an action space $\mathbb{Y}$
3. a feasibility correspondence $\Lambda\colon \mathbb{X} \to \mathbb{Y}$
4. a transition function $\Phi\colon \operatorname{Gr}\Lambda \to \mathbb{X}$
5. a per-period action pay-off $u\colon \operatorname{Gr}\Lambda \to \mathbb{R}$
6. a discount fact $\beta \in (0,1)$

Define the correspondence mapping a current state to sequences of feasible actions and states, $\mathcal{H}_T\colon \mathbb{X} \to (\mathbb{X}\times\mathbb{Y})^{\mathbb{N}}$ as follows:

$$\mathcal{H}_T(x)\colon = \left\{(y_t, x_t)_{t=T}^{\infty} \,\middle|\, y_t \in \Lambda(x_t), x_{t+1} = \Phi(x_t, y_t), x_T = x, t \in \mathbb{N}, t \geq T\right\}$$

for any $x \in \mathbb{X}$. For $x \in \mathbb{X}$, define the primitive form value function:

$$V(x)\colon = \sup_{(x_t,y_t)_{t=0}^{\infty}\in\mathcal{H}_0(x)} \sum_{t=0}^{\infty} \beta^t u(x_t, y_t) \tag{E.16}$$

and the primitive form Bellman Equation:

$$V(x)\colon = \sup_{y\in\Lambda(x)} \left\{u(x_0, y) + \beta V(\Phi(x_0, y))\right\} \tag{E.17}$$

ASSUMPTION E.1 (**Growth Condition**) For any $x \in \mathbb{X}$, there exists a sequence of real numbers, $(m_t)_{t=0}^{\infty}$, such that for any $(x_t, y_t)_{t=0}^{\infty} \in \mathcal{H}(x)$, the following holds

$$|u(x_t, y_t)| \leq m_t, \qquad \forall t \in \mathbb{N} \tag{E.18}$$

and

$$\sum_{t=0}^{\infty} \beta^t m_t < \infty \tag{E.19}$$

THEOREM E.1 *Let Assumption E.1 hold. If a function $V\colon \mathbb{X} \to \mathbb{R}$ is the value function defined by* (E.16), *then V satisfies the Bellman Equation* (E.17). *Conversely, if a function $V\colon \mathbb{X} \to \mathbb{R}$ satisfies the Bellman Equation* (E.17) *and $\lim_{t\to\infty} \beta^t V(x_t) = 0$, then V is the value function defined by* (E.16).

PROOF: Let V be a value the function defined by (E.16). For any $x_0 \in \mathbb{X}$, by definition, the function V will satisfy

$$\begin{aligned}
V(x_0) &= \sup_{(x_t,y_t)_{t=0}^{\infty} \in \mathcal{H}_0(x_0)} \sum_{t=0}^{\infty} \beta^t u(x_t, y_t) \\
&= \sup_{(x_t,y_t)_{t=0}^{\infty} \in \mathcal{H}_0(x_0)} \left\{ u(x_0, y_0) + \sum_{t=1}^{\infty} \beta^t u(x_t, y_t) \right\} \\
&= \sup_{y_0 \in \Lambda(x_0), x_0} \sup_{(x_t,y_t)_{t=1}^{\infty} \in \mathcal{H}_1(\Phi(x_0,y_0))} \left\{ u(x_0, y_0) + \sum_{t=1}^{\infty} \beta^t u(x_t, y_t) \right\} \\
&= \sup_{y_0 \in \Lambda(x_0)} \left\{ u(x_0, y_0) + \sup_{(x_t,y_t)_{t=1}^{\infty} \in \mathcal{H}_1(\Phi(x_0,y_0))} \sum_{t=1}^{\infty} \beta^t u(x_t, y_t) \right\} \\
&= \sup_{y_0 \in \Lambda(x_0)} \left\{ u(x_0, y_0) + \beta V(\Phi(x_0, y_0)) \right\}
\end{aligned}$$

The second equality is a simple expansion of the infinite sum. The third equality follow from Lemma 1 by Kamihigashi (2008), which confirms we can split the supremum over into two suprema. The final equality holds from the definition of the value function.

Next, suppose V satisfies the Bellman Equation, Equation (E.17). Let $(x_t, y_t)_{t=0}^{\infty}$ be any sequence satisfying $(x_t, y_t)_{t=0}^{\infty} \in \mathcal{H}_0(x_0)$. Note

$$\begin{aligned}
V(x_0) &= \sup_{y \in \Gamma(x)} \{u(x_0, y) + \beta V(\Phi(x_0, y))\} \\
&\geq u(x_0, y_0) + \beta V(\Phi(x_0, y_0)) \\
&\geq u(x_0, y_0) + \beta u(x_1, y_1) + \beta^2 V(\Phi(x_1, y_1))
\end{aligned}$$

In particular, for any T,

$$\begin{aligned}
V(x_0) &= \sup_{y \in \Gamma(x)} \{u(x_0, y_0) + \beta V(\Phi(x_0, y_0))\} \\
&\geq \sum_{t=0}^{T-1} \beta^t u(x_t, y_t) + \beta^T V(x_T)
\end{aligned}$$

By Assumption E.1, $|u(x_t, y_t)| \leq m_t$ where $\sum_{t=0}^{\infty} m_t < \infty$. We thus satisfy the requirements of the dominated convergence theorem and conclude

$$\begin{aligned}
V(x_0) &\geq \lim_{T \to \infty} \left\{ \sum_{t=0}^{T-1} \beta^t u(x_t, y_t) + \beta^T V(x_T) \right\} \\
&= \lim_{T \to \infty} \sum_{t=0}^{T-1} \beta^t u(x_t, y_t) + \lim_{T \to \infty} V(x_T) \\
&= \sum_{t=0}^{\infty} \beta^t u(x_t, y_t)
\end{aligned}$$

where the third line uses the assumption $\lim_{t\to\infty} V(x_t) = 0$. Since $(x_t, y_t)_{t=0}^{\infty} \in \mathcal{H}(x_0)$ was arbitrary, $V(x_0)$ must satisfy (E.16), completing the proof.

Q.E.D.

THEOREM E.2 *Let Assumption E.1 hold and fix $x_0 \in \mathbb{X}$. If a sequence $(x_t, y_t)_{t=0}^{\infty} \in \mathcal{H}(x_0)$ achieves the value function*

$$V(x_0) = \sum_{t=0}^{\infty} \beta^t u(x_t, y_t) \tag{E.20}$$

then

$$V(x_t) = u(x_t, y_t) + \beta V(\Phi(x_t, y_t)), \qquad \forall t \in \mathbb{N} \tag{E.21}$$

Conversely, if any sequence $(x_t, y_t)_{t=0}^{\infty} \in \mathcal{H}_0(x_0)$ and function $V\colon \mathbb{X} \to \mathbb{R}$ satisfies (E.21) *and $\beta^t V(x_t) \to 0$, then $(x_t, y_t)_{t=0}^{\infty}$ achieves the value function at x.*

PROOF: Let the sequence $(x_t, y_t)_{t=0}^{\infty}$ achieve the value function. We will proceed by induction. Let $t = 0$, for any sequence $(\tilde{x}_t, \tilde{y}_t)_{t=0}^{\infty} \in \mathcal{H}_0(x_0)$, we have

$$\begin{aligned} V(x_0) &= u(x_0, y_0) + \beta U((x_t, y_t)_{t=1}^{\infty})) \\ &\geq u(x_0, \tilde{y}_0) + \beta U((\tilde{x}_t, \tilde{y}_t)_{t=1}^{\infty})) \end{aligned}$$

where we define $U((\tilde{x}_t, \tilde{y}_t)_{t=T}^{\infty}))\colon = \sum_{t=T}^{\infty} \beta^{t-T} u(\tilde{x}_t, \tilde{y}_t)$ for any $T \in \mathbb{N}$.

In particular, for any $(\tilde{x}_t, \tilde{y}_t)_{t=0}^{\infty} \in \mathcal{H}_0(x_0)$ such that $\tilde{y}_0 = y_0$, we have

$$u(x_0, y_0) + \beta U((x_t, y_t)_{t=1}^{\infty})) \geq u(x_0, y_0) + \beta U((\tilde{x}_t, \tilde{y}_t)_{t=1}^{\infty}))$$

implying

$$U((x_t, y_t)_{t=1}^{\infty})) \geq U((\tilde{x}_t, \tilde{y}_t)_{t=1}^{\infty}))$$

for any $(\tilde{x}_t, \tilde{y}_t)_{t=1}^{\infty} \in \mathcal{H}_1(\Phi(x_0, y_0))$. This gives $U((x_t, y_t)_{t=1}^{\infty}) = V(\Phi(x_0, y_0))$ and allows us to conclude (E.21) holds for $t = 0$.

Now make the inductive assumption and let (E.21) hold for any $T - 1$. We have

$$V(x_T) = u(x_T, y_T) + \beta U((x_t, y_t)_{t=T+1}^{\infty}) \geq u(x_T, \tilde{y}_T) + \beta U((\tilde{x}_t, \tilde{y}_t)_{t=T+1}^{\infty})$$

holds for any $(\tilde{x}_t, \tilde{y}_t)_{t=T}^{\infty} \in \mathcal{H}_T(x_T)$. In particular, for any $(\tilde{x}_t, \tilde{y}_t)_{t=T}^{\infty} \in \mathcal{H}_T(x_T)$ such that $\tilde{y}_T = y_T$, we have

$$u(x_T, y_T) + \beta U((x_t, y_t)_{t=T+1}^{\infty}) \geq u(x_T, y_T) + \beta U((\tilde{x}_t, \tilde{y}_t)_{t=T+1}^{\infty})$$

implying

$$U((x_t, y_t)_{t=T+1}^{\infty}) \geq U((\tilde{x}_t, \tilde{y}_t)_{t=T+1}^{\infty})$$

for any $(\tilde{x}_t, \tilde{y}_t)_{t=T+1}^{\infty} \in \mathcal{H}_{T+1}(\Phi(x_t, y_t))$. As such, $V(\Phi(x_t, y_t)) = U((x_t, y_t)_{t=T+1}^{\infty})$, allowing us to conclude (E.21) holds for all $t \in \mathbb{N}$.

Now suppose (x_t, y_t) with $(x_t, y_t)_{t=0}^{\infty} \in \mathcal{H}_0(x_0)$ and a function $V\colon \mathbb{X} \to \mathbb{R}$ satisfies (E.21) and $\beta^t V(x_t) \to 0$. By Theorem (E.1), V will be the value function and

$$V(x_0)\colon = \sup_{(x_t,y_t)_{t=0}^{\infty} \in \mathcal{H}_0(x)} \sum_{t=0}^{\infty} \beta^t u(x_t, y_t)$$

Since V and $(x_t, y_t)_{t=0}^{\infty}$ satisfies the Bellman Equation, note

$$\begin{aligned} V(x_0) &= u(x_0, y_0) + \beta V(x_1) \\ &= u(x_0, y_0) + \beta u(x_1, y_1) + \beta^2 V(x_2) \end{aligned}$$

moreover,

$$V(x_0) = \sum_{t=0}^{T-1} \beta^t u(x_t, y_t) + \beta^T V(x_T)$$

By Assumption E.1, $|u(x_t, y_t)| \leq m_t$ where $\sum_{t=0}^{\infty} m_t < \infty$. We satisfy the requirements of Dominated Convergence Theorem and conclude

$$V(x_0) = \lim_{T\to\infty} \sum_{t=0}^{T-1} \beta^t u(x_t, y_t) + \lim_{T\to\infty} \beta^T V(x_T) = \sum_{t=0}^{\infty} \beta^t u(x_t, y_t)$$

Thus $(x_t, y_t)_{t=0}^{\infty}$ achieves the value function as was to be shown. *Q.E.D.*

Define the optimal policy correspondence $H\colon \mathbb{X} \to \mathbb{Y}$ as

$$H(x) = \arg\max_{y \in \Lambda(x)} \left\{ u(x, y) + \beta V(\Phi(x, y)) \right\} \tag{E.22}$$

COROLLARY E.1 *Suppose Assumption E.1 holds. If the value function $V(x)$ is achieved by an optimal sequence for each x, then*

1. *there exists an optimal policy correspondence H defined by* (E.22)
2. *there exists an optimal policy function $h\colon \mathbb{X} \to \mathbb{Y}$ such that $h(x) \in H(x)$ for each x*
3. *for any $x \in \mathbb{X}$, the sequence $(x_t, h(x_t))_{t=0}^{\infty}$, where $x_{t+1} = \Phi(x_t, h(x_t))$ and $x_0 = x$, achieves the value function.*

PROOF: If the value function $V(x)$ is achieved by an optimal sequence for each x, then

$$V(x)\colon = \sup_{y \in \Lambda(x)} \left\{ u(x, y) + \beta V(\Phi(x, y)) \right\} = u(x, y_0) + \beta V(\Phi(x, y_0))$$

for $y_0 \in \Lambda(x)$. Since the above holds for any $x \in \mathbb{X}$, $H(x)$ will be non-empty for each x. By the axiom of choice, there exists a function $h\colon \mathbb{X} \to \mathbb{Y}$ such that $h(x) \in H(x)$ for each x (see Section 17.11 in Aliprantis and Border (2006)).

Finally, fix $x \in \mathbb{X}$. To show 3., for each t, we have

$$V(x_t) \colon = \sup_{y \in \Lambda(x_t)} \left\{ u(x_t, y) + \beta V(\Phi(x_t, y)) \right\}$$
$$= u(x_t, h(x_t)) + \beta V(\Phi(x_t, h(x_t)))$$

establishing, by Theorem E.2, that $(x_t, h(x_t))_{t=0}^{\infty}$ achieves the value function $V(x)$.

Q.E.D.

APPENDIX F: FURTHER DISCUSSION ON DYNAMIC PROGRAMMING LIMITATIONS

F.0.1. *Non-Compactness of The Feasibility Correspondence*

I consider the following topologies:

1. The weak topology if we let $\mathbb{Y} = L^2(S, \lambda)$, where λ is the Lebesgue measure (weakly closed and norm-bounded sub-sets are compact)
2. The weak topology if we let $\mathbb{Y} = L^1(S, \lambda)$ (order intervals are weakly compact)
3. The weak-star topology if we let $\mathbb{Y} = L^\infty(S, \lambda)$ (weak-star closed and norm-bounded sub-sets will are compact)
4. The weak topology if we let $\mathbb{Y} = \mathscr{C}b(S, \lambda)$, the space of continuous bounded functions on S

Consider the weak topology on $L^2(S, \lambda)$. If we let $\mathbb{M}$ be the space of Borel probability measures on S, then Φ will not be defined. To see why, specialise to the case where shocks are IID and take on discrete values $\{e_1, e_2\}$ with probability π_1 and π_2. Since the endowment shocks are independent of agents' previous shocks and current assets, we can track the state by tracking the marginal distribution on A at each t, $\tilde{\mu}_t$. The marginal distribution evolves according to

$$\tilde{\mu}_{t+1}(B) = \sum_{j \in \{1,2\}} \pi_j \int \mathbb{1}_B\{h(a, e_j)\} \tilde{\mu}_t(da), \qquad B \in \mathscr{B}(A)$$

Now suppose $\tilde{\mu}_t = \delta_x$ is the Dirac delta measure which puts all weight on a point $x \in A$. We write

$$\tilde{\mu}_{t+1}(B) = \sum_{j \in \{1,2\}} \pi_j \mathbb{1}_B\{h(x, e_j)\}, \qquad B \in \mathscr{B}(A)$$

Recalling $h(\cdot, e_j)$ satisfies $h(\cdot, e_j) \in L^2(S, \lambda)$, $h(\cdot, e_j)$ is an equivalence class of functions equal λ -almost everywhere; μ_{t+1} as defined above will not be a measure on $\mathscr{B}(\mathbb{R})$ because the evaluation

$$\mathbb{1}_B\{h(x, e_j)\} = \begin{cases} 1 & \text{if } \; h(x, e_j) \in (0, 1) \\ 0 & \text{if } \; h(x, e_j) \notin (0, 1) \end{cases}$$

is not defined. In particular, let h' and h'' be two functions belonging to the equivalence class h. Since the functions can differ on measure zero sets, we can have $h'(x, e_j) \in (0, 1)$ but $h''(x, e_j) \notin (0, 1)$.

We could take $\mathbb{M}$ to be the space of absolutely continuous measures on S. However, in this case, take h to be a constant function, then $\mu \circ h^{-1}$ will be the Dirac delta function, which is not absolutely continuous. The operator Φ will then map to values outside of $\mathbb{M}$.

Similar problems arise if we consider $\mathbb{Y} = L^1(S, \lambda)$ and $\mathbb{Y} = L^\infty(S, \lambda)$.

Finally, the space of continuous bounded real functions on S with the weak topology does not present useful compact sets: unit balls will not be weakly compact since the space is not reflexive and order intervals are only compact if the dual pairing is a symmetric Reisz pair (see section 8.16 in Aliprantis and Border (2006)).

F.1. *Discussion of Standard Dynamic Optimisation Theory*

Standard theory uses two approaches to verify existence in an infinite horizon dynamic optimisation problem: dynamic programming and product topology approaches. Both these approaches require the feasibility correspondences to have compact image sets. I briefly discuss why standard theory requires compact image sets below.

Write the Bellman Operator, T, for the recursive constrained planner's problem as

$$TV'(\mu) :=\sup_{h\in\Lambda(\mu)} \left\{u(\mu, h) + \beta V'(\Phi(\mu, h))\right\} \tag{F.23}$$

where V' is an extended real-valued function on $\mathbb{M}$.

Recall the standard dynamic programming procedure shows a fixed point to the Bellman Operator is the value function, V. If V is (semi) continuous and the feasibility correspondence compact valued, we can confirm existence of an optimal policy. (Semi) continuity of V is usually achieved by showing the Bellman Operator maps (semi) continuous functions to (semi) continuous functions,[18] which allows us to show sequences of iterations on the Bellman Operator converge to a (semi) continuous fixed point. To show the Bellman Operator preserves (semi) continuity, the standard approach (see Stachurski (2009), Appendix B and Stokey and Lucas (1989), chapter 4) is to use Berge's Theorem, which requires Λ to be upper hemicontinuous and compact-valued (Aliprantis and Border (2006), Lemma 17.30). And compact-valued upper hemicontinuous correspondences have compact image sets (Aliprantis and Border (2006), Lemma 17.8).

I discuss the product space approach using a reduced form stationary problem for easier notation, though similar ideas can work on the primitive form problem. Let $\mathbb{S}$, ρ and Γ be the state space, the feasibility correspondence and pay-offs for each t. The approach (see Le Van and Morhaim (2002) Theorem 1, Acemoglu (2009) Theorem 6.3, or Kamihigashi (2017) Proposition 6.1), assuming $\mathbb{S}$ is a metric space, works by showing the function $(x_t)_{t=0}^{\infty} \mapsto \sum_{t=0}^{\infty} \rho(x_t, x_{t+1})$ is upper semicontinuous on a compact space of feasible sequences $\mathcal{G}(x_0)$. To show $\mathcal{G}(x_0)$ is compact, we assume Γ is upper hemicontinuous and compact-valued, hence $\Pi_{t=0}^{\infty}\Gamma^t(x_0)$ is sequentially compact in the product topology. Since Γ and will also have closed graph (see Theorem 17.10 in Aliprantis and Border (2006)), and $\mathcal{G}(x_0) \subset \Pi_{t=0}^{\infty}\Gamma^t(x_0)$, $\mathcal{G}(x_0)$ will be compact.

[18]We can either show the Bellman Operator maps a space of bounded continuous functions to bounded continuous functions, which is complete, or use more general results, say by Kamihigashi (2014) Section 3.2, to show the limit of Bellman iteration can preserve (semi) continuity.

REFERENCES

Acemoglu, D., 2009. Introduction to Modern Economic Growth. Princeton University Press, Princeton, New Jersey.

Aliprantis, C. D., Border, K. C., 2006. Infinite Dimensional Analysis: A Hitchhiker's Guide. Springer-Verlag, Berlin.

Alt, H. W., Nürnberg, R., 2016. Linear Functional Analysis: An Application-Oriented Introduction. Universitext. Springer London.
URL `https://books.google.com.au/books?id=qyCcDAEACAAJ`

Barbu, V., Precupanu, T., 2012. Convexity and Optimization in Banach Spaces. Springer Monographs in Mathematics. Springer Netherlands.
URL `https://books.google.com.au/books?id=aKljTkrZgBIC`

Borovkov, A. A., 2013. Probability Theory, 4th Edition. Springer-Verlag, London.

Borwein, J. M., Vanderwerff, J. D., 2010. Convex Functions: Constructions, Characterizations and Counterexamples. Cambridge University Press, New York.

Çinlar, E., 2011. Probability and Stochastics. Springer-Verlag New York.

Halmos, P. R., 1974. Measure Theory. Springer-Verlag New York.

Kallenberg, O., 1997. Foundations of Modern Probability. Springer-Verlag New York.

Kamihigashi, T., 2008. On the Principle of Optimality for Nonstationary Deterministic Dynamic Programming. International Journal of Economic Theory 4, 519–525.

Kamihigashi, T., 2014. Elementary Results on Solutions to the Bellman Equation of Dynamic Programming: Existence, Uniqueness, and Convergence. Economic Theory 56 (2), 251–273.

Kamihigashi, T., 2017. A Generalisation of Fatou's Lemma for Extended Real-Valued Functions on sigma-Finite Measure spaces: with an Application to Infinite-Horizon Optimization in Discrete Time. Journal of Inequalities and Applications 2017 (1), 24.

Kelley, J. L., 1975. General Topology. Springer-Verlag New York.

Le Van, C., Morhaim, L., 2002. Optimal Growth Models with Bounded or Unbounded Returns: A Unifying Approach. Journal of Economic Theory 105 (1), 158–187.

Mas-colell, A., Zame, W. R., 1991. Equilibrium Theory in Infinite Dimensional Spaces. In: Hildenbrand, W., Sonnenschein, H. (Eds.), Handbook of Mathematical Economics, Vol. IV. pp. 1835–1890.

Megginson, R. E., 1998. An Introduction to Banach Spaces. Springer-Verlag New York.

Stachurski, J., 2009. Economic Dynamics: Theory and Computation. MIT Press Books, Cambridge, MA.

Stokey, N., Lucas, R., 1989. Recursive Methods in Economic Dynamics. Harvard University Press, Cambridge, MA.

Tao, T., 2010. Epsilon of Room, One: Volume 117 of Graduate Studies in Mathematics. American Mathematical Soc.
URL `https://books.google.com.au/books?id=DhWarYBllZAC`

Williams, D., 1991. Probability with Martingales. Cambridge University Press.