ONLINE APPENDIX FOR EXISTENCE OF CONSTRAINED OPTIMAL POLICIES IN THE HETEROGENEOUS AGENT GROWTH MODEL

Akshay Shanker[1]

Australian National University (ANU)

8. MATHEMATICAL PRELIMINARIES

8.1. *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

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

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 $\mathrm{Gr}\,\Gamma: \; = \{(x,y)|y \in \Gamma(x)\}$. A correspondence will have a **closed graph** if $\mathrm{Gr}\,\Gamma$ is closed.

A correspondence 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 8.1 *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 Aliprantis and Border (2005)) for a proof.

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

8.2. *Semicontinuity*

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

$$UC_f(\epsilon) \colon = \{x \in X \mid f(x) \geq \epsilon\} \tag{1}$$

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.

Equivalently, f is sequentially upper semi-continuous if for $x_n \to x$ with $x \in X$, we have $\limsup\limits_{n\to\infty} f(x_n) \leq f(x)$.

Let D be a subset of $\bar{\mathbb{R}}$. A function $f\colon X \to D$ will be called **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 $\epsilon > 0$.

8.3. *Probability and Conditional Expectation*

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

DEFINITION 8.1 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

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

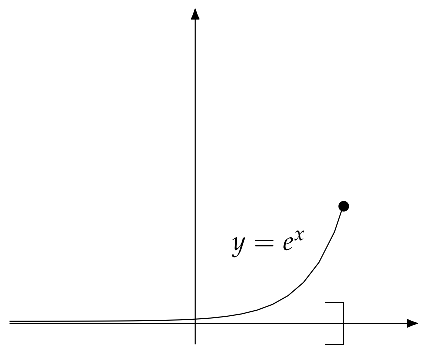
If y is a conditional expectation of x given $\mathcal{H}$, we write $y = \mathbb{E}(x|\mathcal{H})$.

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

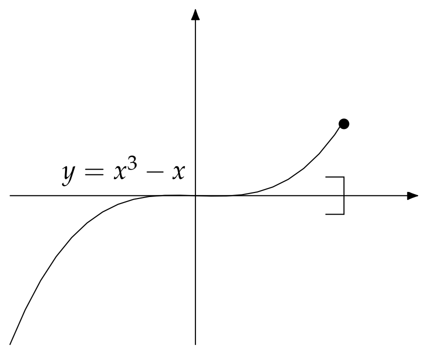
$$\sigma(\{y_i\}_{i\in F}) \colon = \sigma(\cup_{i\in F}\sigma(y_i))$$

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.



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.

FACT 8.1 *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 8.2 *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 8.3 ***(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 $h\colon \mathbb{R}^k \to \mathbb{R}^n$ such that $f = h \circ g$.*

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

FACT 8.4 ***(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 8.5 ***(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 $f_n \leq g$, 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).

Markov Properties

Let $S \subset \mathbb{R}$. Following Stachurski (2009), section 9.2, we can characterise any S- valued Markov process $(e_t)_{t=0}^{\infty}$ on a probability space $(\Omega, \Sigma, \mathbb{P})$ recursively. In particular, let $G\colon S \times \Omega \to S$, we will have

$$e_{t+1} = G(e_t, \eta_{t+1}), \qquad t \in \mathbb{N} \tag{2}$$

where $(\eta_t)_{t=1}^{\infty}$ is an I.I.D. sequence of random variables defined on $(\Omega, \Sigma, \mathbb{P})$ and e_0 is given.

The dynamics of an S-valued Markov process on $(Z, \mathscr{F}, P)$ can be summarised by a stochastic kernel Q. The value $Q(x, B)$ represents the probability that the Markov process moves from x to B, with $x \in S$ and $B \in \mathscr{B}(S)$ in a unit of time. The Markov kernel satisfies $Q(x, \cdot) \in \mathscr{P}(S)$ for each $x \in S$, where $\mathscr{P}(S)$ is the space of probability measures on S. Moreover, $Q(\cdot, B)$ is measurable for each B.

Letting $\eta_t \sim \phi$ for each t, we can relate the recursive characterisation of the Markov process to the stochastic kernel as follows

(3)
$$Q(x,B) = \mathbb{P}\{G(x,\eta_{t+1}) \in B\} = \mathbb{E}\mathbb{1}_B\{G(x,\eta_{t+1})\} = \int \mathbb{1}_B\{G(x,z)\}\phi(dz)$$

for $x \in S$ and $B \in \mathscr{B}(S)$.

9. PROOFS

Proofs for Section 2.1

Consider the setting described in section 2.1. 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 9.1 *Let Assumptions 2.1 and 2.1 hold. If $(x_t^i)_{t=0}^\infty$ is defined by the recursion in Equation* (4) *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* (5) *of the main paper.*

PROOF: We will proceed inductively 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. Moreover, the joint distribution of x_1^i and e_1^i will be

$$\begin{aligned}
\mu_1(B_A \times B_E) :&= \mathbb{P}\left\{x_1^i \in B_A, e_1^i \in B_E\right\} \\
&= \int \mathbb{1}_{B_A}\{x_1^i\} \times \mathbb{1}_{B_E}\{e_1^i\}\, d\mathbb{P} \\
&= \int\int\int \mathbb{1}_{B_A}\left\{h_0(x,e)\right\} \\
&\quad \times \mathbb{1}_B\left\{G(e,\eta)\right\}\mu_0(dx,de)\phi(d\eta) \\
&= \int\int \mathbb{1}_{B_A}\left\{h_0(x,e)\right\} \\
&\quad \times \left[\int \mathbb{1}_{B_E}\left\{G(e,\eta)\right\}\phi(d\eta)\right]\mu_0(dx,de) \\
&= \int\int \mathbb{1}_{B_A}\left\{h_0(x,e)\right\} Q(e,B_E)\mu_0(dx,de)
\end{aligned}$$

The first equality is given by the standard definition of expectations. The second equality follows from Equation (4), the recursive characterisation of the Markov process (Equation (2) above) and because μ_0 is the marginal distribution of $\{x_0^i, e_0^i\}$ and ϕ is the marginal distribution of the IID shock η_t^i. The final line follows from the properties of Markov kernel, in particular, Equation (3) above.

The above argument shows μ_1 and μ_0 satisfy the recursion (5) 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 (5) for each i. To see the claim holds for $t+1$,

$$
\begin{aligned}
\mu_{t+2}(B_A \times B_E) \colon &= \mathbb{P}\left\{x_{t+2}^i \in B_A, e_{t+2}^i \in B_E\right\} \\
&= \int \mathbb{1}_{B_A}\{x_{t+2}^i\} \times \mathbb{1}_{B_E}\{e_{t+2}^i\}\, d\mathbb{P} \\
&= \int \int \int \mathbb{1}_{B_A}\left\{h_{t+1}(x, e)\right\} \\
&\quad \times \mathbb{1}_{B_E}\left\{G(e, \eta)\right\} \mu_{t+1}(dx, de) \phi(d\eta) \\
&= \int \int \mathbb{1}_{B_A}\left\{h_{t+1}(x, e)\right\} \\
&\quad \times \left[\int \mathbb{1}_{B_E}\left\{G(e, \eta)\right\} \phi(d\eta)\right] \mu_{t+1}(dx, de) \\
&= \int \int \mathbb{1}_{B_A}\left\{h_{t+1}(x, e)\right\} Q(e, B_E) \mu_{t+1}(dx, de)
\end{aligned}
$$

Q.E.D.

For the following claim, consider the setting of section 2.1 and let $(h_t)_{t=0}^{\infty}$ be a sequence of measurable functions with $h_t \colon S \to A$ for each t. Let $(x_t^i)_{t=0}^{\infty}$ be a sequence of random variables generated by Equation (4) for each t. Let $(x_t^i)_{t=0}^{\infty}$ satisfy Equation (10) for each t. Recall by Claim 9.1 that $\left\{x_t^i, e_t^i\right\} \sim \mu_t$ for each t.

CLAIM 9.2 *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 will proceed 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 (10) in the main paper, the following holds

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

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 9.1 *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 $j \geq t$,*

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

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

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

Thus, we will prove the lemma by using the definition of conditional expectations at section 8.3 of the online appendix, and show

$$\int_B \mathbb{E}(y_{j+1}|\sigma(x_t, e_t, \dots, e_j))\, \mathrm{d}P = \int_B y_{j+1}\, \mathrm{d}P \tag{4}$$

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

We begin by verifying

$$\mathbb{E}(y_{j+1}|\sigma(x_t, e_t, \dots, e_j)) = \mathbb{E}(y_{j+1}|\sigma(x_t, e_t, \dots, e_j, \eta_{j+1})) \tag{5}$$

By construction of the Markov process at Equation (2) in the online appendix, $e_{i+1} = G(e_i, \eta_{i+1})$ for each $i \geq 1$, where $\sigma(\eta_{i+1})$ is independent of $\sigma(x_0, e_0, \eta_1, \dots, \eta_i)$ and $G\colon E \times Z \to E$ is measurable. As such, each e_i is a function of the shocks e_0 and $\eta_1, \dots, \eta_i$; applying the Doob-Dynkin Lemma (Fact 8.3), we have

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

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

Now, use Fact 8.2 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 8.1 in the online appendix, Equation (5) 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

$$\begin{aligned} \sigma(x_t, e_t, \ldots, e_{j+1}) &\subset \sigma\left(\sigma(x_t) \cup \sigma(e_t) \cup \cdots \cup \sigma(e_j) \cup \sigma(e_j, \eta_{j+1})\right) \\ &\subset \sigma\left(\sigma(x_t) \cup \sigma(e_t) \cup \cdots \cup \sigma(e_j) \cup \sigma(e_j) \cup \sigma(\eta_{j+1})\right) \\ &= \sigma\left(\sigma(x_t) \cup \sigma(e_t) \cup \cdots \cup \sigma(e_j) \cup \sigma(\eta_{j+1})\right) \end{aligned} \tag{6}$$

where the second inclusion follows from Fact 8.2.

To complete the proof by showing (4), let B satisfy $B \in \sigma(x_t, e_t, \ldots, e_{j+1})$. Recall Equation (5) and write

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

where the final equality comes from the definition of conditional expectation and since, by (6), B will satisfy $B \in \sigma(x_t, e_t, \ldots, e_j, \eta_{j+1})$.

Q.E.D.

PROOF OF CLAIM 6.1: We will prove the claim by using the definition of conditional expectation from section 8.3 showing $\mathbb{E}(y_t | \sigma(x_{t-1}, e_{t-1}))$ is $\sigma(x_t, e_t)$ measurable and satisfies

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

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 8.3).

Next, by Lemma 9.1, 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

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

as was to be shown to prove the claim.

Q.E.D.

Proofs for section 5

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

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

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 $L^2(Z, P)$.

Consider the set

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

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

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

The first and second lines follow from the definition of interest rates and aggregate capital given by Equation (17) 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.

10. THE BELLMAN PRINCIPLE OF OPTIMALITY

This section presents the elementary Bellman Principle of Optimality theorems 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:

$$V(x) \colon = \sup_{(x_t, y_t)_{t=0}^{\infty} \in \mathcal{H}_0(x)} \sum_{t=0}^{\infty} \beta^t u(x_t, y_t) \tag{7}$$

and the primitive form Bellman Equation:

$$V(x) \colon = \sup_{y \in \Lambda(x)} \left\{ u(x_0, y) + \beta V(\Phi(x_0, y)) \right\} \tag{8}$$

ASSUMPTION 10.1 (**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

$$|u(x_t, y_t)| \leq m_t, \qquad \forall t \in \mathbb{N} \tag{9}$$

and

$$\sum_{t=0}^{\infty} \beta^t m_t < \infty \tag{10}$$

THEOREM 10.1 *Let Assumption 10.1 hold. If a function $V \colon \mathbb{X} \to \mathbb{R}$ is the value function defined by* (7)*, then V satisfies the Bellman Equation* (8)*. Conversely, if a function $V \colon \mathbb{X} \to \mathbb{R}$ satisfies the Bellman Equation* (8) *and $\lim_{t \to \infty} \beta^t V(x_t) = 0$, then V is the value function defined by* (7)*.*

PROOF: Let V be a value the function defined by (7). 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 (8). 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 10.1, $|u(x_t, y_t)| \leq m_t$ 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 (7), completing the proof.

Q.E.D.

THEOREM 10.2 *Let Assumption 10.1 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*

$$V(x_0) = \sum_{t=0}^{\infty} \beta^t u(x_t, y_t) \tag{11}$$

then

$$V(x_t) = u(x_t, y_t) + \beta V(\Phi(x_t, y_t)), \qquad \forall t \in \mathbb{N} \tag{12}$$

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 (12) *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

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

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 (12) holds for $t = 0$.

Now make the inductive assumption and let (12) 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 (12) 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 (12) and $\beta^t V(x_t) \to 0$. By Theorem (10.1), 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 10.1, $|u(x_t, y_t)| \leq m_t$ 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

$$H(x) = \arg\max_{y \in \Lambda(x)} \left\{ u(x,y) + \beta V(\Phi(x,y)) \right\} \tag{13}$$

COROLLARY 10.1 *Suppose Assumption 10.1 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* (13)
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 (2005)).

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

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

establishing, by Theorem 10.2, that $(x_t, h(x_t))_{t=0}^{\infty}$ achieves the value function $V(x)$.

Q.E.D.

11. FURTHER DISCUSSION ON DYNAMIC PROGRAMMING LIMITATIONS

11.0.1. *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

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

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 (2005)).

11.0.2. *Discussion of Standard Dynamic Optimisation Theory*

Standard theory uses two approaches can be used 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, and as discussed in section 5.1 of the main paper, the constrained planner's sequential and recursive problems will not 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

$$TV'(\mu) := \sup_{h \in \Lambda(\mu)} \left\{ u(\mu, h) + \beta V'(\Phi(\mu, h)) \right\} \tag{14}$$

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,[1] 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 (2005) Lemma 17.30). And compact-valued upper hemicontinuous correspondences have compact image sets (Aliprantis and Border (2005) Lemma 17.8).

[1] 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.

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 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 (2005)), and $\mathcal{G}(x_0) \subset \Pi_{t=0}^{\infty}\Gamma^t(x_0)$, $\mathcal{G}(x_0)$ will be compact.

REFERENCES

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

Aliprantis, C. D., Border, K. C., 2005. Infinite Dimensional Analysis: A Hitchhiker's Guide. Springer-Verlag, Berlin.

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

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, 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.

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. pp. 1835–1890.

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

Stokey, N., Lucas, R., 1989. Recursive Methods in Economic Dynamics.

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