Existence of Constrained Optima in the Heterogeneous Agent Neoclassical Growth Model

Akshay Shanker*

Australian National University (ANU)

May 16, 2017

Draft

Abstract

This paper establishes the existence of sequential and recursive constrained social optima in a neoclassical growth model with infinitely many heterogeneous agents, incomplete markets and production, also known as the Aiyagari model. A constrained planner chooses individual saving and consumption levels, constrained by infinitely many agents' individual budget constraints, to maximise aggregate welfare. Due to the infinite dimensional state and action space, the dynamic optimisation problem for the constrained planner has discontinuous and non-compact feasibility correspondences, preventing the use of standard dynamic programming arguments on the Bellman Operator to show existence of optimal policies. To address the challenge, the paper first shows sequential optima are also recursive optima, without using a Bellman Equation. The paper then utilises a generalisation of Berge's Theorem for non-compact dynamic optimisation problems to verify existence of sequential optima.

*akshay.shanker@me.com.

1 Introduction

By relaxing traditional complete market and representative agent assumptions, growth models with heterogeneous agents and incomplete markets provide a more realistic framework for macroeconomic analysis. The foundations of heterogeneous agent models were set by Aiyagari (1994), Bewley (1977) and Huggett (1993) and the models are now said to have become a 'workhorse' in modern macroeconomics; macroeconomists use heterogeneous agent models to study issues such as consumption dynamics (Berger and Vavra, 2015; Kaplan and Violante, 2010), business cycles (Mckay and Reis, 2016), shapes of wealth distributions (Benhabib et al., 2015), asset pricing (Krusell et al., 2011) and monetary policy transmission Kaplan et al. (2016), to name a few. More recently, the literature (Dávila et al., 2012; Nuño and Moll, 2015; Chen and Yang, 2017; Nuno and Thomas, 2017) has begun asking questions about optimal government policy in heterogeneous agent models using the notion of a constrained planner, a planner who cannot complete insurance markets or overcome individual heterogeneity, but must pick personal saving and consumption levels while respecting each individual's budget constraint.

So far, the constrained planner's literature has established necessary conditions (Dávila et al. (2012) in discrete time Nuño and Moll (2015) in continuous time) and computed optimal solutions. However, whether or not constrained optima exist has not been verified and remains an open question. This paper provides an existence proof for discrete time recursive and sequential constrained optima in the standard and most popular heterogeneous agent model, the model considered by Dávila et al. (2012), the Aiyagari (1994) model. I also present a general result that can be applied to constrained planner problems in other heterogeneous agent models.

The constrained planner problem is the natural way to study optimal policy in an heterogeneous agent model because any planner wishing to implement policy as a competitive equilibrium will be a constrained planner. In particular, any planner wishing to implement policy uses instruments such as taxes and transfers such that given the taxes and transfers, decentralised agents whose decisions are constrained by their budget constraints also reach the planner's outcome. (See Chen and Yang (2017) and Nuno and Thomas (2017) for Ramsey constrained planner problems).

Computation of constrained optima have so far revealed interesting and unexpected policy implications; an existence proof is relevant for policy making since it confirms the results are well-defined. Dávila et al. (2012) show a decentralised equilibrium in the Aiyagari model may under-save compared to the constrained optima, which justifies an argument for saving subsidies, in contrast to the long held belief of sub-optimal over-saving, which justifies an argument for capital taxation (see Aiyagari (1995)). Dávila et al. (2012) also compute a planner's solution where sequences of asset distributions do not converge to a steady-state, but display ever increasing wealth inequality. Verifying existence helps ensure such policy conclusions and solution sequences are not pathological, and creates a foundation for further progress on understanding the dynamics of optimal policy in heterogeneous agent models.

The constrained planner's dynamic optimisation problem, however, presents mathematical challenges. Because we assume a continuum of heterogeneous agents, not only does the planner's action depend on the infinite dimensional distribution of agents, but the planner's actions are themselves infinite dimensional policy functions. In infinite dimensional spaces, it is more difficult to find continuous functions and compact sets, properties critical to assumptions made by existing dynamic optimisation theory (Stokey and Lucas (1989), Acemoglu (2009) ch.6 and Stachurski (2009)). As demonstrated by computations in Dávila et al. (2012), the infinite dimensional structure of the constrained planner's problem makes it possible for the planner to pick sequences of wealth distributions such that total capital is bounded but variance increases without bound, by increasing the wealth of a smaller and smaller measure of individuals. As a result, the constrained planner's state-space will not be compact. Moreover, for similar but more technical reasons I discuss in the paper, the constrained planner's problem fails to have continuous and compact-valued feasibility correspondences.

The constrained planner's problem in the literature is posed as a *recursive constrained planner's problem*. In each period, the planner selects a policy function that instructs agents on their assets tomorrow, based on their asset and shock today. The discontinuity and non-compactness of the feasibility correspondence for the recursive problem comes from two sources. First, it is difficult to bound individual agents' asset spaces, even though aggregate capital will be bounded above through properties of neoclassical production. This rules out Arzela-Arscoli type theorems to verify compactness. Second, it is difficult to verify a lower-bound on aggregate capital. As capital converges to zero, interest rates diverge and the norm of feasible asset distributions could diverge even though capital goes to zero. Both these features of the constrained planner's problem preclude the use of standard dynamic programming arguments — iteration on the Bellman Operator — to verify existence.

The strategy to address the mathematical challenges and establish existence for the recursive problem has two steps. First, the paper considers a *sequential constrained planner's* *problem*, where the planner instructs each agent on assets at a given time based on the agent's history of shocks up to that time. The sequential planner's state-space is the space of square integrable random variables representing assets, and with the weak topology, the state-space is a reflexive space, where compact sets are easier to find. The first technical innovation of the paper is a projection argument to show sequential solutions are also recursive solutions without using the Bellman Equation.

The sequential problem still does not resolve all of the non-compactness, particularity around regions where capital is zero. As the second technical innovation, the paper uses a generalisation of Berge's Maximum Theorem for sequence problems on non-compact spaces. The theorem extends ideas on non-compact dynamic programming found in Hernández-Lerma and Lasserre (1996), Feinberg et al. (2012) and Feinberg et al. (2013) by allowing functions to be bounded below. A proof of the theorem, along with a generalisation to arbitrary topological vector spaces, is presented in Shanker (2017b). In practice, the main assumption of the theorem can be verified by checking the variance of feasible sequences of asset distributions that give a strictly positive per-period pay-off at a time in the future are bounded.

Regarding previous work on dynamic optimisation on infinite dimensional spaces, a large literature Brock et al. (2014); Boucekkine et al. (2009); Fabbri et al. (2015) has shown existence and characterised optimal solutions in models of economic geography. In its current form, the economic geography planner *does not encounter non-compactness* as the constrained planner. The key difference is that the constrained planner endogenises common prices such as interest rates, or pecuniary externalities, which may become unbounded.

However, in addition to the models mentioned in the introductory paragraph, the results here may be relevant to optimal policy in models with aggregate shocks (Krusell and Smith, 1998), heterogeneous agent new Keynesian models (Bhandari et al., 2017; Brunnermeier and Sannikov, 2016; Kaplan et al., 2016), models of credit frictions and firm heterogeneity (Khan and Thomas, 2013; Buera and Moll, 2015), models of industry dynamics and trade (Hopenhayn, 1992; Melitz, 2003; Sampson, 2016) or models of social status (Genicot and Ray, 2015; Ray and Robson, 2012).[1]

[1]Even though in many of the cited models there is some form of heterogeneity, the distributions are static through time and the economy can be characterised by the mean of the distribution as in Buera and Moll (2015) or a cut-off value as in Hopenhayn (1992). However, extensions, where for example, there are spillovers between firms, as in Sampson (2016), the economy will need to be characterised by the entire distribution.

2 Constrained Planner's Problems

This section presents the recursive and sequential constrained planner's problems in a standard Aiyagari (1994) model. In the recursive problem, the planner selects policy functions which drive the wealth distribution of the economy and in the sequential problem, the planner will select random variables that map histories of shocks to assets. The recursive problem will be a stationary primitive form dynamic optimisation problem, where the planner selects an action (policy function) to drive a state (wealth distribution); by contrast the sequential problem will be a non-stationary reduced form dynamic optimisation problem, where the planner selects a sequence of states (random variables).

In the existing literature, the constrained planner's problem has been stated as a recursive problem, see for example, Dávila et al. (2012). The definition of recursive and sequential problems remains consistent with the use of the terms by Cao (2016) and Miao (2006), who consider existence of equilibria in heterogeneous agent models with aggregate shocks, and by Ljungqvist and Sargent (2004), sections 8.8- 8.9, who describe asset pricing equilibria.[2]

Some introduction to mathematical concepts used in the paper is in the Appendix.

2.1 The Aiyagari Model

Time is discrete and indexed by $t \in \mathbb{N}$. There are a continuum of identical individuals indexed by $i \in [0, 1]$. Define $A\colon = [\underline{a}, \infty)$ as the agents' asset space and define E as the agents' labour endowment space. Assume $E \subset \mathscr{B}(\mathbb{R}_+)$, where $\mathscr{B}(\mathbb{R}_+)$ are the Borel sub-sets of $\mathbb{R}_+$.

At time zero, each agent i draws an initial asset level x_0^i, with x_0^i taking values in A. In subsequent periods, each agent receives a sequence of labour endowment shocks $(e_t^i)_{t=0}^\infty$, with e_t^i taking values in E for each t and i. Assume a common probability space $(\Omega, \Sigma, \mathbb{P})$ for all uncertainty, that is, x_0^i and $(e_t^i)_{t=0}^\infty$ for each i are random variables defined on $(\Omega, \Sigma, \mathbb{P})$.

Assumption 2.1. The shocks $(e_t^i)_{t=0}^\infty$ and x_0^i have finite variance and are independent and identically distributed across i.

[2]The terminology of 'sequence' and 'recursive' problems may be overloaded. The distinction here concerns whether the planner assigns assets according to a history of shocks or according to the present shock, asset and aggregate distribution. For infinite dimensional problems, both the sequential and recursive problems can be written out as a deterministic sequence problem (maximising the sum of discounted pay-offs) and as a deterministic Bellman Equation — for example, (21) compared to (15).

Let μ_0 denote the common joint distribution of x_0^i and e_0^i. That is,

$$\mu_0(B) = \mathbb{P}\{\omega \in \Omega \,|\, x_0^i(\omega), e_0^i(\omega) \in B\}, \qquad\qquad B \in \mathscr{B}(S),\ i \in [0,1] \tag{1}$$

Let P denote the joint distribution of x_0^i and $(e_t^i)_{t=0}^{\infty}$.

Assumption 2.2. The shocks $(e_t^i)_{t=0}^{\infty}$ are a stationary Markov process with common Markov kernel Q and stationary marginal distribution ψ.

Assumption 2.2 can be relaxed to boundedness of the mean of the endowment shock, however, the stationarity assumption simplifies notation since total labour supply will be constant.

We do not need any further assumptions on E for the proofs in this paper to work. However, A will, in general, be unbounded above, even if E is bounded. While we can place a natural upper-bound on aggregate assets through standard properties of neoclassical production, it is still not clear whether sequences of asset distributions with a common upper bound on individual assets dominate all unbounded sequences of assets for the constrained planner. Dávila et al. (2012) assume an upper-bound each period on A, however, simulations by Dávila et al. (2012) (see fig.3) and by Nuño and Moll (2015) show a solution with diverging variance, implying a sequence of assets on an unbounded space.[3]

No Aggregate Uncertainty

No aggregate uncertainty formalises the idea that despite individual uncertainty, aggregate variables are almost deterministic; aggregate variables only depend on the common theoretical distribution of individual shocks, rather than realisations of the shocks, with probability one.

Letting λ denote Lebesgue measure:

Proposition 2.1. *Let $g\colon S \to \mathbb{R}$ be a measurable function such that $g(x_0^i, e_0^i)$ has finite variance. If Assumption 2.1 holds, then*

$$\int g(x_0^i, e_0^i)\lambda(di) = \int\int g(x,e)\mu_0(dx, de) \tag{2}$$

holds $\mathbb{P}$-almost everywhere.

[3]Popoviciu's inequality for variance states the variance of any bounded random variable is bounded. Dávila et al. (2012) compute a solution path with ever increasing variance that does not converge to an upper-bound.

Proof. The random variables $g(x_0^i, e_0^i)$ are uncorrelated across i, have a common bounded variance and a common mean $\int\int g(x,e)\mu_0(dx,de)$. We thus satisfy the conditions of Theorem 2 in Uhlig (1996), which gives the result. □

The following will then hold $\mathbb{P}$ - almost everywhere:

$$\lambda\{i \in [0,1] | x_0^i, e_0^i \in B\} = \int \mathbb{1}_B\{x_0^i, e_0^i\}\lambda(di) = \int\int \mathbb{1}_B(x,e)\mu_0(dx,de) = \mu_0(B)$$

The expression says that, with probability one, the empirical distribution of agents over S is the probability distribution of the individual shocks. This is the standard definition of no aggregate uncertainty used, for example, by Miao (2006) in Assumption 6 and by Bergin and Bernhardt (1992) in Definition 1.

The distribution μ_0 will be the initial state for the recursive constrained planner problem. The recursive problem we consider is one where the planner selects a measurable policy function $h_t \colon S \to A$ that instructs agents on $t+1$ assets given time t asset and shock. I will describe the constraints the planner faces below, however, now let us note the evolution of the distribution of agents through time. A sequence of policy functions $(h_t)_{t=0}^\infty$ chosen by the constrained planner generates a sequence of assets for each agent, $(x_t^i)_{t=0}^\infty$, by

$$x_{t+1}^i = h_t(x_t^i, e_t^i), \qquad t \in \mathbb{N},\ i \in [0,1] \tag{3}$$

Since h_t applies to all agents i, the distribution of $\{x_t^i, e_t^i\}$ will be identical across i. Moreover, I verify in the appendix (Claim 5.1) $\{x_t^i, e_t^i\} \sim \mu_t$ for each i, where $(\mu_t)_{t=0}^\infty$ satisfies the recursion

$$\mu_{t+1}(B_A \times B_E) = \int\int \mathbb{1}_{B_A}\{h_t(x,e)\}Q(e, B_E)\mu_t(dx,de), \qquad B_A \times B_E \in \mathscr{B}(S),\ \ t \in \mathbb{N} \tag{4}$$

If each x_t^i has finite variance, once again by Proposition 2.1, the time t empirical distribution of agents over S will satisfy

$$\lambda\{x_t^i, e_t^i \in B\} = \mu_t(B), \qquad B \in \mathscr{B}(S),\ \ t \in \mathbb{N} \tag{5}$$

with probability one.

Production

Turning to consumer utility, let $\nu \colon \mathbb{R}_+ \to \mathbb{R}_+$ be each consumer's utility function.

Assumption 2.3. The utility function ν is strictly increasing, bijective, concave and upper semi-continuous.

Regarding production, assume a representative firm renting capital from consumers and hiring workers to produce output Y_t

$$Y_t = F(K(\mu_t), L) - \delta K(\mu_t) \tag{6}$$

where $F\colon \mathbb{R}^2_+ \to \mathbb{R}_+$. Using again the LLN argument from Proposition 2.1, total capital in the economy is

$$K(\mu_t) \colon = \int \int x \mu_t(dx, de) = \int x_t^i \lambda(di) \tag{7}$$

Total supply of labour, L, will be constant according to Assumption 2.2.

Assumption 2.4. The production function F is differentiable on $\mathbb{R}_{++}$, homogeneous of degree one, strictly concave and for any $\hat{L} > 0$ and $\hat{K} > 0$ satisfies

1. $\lim_{K\to\infty} F_1(K, \hat{L}) = 0$ and $\lim_{K\to 0} F_1(K, \hat{L}) = \infty$
2. $F(0, \hat{L}) = F(\hat{K}, 0) = 0$
3. $K \mapsto F(K, \hat{L})$ is bijective

Budget Constraints and Utility

The interest and wage rates in the economy will be

$$r(\mu_t) \colon = F_1\left(K(\mu_t), L\right) - \delta, \quad w(\mu_t) \colon = F_2\left(K(\mu_t), L\right)$$

Given the economy-wide state μ_t, an agent i with asset x_t^i and endowment shock e_t^i must satisfy their budget constraint

$$\underline{a} \leq x_{t+1}^i \leq (1 + r(\mu_t)) x_t^i + w(\mu_t) e_t^i \tag{8}$$

where x_{t+1}^i is the next period asset. If x_0^i has finite variance and $r(\mu_t)$ us real-valued for each t, then if $(x_t^i)_{t=0}^\infty$ satisfies (8), x_t^i will have finite variance for each t, see Claim 5.2 in the Appendix.

Integrating across agents' budget constraints at Equation (8) and using the definition of interest and wages rates, along with homogeneity of the production function (see Theorem 2.1 in Acemoglu (2009)) gives us a law of motion for aggregate capital

$$K(\mu_{t+1}) \leq (1 + r(\mu_t))K(\mu_t) + w(x_t)L = F(K(\mu_t), L) + (1 - \delta)K(\mu_t) \tag{9}$$

From the law of motion above, we can assume an upper-bound $\bar{K} > 0$ on capital such that given any initial aggregate level of capital below $\bar{K}$, aggregate capital for wealth distributions satisfying (8) will never exceed $\bar{K}$ (see fig 9.3 in Acemoglu (2009)).

Assumption 2.5. The initial wealth distribution μ_0 satisfies $K(\mu_0) < \bar{K}$.

Time t utility for agent i will be

$$\nu\big((1 + r(\mu_t))x_t^i + w(\mu_t)e_t^i - x_{t+1}^i\big)$$

I leave out a definition of a competitive equilibrium as it is standard, for example, see Aiyagari (1994), Dávila et al. (2012), Kuhn (2013), Miao (2002) or Acikgoz (2015).

2.2 Recursive Constrained Planner

The recursive planner's state-space will be the space of probability distributions on S that satisfy assumption 2.1. Let $\mathbb{M} \subset \mathscr{P}(S)$ be such that each $\mu \in \mathbb{M}$ satisfies

- the marginal distribution across E, $\int \mu(dx, \cdot)$, agrees with ψ
- the marginal distribution across A, $\int \mu(\cdot, de)$, has finite variance
- aggregate assets satisfy $\int \int x\mu(dx, de) \in [0, \bar{K}]$.

To formalise the space of policy functions or *action-space*, let $\mathbb{Y}$ denote the space of measurable functions $h\colon S \to A$. The constrained planner picks a policy $h_t \in \mathbb{Y}$ for each t and agents' assets transition according to Equation (3).

Denote a correspondence $\Lambda\colon \mathbb{M} \twoheadrightarrow \mathbb{Y}$ mapping an economy's state to feasible policy functions as follows:

$$\Lambda(\mu)\colon = \begin{cases} h \in \mathbb{Y} \; | \underline{a} \leq h(x, e) \leq (1 + r(\mu))x + w(\mu)e, & \text{if } K(\mu) > 0 \\ h \in \mathbb{Y} \; | \displaystyle\int\int h(x, e)\mu(dx, de) = 0, \int\int h(x, e)^2\mu(dx, de) < \infty, & \text{if } K(\mu) = 0 \end{cases} \tag{10}$$

The inequality above is understood to hold μ_t - almost everywhere.

Following Equation (4), given a time t distribution of agents on S, μ_t, under a policy function h_t, the time $t+1$ distribution of agents is given by the operator $\Phi\colon \operatorname{Gr}\Lambda \to \mathbb{M}$ defined by:

$$\Phi(\mu, h)(B) = \int\int \mathbb{1}_{B_A}\{h(x,e)\}Q(e, B_E)\mu(dx, de) \qquad B_A \times B_E \in \mathscr{B}(S) \tag{11}$$

where $\mathscr{B}(S)$ are the Borel sets of S. We will write $\mu_{t+1} = \Phi(\mu_t, h_t)$.

TBC: show Φ is well-defined to $\mathbb{M}$

The constrained planner's utility, $u\colon \operatorname{Gr}\Lambda \to \mathbb{R}_+$ integrates across the empirical distribution of all agents

$$u(\mu, h) \colon = \int\int v\big((1 + r(\mu))x + w(\mu)e - h(x,e)\big)\mu(dx, de)$$

It is a straight-forward use of Jensen's inequality (Fact 5.4 in the appendix) and homogeneity of the production function to show the first integral is well-defined and real-valued.

Finally, let $\beta \in (0, 1)$ be a discount factor and let V denote the constrained planner's value function:

$$V(\mu_0) \colon = \sup_{(h_t, \mu_t)_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t u(\mu_t, h_t) \tag{12}$$

subject to

$$h_t \in \Lambda(\mu_t), \qquad \mu_{t+1} = \Phi(\mu_t, h_t), \qquad t \in \mathbb{N}, \qquad \mu_0 \,\text{given} \tag{13}$$

Definition 2.1. (**Recursive Constrained Planner's Problem**)

Given μ_0, a solution to the recursive constrained planner's problem is a sequence of measurable policy functions $(h_t)_{t=0}^{\infty}$, with $h_t\colon S \to A$ and a sequence of Borel probability measures on S, $(\mu_t)_{t=0}^{\infty}$ satisfying (13) that achieves the value function:

$$V(\mu_0) = \sum_{t=0}^{\infty} \beta^t u(\mu_t, h_t) \tag{14}$$

If the recursive constrained planner's problem has a solution, $(\mu_t, h_t)_{t=0}^{\infty}$, for each $\mu_0 \in \mathbb{M}$, then following standard arguments, we can show there exists a policy operator $H\colon \mathbb{M} \to \mathbb{Y}$ such that the sequence $(\mu_t, H(\mu_t))_{t=0}^{\infty}$ with $\mu_{t+1} = \Phi(\mu_t, H(\mu_t))$ solves the recursive problem. See Corollary 1.1 in Shanker (2017a) for a proof. The recursive characterisation of the policy implies that if a solution to the recursive constrained planner's problem exists, then the *policy function* that maps assets and shocks to next period assets depends only on the current distribution.

If it is defined, the Bellman Operator, T, for the recursive problem will be

$$TV'(\mu) :=\sup_{h\in\Lambda(\mu)} \{u(\mu,h)+\beta V'(\Phi(\mu,h))\} \tag{15}$$

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

However, we cannot use iteration on the Bellman Operator to show existence of a policy functions. In particular, we are unable to find a topology on sub-spaces of $\mathbb{Y}$ and $\mathbb{M}$ to ensure Φ is continuous and Λ compact valued. Further details are in Section 4. The challenge is partly addressed by switching to a sequential version of the constrained planner's problem.

2.3 Sequential Constrained Planner

For any initial state for the recursive problem, $\mu_0 \in \mathbb{M}$, let $\{x_0^i, e_0^i\} \sim \mu_0$ for each i and let P denote the joint probability distribution of $\{x_0^i, e_0^i, e_1^i \dots\}$.[4] Define $Z := A\times E^{\mathbb{N}}$ and consider the probability space $(Z,\mathscr{B}(Z),P)$. Let $\{x_0, e_0, e_1 \dots\}$ be a sequence of random variables representing draws from $(Z,\mathscr{B}(Z),P)$. By Proposition 2.1, we can see $\{x_0, e_0, e_1 \dots\}$ as a random variable whose realisation is a draw from the empirical distribution of individual shock values in Z.[5]

The sequential planner assigns period $t+1$ assets based on history to t, $\{x_0, e_0, e_1,\dots,e_t\}$. The sequential planner assumes strategies are symmetric across individuals, individuals with the same history are given the same instructions; the assumption is sufficient for our objective, which is show a sequential planner's solution also solves the recursive problem.

Let $\mathbb{X} := L^2(Z,P)$ be the space of square integrable real-valued functions on Z (further details about L^2 spaces is in section 3). The sequential planner's state-spaces will be sub-spaces of $\mathbb{X}$. Equip $\mathbb{X}$ with the weak topology. And define $(\mathscr{F}_i)_{i=0}^{\infty}$ as the filtration generated by the sequence of functions $\{a_0, e_0, e_1,\dots\}$. That is, $\mathscr{F}_i$ is the σ-algebra generated by the family of functions $\{a_0, e_0,\dots,e_{i-1}\}$.

[4]TBC: More formally, using Kolomogorov extension theorem, P should be constructed without reference to the random variables.

[5]Formally the random variable is an identity function id_Z, that is, for a realisation $z\in Z$, $x_0(z), e_0(z), e_1(z),\dots = \mathrm{id}_Z(z)$

For a random variable x with strictly positive mean defined on the space $Z\colon = A \times E^{\mathbb{N}}$, define

$$\begin{aligned}
\tilde{K}(x)\colon &= \int x \,\mathrm{d}P \\
\tilde{r}(x)\colon &= F_1(\tilde{K}(x), L) - \delta \\
\tilde{w}(x)\colon &= F_2(\tilde{K}(x), L)
\end{aligned}$$

Define the time t state-space for the sequential planner as follows:

$$\mathbb{S}_t\colon = \left\{ x \in m\mathscr{F}_t \,\middle|\, \underline{a} \leq x,\, \tilde{K}(x) \in [0, \bar{K}] \right\} \tag{16}$$

where $m\mathscr{F}_t$ is the space of $\mathscr{F}_t$-measurable functions in $\mathbb{X}$.

Next, for each t, define the feasibility correspondence $\Gamma_t\colon \mathbb{S}_t \to \mathbb{S}_{t+1}$ as follows:

$$\Gamma_t\left(x\right)\colon = \begin{cases} y \in \mathbb{S}_{t+1} \,\middle|\, \underline{a} \leq y \leq \left(1 + \tilde{r}\left(x\right)\right) x + \tilde{w}\left(x\right) e_t & \text{if } \tilde{K}(x) > 0 \\ y \in \mathbb{S}_{t+1} \,\middle|\, \int y \,\mathrm{d}P = 0 & \text{if } \tilde{K}(x) = 0 \end{cases} \tag{17}$$

For each t, define the sequential planner's pay-offs $\rho_t\colon \operatorname{Gr}\Gamma_t \to \mathbb{R}_+$ as follows:

$$\rho_t\left(x, y\right)\colon = \begin{cases} \displaystyle\int \nu\left(\left(1 + \tilde{r}\left(x\right)\right) x + \tilde{w}\left(x\right) e_t - y\right) \,\mathrm{d}P & \text{if } \tilde{K}(x) > 0 \\ 0 & \text{if } \tilde{K}(x) = 0 \end{cases} \tag{18}$$

Finally, let $\tilde{V}_0$ denote the time 0 sequential planner's value function

$$\tilde{V}_0(x_0)\colon = \sup_{(x_t)_{t=0}^{\infty}} \sum_{t=0}^{\infty} \beta^t \rho_t(x_t, x_{t+1})$$

subject to

$$x_{t+1} \in \Gamma_t(x_t),\ \forall t \in \mathbb{N}, \qquad x_0 \in \mathbb{S}_0 \text{ given} \tag{19}$$

Definition 2.2. (Sequential Constrained Planner's Problem)

Given $x_0 \in \mathbb{S}_0$, a solution to the sequential constrained planner's problem is a sequence of random variables $(x_t)_{t=0}^{\infty}$ satisfying (19) that achieve the sequential planner's value function:

$$\tilde{V}_0(x_0) = \sum_{t=0}^{\infty} \beta^t \rho_t(x_t, x_{t+1}) \tag{20}$$

I have used the notation $\tilde{V}_0$ to distinguish the sequential planner's value function from the recursive planner's value function, V.

2.4 Equivalence of Sequential and Recursive Problems

Given x_0, let $(y_t)_{t=0}^{\infty}$ be a solution to the sequential planner's problem. The solution $(y_t)_{t=0}^{\infty}$ will be a sequence of random variables mapping the history $\{x_0, e_0, e_1, \ldots, e_{t-1}\}$ to assets at each t. Construct a candidate sequence, $(x_t)_{t=0}^{\infty}$, as follows:

$$\begin{aligned} & x_0 = y_0, \qquad x_1 = \mathbb{E}(y_1|\sigma(x_0, e_0)) \\ & \text{and} \\ & x_{t+1} = \mathbb{E}(y_t|\sigma(x_t, e_t)), \qquad\qquad \forall t \in \mathbb{N} \end{aligned} \tag{21}$$

Our objective in this section is to show $(x_t)_{t=0}^{\infty}$ solves the recursive problem.

The term $\sigma(x_t, e_t)$ refers to the σ - algebra generated by the random variables x_t and e_t. The term $\mathbb{E}(y_t|\sigma(x_t, e_t))$ is the conditional expectation of y_t with respect to x_t and e_t. See Section 5.3 in the Appendix for details.

Since x_{t+1} is $\sigma(x_t, e_t)$ measurable, $x_{t+1} = h_t(x_t, e_t)$ for a measurable function h_t for each t.

The following claim will aid us with the proof show $(x_t)_{t=0}^{\infty}$ is a solution to the sequential problem.

Claim 2.1. *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* (21)*, then*

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

The proof requires some further measure theory preparation and is in the Appendix. The following proposition and theorem are the first technical innovation of the paper.

Proposition 2.2. *If $(y_t)_{t=0}^{\infty}$ is a solution to the sequential problem (Definition 2.2), then $(x_t)_{t=0}^{\infty}$ defined by* (21) *is a solution to the sequential problem.*

Proof. We first show the sequence $(x_t)_{t=0}^{\infty}$ is feasible and then show $(x_t)_{t=0}^{\infty}$ achieves the sequential planner's value function.

To show feasibility, first we need to verify $x_{t+1} \in \Gamma_t(x_t)$ for each t, where Γ_t is defined by (17). To check $x_t \in \mathbb{S}_t$ for each t, note $x_{t+1} = h_t(x_t, e_t)$ for a measurable function h_t for each t with x_0 given. Thus each 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. Furthermore, by the Tower Property of conditional expectation,

$$\int x_t \, \mathrm{d}P = \int \mathbb{E}(y_t|\sigma(x_{t-1}, e_{t-1})) \, \mathrm{d}P = \int y_t \, \mathrm{d}P \geq 0$$

Hence $\int x_t \, \mathrm{d}P \geq 0$ for each t, and we conclude $x_t \in \mathbb{S}_t$.

To check the second condition defining Γ_t, that individual assets satisfy feasibility, there are two cases to consider: first $\int x_t \, \mathrm{d}P > 0$ and second $\int x_t \, \mathrm{d}P = 0$.

Before proceeding, note the following holds due to the Tower Property,

$$\int y_t \, \mathrm{d}P = \int \mathbb{E}(y_t | \sigma(x_{t-1}, e_{t-1})) \, \mathrm{d}P = \int x_t \, \mathrm{d}P > 0 \tag{22}$$

Now suppose $\int x_t \, \mathrm{d}P > 0$. Since $(y_t)_{t=0}^{\infty}$ is a solution to the sequential planner's problem and satisfies (17), $y_{t+1} \leq (1 + \tilde{r}(y_t))y_t + \tilde{w}(y_t)e_t$. 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 (22), the third line follows from

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

A similar argument shows $\tilde{w}(y_t) = \tilde{w}(x_t)$. The fourth line follows from Claim 2.1 and the final line follows from the definition of x_t.

On the other hand, consider the case when $\int x_t \, \mathrm{d}P = 0$. We have $\int y_t = \int x_t \, \mathrm{d}P = 0$ by (22). As such, by (17), $\int x_{t+1} \, \mathrm{d}P = \int y_{t+1} \, \mathrm{d}P = 0$. Thus we have satisfied all the conditions, stated at (17), for x_{t+1} to belong to $\Gamma_t(x_t)$.

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 $\int x_t \, \mathrm{d}P > 0$. We have

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

where the second line is due to the definition x_t and x_{t+1}. The third line follows from Claim 2.1, the fourth line follows from Jensen's inequality (Fact 5.4 in the appendix) and the final line is due to the Tower Property.

If $\int x_t \,\mathrm{d}P = 0$, then $\rho_t(x_t, x_{t+1}) = 0$ by definition of ρ_t. Since $\int y_t \,\mathrm{d}P = 0$ by (22), $\rho_t(y_t, y_{t+1}) = 0$. We can then conclude

$$\tilde{V}(x_0) = \sum_{t=0}^{\infty} \beta^t \rho_t(y_t, y_{t+1}) \leq \sum_{t=0}^{\infty} \beta^t \rho_t(x_t, x_{t+1}) \tag{23}$$

Since $\tilde{V}(x_0)$ 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, we must have $\tilde{V}(x_0) = \sum_{t=0}^{\infty} \beta^t \rho_t(x_t, x_{t+1})$, allowing us to conclude $(x_t)_{t=0}^{\infty}$ is a solution to the sequential problem.

□

For any $(x_t)_{t=0}^{\infty}$ that is a solution to the sequential problem, we can define a sequence of Borel probability measures on S as follows:

$$\mu_t(B) :\, = P\{x_t, e_t \in B\} \tag{24}$$

Theorem 2.1. *Let μ_0 be given with $\{x_0, e_0\} \sim \mu_0$. If given x_0, $(x_t)_{t=0}^{\infty}$ is a solution to the sequential problem (Definition 2.2) that satisfies $x_{t+1} = h_t(x_t, e_t)$ for a sequence of measurable functions $(h_t)_{t=0}^{\infty}$, then:*

1. *the sequence of policy functions $(h_t)_{t=0}^{\infty}$ and the sequence of measures $(\mu_t)_{t=0}^{\infty}$ defined by* (24) *are a solution to the recursive problem (Definition 2.1)*
2. *the sequential and recursive planner's value functions satisfy $V(\mu_0) = \tilde{V}_0(x_0)$.*

The details of the proof are less interesting and are in the Appendix. In brief, the proof first verifies $(h_t)_{t=0}^{\infty}$ and $(\mu_t)_{t=0}^{\infty}$ constructed by (24) satisfy feasibility for the recursive problem. That is, they satisfy Equations (10) and (11) for each $t \in \mathbb{N}$. The proof then shows that the discounted sum of pay-offs from $(h_t)_{t=0}^{\infty}$ and $(\mu_t)_{t=0}^{\infty}$ given by (14) dominate the discounted sum of pay-offs from any other feasible sequence of Borel probability measures and policy functions.

3 Existence

3.1 General Existence Theorem for L^2 Spaces

To prepare the existence proof for the sequential problem, I will state a general result for a non-stationary reduced form dynamic optimisation problem. The result here can be used for other heterogeneous agent models.

Let $(Z, \mathscr{F}, P)$ be a probability space. The space $L^p(Z, P)$ denotes the space of measurable functions $x \colon Z \to \mathbb{R}$ such that $\|x\| :\, = \left(\int |x|^p \,\mathrm{d}P\right)^{\frac{1}{p}} < \infty$. A sequence of random variables $(x^n)_{n=0}^{\infty}$ with $x^n \in$

$L^p(Z,P)$ for each n, converges to x in the weak topology if $\int x^n h \,\mathrm{d}P \to \int xh \,\mathrm{d}P$ for each $L^p(Z,P)$. Subsets of $L^p(Z,P)$ that are norm-bounded and weakly closed will be sequentially compact by Alaoglu's Theorem (see Aliprantis and Border (2005) Theorem 6.21).

Convergence of random variables and compactness of sets of random variables from now will be with respect to the weak topology.

A non-stationary dynamic optimisation problem is a 5-tuple

$$\mathscr{E} = ((\mathbb{X}, \tau), (\mathbb{S}_t)_{t=0}^{\infty}, (\Gamma_t)_{t=0}^{\infty}, (\rho_t)_{t=0}^{\infty}, \beta)$$

consisting of:

1. A topological vector space $(\mathbb{X}, \tau)$, where $\mathbb{X} = L^2(Z,P)$ and τ is the weak topology
2. A collection of state-spaces $\{\mathbb{S}_0, \mathbb{S}_1, \dots\}$, with $\mathbb{S}_t \subset \mathbb{X}$ for each t
3. A collection of (possibly non-compact and discontinuous) feasibility correspondences $\{\Gamma_0, \Gamma_1, \dots\}$, with $\Gamma_t \colon \mathbb{S}_t \twoheadrightarrow \mathbb{S}_{t+1}$ for each t
4. A collection of per-period pay-offs $\{\rho_0, \rho_1, \dots\}$, with $\rho_t \colon \operatorname{Gr}\Gamma_t \to \mathbb{R}_+$ for each t
5. A discount factor $\beta \in (0,1)$.

Define the correspondence of **feasible sequences** $\mathcal{G}^T \colon \mathbb{S}_0 \twoheadrightarrow \prod_{t=0}^{T} \mathbb{S}_t$ starting at time $t = 0$ from $x \in \mathbb{S}_0$ and ending at time T

$$\mathcal{G}^T(x) : = \left\{ (x_t)_{t=0}^{T} \mid x_{t+1} \in \Gamma_t(x_t), x_0 = x \right\}$$

The value function for the non-stationary problem is

$$\tilde{V}_0(x) : = \sup_{(x_t)_{t=0}^{\infty} \in \mathcal{G}^{\infty}(x)} \sum_{t=0}^{\infty} \beta^t \rho_t(x_t, x_{t+1})$$

where $x \in \mathbb{S}_0$.

Given $x \in \mathbb{S}_0$, a solution to the reduced form, non-stationary problem is a sequence of random variables $(x_t)_{t=0}^{\infty} \in \mathcal{G}^{\infty}(x)$ that achieves the value function

$$\tilde{V}_0(x) = \sum_{t=0}^{\infty} \beta^t \rho_t(x_t, x_{t+1})$$

Let's turn to assumptions to guarantee existence of a solution to the general non-stationary reduced form problem.

Assumption 3.1. If C is a sequentially compact sub-set of $\mathbb{S}_0$, then there exists a sequence of non-negative real numbers $(m_t)_{t=0}^{\infty}$ such that any $(x_t)_{t=0}^{\infty} \in \mathcal{G}^{\infty}(C)$ satisfies

$$\rho_t(x_t, x_{t+1}) \leq m_t, \qquad \forall t \in \mathbb{N} \tag{25}$$

and

$$\sum_{t=0}^{\infty} \beta^t m_t < \infty \tag{26}$$

Assumption 3.2. The functions $(\rho_t)_{t=0}^{\infty}$ are upper semi-continuous for each $t \in \mathbb{N}$.

The following assumptions replace compactness and continuity of the feasibility correspondences.

Assumption 3.3. The correspondences $(\Gamma_t)_{t=0}^{\infty}$ have a closed graph for each $t \in \mathbb{N}$.

Assumption 3.4. Let $t \in \mathbb{N}$ and consider any $\epsilon > 0$. If C is a sequentially compact sub-set of $\mathbb{S}_0$, then there exists a constant $\bar{M}$ such that if $(x_i)_{i=0}^{t+1} \in \mathcal{G}^{t+1}(C)$ and $\rho_t(x_t, x_{t+1}) \geq \epsilon$, then $\|x_i\| \leq \bar{M}$ for any $i \in \{0, \ldots, t+1\}$.

Theorem 3.1. *If $\mathscr{E}$ satisfies Assumption 3.1 - 3.4, then*

a) the value function will satisfy $\tilde{V}_0 < \infty$ and for every $x \in \mathbb{S}_0$, there will exist a solution $(x_t)_{t=0}^{\infty}$ to the non-stationary problem

b) the value function $\tilde{V}_0$ will be upper semi-continuous.

The proof for the theorem still cannot use iteration of a non-stationary Bellman Operator to validate existence of an optimal policy. (See Hernández-Lerma (1989), Section 1.3 on how to use Bellman iteration on non-stationary problems.) Rather, the proof works in the product space $\Pi_{t=0}^{\infty}\mathbb{S}_t$ and uses Assumption 3.4 to show *some, but not all* upper-contour sets of the function $(x_t)_{t=0}^{\infty} \mapsto \sum_{t=0}^{\infty} \beta^t u_t(x_t, x_{t+1})$ on $\mathcal{G}^{\infty}(x_0)$ are compact in the product topology. For the proof and further discussion, see Shanker (2017b), Proposition 3.1.

3.2 Checking Conditions for the Sequential Constrained Planner's Problem

Consider again the setting of the Aiyagari model in sections 2.1 - 2.3. Let Assumptions 2.1- 2.5 hold and consider the economy $\mathscr{E} = ((\mathbb{X}, \tau), (\mathbb{S}_t)_{t=0}^{\infty}, (\Gamma_t)_{t=0}^{\infty}, (\rho_t)_{t=0}^{\infty}, \beta)$ where:

- $\mathbb{X} = L^2(Z, P)$, where $Z = A \times E^N$ and P is the joint distribution of $\{x_0^i, e_0^i, e_1^i, \ldots\}$

- The topology τ is the weak topology
- The sequence of state-spaces $(\mathbb{S}_t)_{t=0}^{\infty}$ are defined by (16)
- The sequence of correspondences $(\Gamma_t)_{t=0}^{\infty}$ are defined by (17)
- The sequence of pay-offs $(\rho_t)_{t=0}^{\infty}$ are defined by (18).

Proposition 3.1. ***(Checking Assumption 3.1)*** *If C is a sequentially compact sub-set of $\mathbb{S}_0$, then there exists a sequence of non-negative real numbers $(m_t)_{t=0}^{\infty}$ such that $\sum_{t=0}^{\infty} \beta^t m_t < \infty$ and any $(x_t)_{t=0}^{\infty} \in \mathcal{G}^{\infty}(C)$ satisfies $\rho_t(x_t, x_{t+1}) \leq m_t$ for each t.*

Proof. Let C be a sequentially compact sub-set of $\mathbb{S}_0$. By Assumption 2.5, aggregate capital is bounded above by $\bar{K}$. As such, for any $(x_t)_{t=0}^{\infty} \in \mathcal{G}^{\infty}(C)$ and for any t, we can use Jensen's inequality (Fact 5.4 in the appendix) to arrive at

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

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 $(x_t)_{t=0}^{\infty} \in \mathcal{G}^{\infty}(C)$,

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

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

$\square$

Remark 3.1. Let $(Z, \mathscr{F}, \mu)$ be a finite measure space. Let D be a convex set with $D \subset \mathbb{R}$. Consider a function $g \colon D \to \mathbb{R}$. Now define $G \colon L^2(Z, \mu) \to \mathbb{R}$ by

$$G(\zeta) = \int g(\zeta) \, d\mu$$

for $\zeta \in L^2(Z, \mu)$. If g is concave, then G will be weakly (sequentially) upper semi-continuous. For a proof, see Balder (1987) or Berkovitz (1974).

Proposition 3.2. ***(Checking Assumption 3.2)*** *The functions $(\rho_t)_{t=0}^{\infty}$ are upper semi-continuous for each t*

Proof. Recall the definition of upper semi-continuity from Section 5.2 in the Appendix.

Set any $t \in \mathbb{N}$ and consider sequences $(x^n)_{n=0}^{\infty}$ and $(y^n)_{n=0}^{\infty}$ with $\{x^n, y^n\} \in \operatorname{Gr}\Gamma_t$ for each n. Let $x^n \to x$ and $y^n \to y$ weakly with $y \in \Gamma_t(x)$. To verify upper semi-continuity, we will show

$$\limsup_{n\to\infty} \rho_t(x^n, y^n) = \limsup_{n\to\infty} \int \nu((1 + \tilde{r}(x^n))x^n + \tilde{w}(x^n)e_t - y^n) \, \mathrm{d}P \leq \rho_t(x, y) \tag{27}$$

Define $g(x,y) \colon = (1+\tilde{r}(x))x + \tilde{w}(x)e_t - y$. By remark 3.1, the mapping $\zeta \mapsto \int v(\zeta)\,\mathrm{d}P$ for $\zeta \in L^2(Z,P)$ will be upper semi-continuous since v is concave. As such, for any sequence in $L^2(Z,\mu)$ satisfying $\zeta^n \to \zeta$ weakly, we have

$$\limsup_{n\to\infty} \int v(g(\zeta^n))\,\mathrm{d}P \leq \int v(g(\zeta))\,\mathrm{d}P \tag{28}$$

First, we show (27) for the case $\int x\,\mathrm{d}P > 0$. If $\int x\,\mathrm{d}P > 0$, then $g(x^n,y^n) \to g(x,y)$ weakly, implying by (28),

$$\limsup_{n\to\infty} \int \nu(g(x^n,y^n))\,\mathrm{d}P \leq \int \nu(g(x,y))\,\mathrm{d}P = \rho_t(x,y)$$

On the other hand, if $\int x\,\mathrm{d}P = 0$, then

$$\begin{aligned}
\limsup_{n\to\infty} \int \nu(g(x^n,y^n))\,\mathrm{d}P &\leq \limsup_{n\to\infty} \nu\left(\int (1+\tilde{r}(x^n))x^n + \tilde{w}(x^n)e_t - y^n\,\mathrm{d}P\right) \\
&\leq \lim_{n\to\infty} \nu(F(\tilde{K}(x^n),L) + (1-\delta)\tilde{K}(x^n)) \\
&= 0 = \rho_t(x,y)
\end{aligned}$$

where the first inequality follows from Jensen's inequality. The second inequality follows from Assumption 2.4 on homogeneity of the production function (see (45) in the Appendix for further detail). $\square$

I have placed the proofs for the following in the Appendix.

Proposition 3.3. *(**Checking Assumption 3.3**) The correspondences $(\Gamma_t)_{t=0}^{\infty}$ have closed graph for each t.*

Proposition 3.4. *(**Checking Assumption 3.4**) Let $t \in \mathbb{N}$ and let $\epsilon > 0$. If C is a sequentially sub-set of $\mathbb{S}_0$, then there exists a constant $\bar{M}$ such that if $(x_i)_{t=0}^{t+1} \in \mathcal{G}^{t+1}(C)$ and $\rho_t(x_t, x_{t+1}) \geq \epsilon$, then*

$$\|x_i\| \leq \bar{M}$$

for all $i \in \{0,1,\dots,t+1\}$.

3.3 Existence for the Recursive Planner's Problem

The following theorem uses all the results we have studied so far to verify existence of recursive constrained optima.

Theorem 3.2. *If the recursive constrained planner's problem (Definition 2.1) satisfies Assumptions 2.1-2.5, then for any $\mu_0 \in \mathbb{M}$, there exists a solution $(\mu_t, h_t)_{t=0}^{\infty}$ such that $V(\mu_0) = \sum_{t=0}^{\infty} \beta^t u(x_t,h_t) < \infty$.*

Proof. Given μ_0, let P and $\{x_0, e_0, e_1, \dots\}$ be as defined in section 2.3. By assumptions 2.1- 2.5 and Propositions 3.1 - 3.4, the economy $\mathscr{E}$ satisfies Assumptions 3.1 - 3.4. Since Assumptions 3.1 - 3.4 satisfy the conditions for Theorem 3.1, there exists a sequence $(x_t)_{t=0}^{\infty}$ with $x_t \in \mathbb{S}_t$ for each t solving the Sequential Planner's problem (Definition 2.2) such that $\tilde{V}(x_0) < \infty$.

By Proposition 2.2, there exists $(h_t)_{t=0}^{\infty}$ with $h_t \colon S \to A$ and $x_{t+1} = h_t(x_t, e_t)$ for each t. Furthermore, by Theorem 2.1, $(h_t)_{t=0}^{\infty}$ and $(\mu_t)_{t=0}^{\infty}$ defined by Equation (24) solve the recursive problem and

$$V(\mu_0) = \tilde{V}(x_0) = \sum_{t=0}^{\infty} \beta^t \rho_t(x_t, x_{t+1}) = \sum_{t=0}^{\infty} \beta^t u(\mu_t, h_t) < \infty \tag{29}$$

□

4 Difficulties using Standard Dynamic Programming

Recall the standard dynamic programming procedure is to show a fixed point to the Bellman Operator is the value function. We may be able to find a fixed point V to the Bellman Operator at (15), even without making topological assumptions on $\mathbb{M}$ or $\mathbb{Y}$ by using new results in Kamihigashi (2014).[6] However, we still face the challenge of showing existence of an optimal policy; to replace sup with max in the Bellman Operator, we need (semi) continuity on V and we require Λ to be compact valued. Continuity of V is usually achieved by showing the Bellman Operator maps (semi) continuous functions to (semi) continuous functions, which allows us to show sequences of iterations on the Bellman Operator converge to a continuous fixed point. To show the Bellman Operator preserves (semi) continuity, the standard approach (see Stachurski (2009), Appendix B and Stokey and Lucas (1989), Section 3.3) is to use Berge's Theorem:

Theorem 4.1. ***Berge's Theorem (Aliprantis and Border (2005) Lemma 17.30)*** *If $f \colon X \times Y \to \bar{\mathbb{R}}$ is an upper semi-continuous function and $\Gamma \colon X \twoheadrightarrow Y$ is a non-empty compact valued and upper hemi-continuous correspondence, then the value function $V(x) = \max_{y \in \Gamma(x)} f(x, y)$ exists and is upper semi-continuous.*

We then require *joint* semi-continuity of

$$\{h, \mu\} \mapsto u(\mu, h) + \beta V'(\Phi(\mu, h))$$

In addition to a compact valued and upper-hemicontinuous Λ. Suppose V' and u were semi-continuous, we then require Φ to be continuous. There are two distinct challenges:

[6] Kamihigashi (2014) gives results for a reduced form problem, though the arguments could be extended to incorporate primitive form problems such as the recursive planner's problem

1. Difficulties in finding a topology on $\mathbb{M}$ and $\mathbb{Y}$ such that Λ is compact valued while Φ is continuous. The issue is resolved by switching to the sequential planner's problem.
2. Once we switch to the sequential planner's problem, difficulties from $(\Gamma_t)_{t=0}^{\infty}$ not being compact valued *and* hemi-continuous near regions where aggregate becomes zero. The issue is addressed by replacing the requirements of Berge's Theorem with Assumption 3.4.[7]

Regarding the first challenge, the most promising candidate topological space for the action-space is to consider continuous functions in $\mathbb{Y}$. However, because A is unbounded, Arzela-Arscoli Theorem breaks down and families of equicontinuous functions on A will not be compact in the sup-norm topology. (Recall the discussion of why A is unbounded in section 2.1.) The appendix gives further detail of pathologies in other topologies.

If we assumed shocks were bounded, we could try restricting $\mathbb{M}$ to spaces of measures that assign zero probability to assets over a certain constant, then $\Lambda(\mu)$ will essentially contain functions on a bounded domain, even though a sequence of distributions through time can still be unbounded. To justify the restriction, we would still need to show an equicontinuous family of functions dominates all other policy functions, for example by showing monotone functions dominate all other policy functions — this line of argument has so far not yielded success.

Turning to the second challenge, since we cannot rule out aggregate capital arbitrarily close to zero, to give the correspondences Γ_t a closed graph, $\Gamma_t(x)$ includes all y such that $\int y \, \mathrm{d}P = 0$ if $\int x \, \mathrm{d}P = 0$. The feasibility correspondence will not be compact valued since if $\tilde{K}(x) = 0$, the correspondence maps to a norm unbounded set. (Recall weakly compact sets are norm bounded as discussed in Appendix A.)

We could try and make the correspondence *compact-valued* by point-wise bounding values of y when $\tilde{K}(x) = 0$. However, we will still not satisfy hemi-continuity. Recall (see Section 5.1 in the Appendix) the image of a compact set under a compact-valued and upper hemi-continuous is compact. To show Γ_t will not be hemi-continuous, it will suffice to construct a compact set C such that $\Gamma_t(C)$ is norm unbounded for some t. Let x_0, the initial assets, be a uniform random variable on the interval $[0, 1]$. Assume the random variable e_0 is large enough to satisfy $\tilde{w}(x_0)e_0 > 1$. Now define the set C as follows

$$C\colon = \{x \in \mathbb{S}_0 \,|\, \underline{a} \le x \le (1 + \tilde{r}(x))x + \tilde{w}(x)e_0\}$$

[7] The generalisation of Berge's Theorem I present in this paper, Assumption 3.4, still requires some upper-contour sets of the sum of per-period pay-offs to be compact in the product topology on $\mathbb{X}^{\mathbb{N}}$. Adapting Assumption 3.4 to the recursive problem still does not resolve the non-compactness and discontinuity encountered (point 1) in the recursive problem. In particular, it will be difficult to find compact upper contour sets in the space of feasible states and actions.

The set C will be compact since it is norm bounded and weakly closed. Consider the set $\Gamma_1(C)$

$$\Gamma_1(C) = \begin{cases} y \in \mathbb{S}_1 \,|\, \underline{a} \leq y \leq (1+\tilde{r}(x))x + \tilde{w}(x)e_1 \quad \text{for some } x \in C & \text{if } \tilde{K}(x) > 0 \\ y \in \mathbb{S}_1 \,|\, \int y \,\mathrm{d}P = 0, y \leq \tilde{M} & \text{if } \tilde{K}(x) = 0 \end{cases}$$

Where $\tilde{M}$ is a constant. Now construct an unbounded sequence in $\Gamma_1(C)$ as follows: first define a sequence x^n, with $x^n \in C$ for all n, as $x^n(a, e_0) = \sin(na)$ and then define $y^n := (1 + \tilde{r}(x^n))x^n$. Clearly $y^n \in \Gamma_1(C)$.

Now use the definition of L^2 norm to write

$$\|y^n\| = \left(\int [(1+\tilde{r}(x^n))x^n]^2 \,\mathrm{d}P \right)^{\frac{1}{2}} = (1+\tilde{r}(x^n)) \left(\int_0^1 \sin^2(na) \,\mathrm{d}a \right)^{\frac{1}{2}}$$

where $\tilde{r}(x^n) = F_1 \left(\int_0^1 \sin(na) \,\mathrm{d}a, L \right)$, whence $\tilde{r}(x^n) \to \infty$. However, we know

$$\left(\int_0^1 \sin^2(na) \,\mathrm{d}a \right)^{\frac{1}{2}} \to \frac{1}{2}$$

We can therefore conclude $\|y^n\| \to \infty$.

4.1 Relationship to Challenges in the General Equilibrium Literature

The issues raised here are also encountered by Miao (2006) and Cao (2016) in proofs for equilibria in the heterogeneous agent model with aggregate shocks. Miao (2006) overcomes the issue of unbounded interest rates by using only production functions which do not satisfy the standard neoclassical production assumptions (Inada conditions). To allow unbounded interest rates, Cao (2016) presents a proof which uses consumer necessary conditions for a sequence of finite horizon equilibria to show capital is bounded below. Cao (2016) is also able to justify search for a general equilibrium on the space of monotone policy functions. The proof using the consumer's Euler equation, shows any equilibrium policy function, if it exists, must be monotone. We know each agent's dynamic optimisation problem has a solution through standard finite dimensional dynamic programming arguments. To show existence of equilibria, it is then sufficient to find *any* sequence of policy functions that solve agents' problems and satisfy the fixed point conditions implied by the general equilibrium.

Unfortunately, I have yet not been able to use similar arguments for the constrained planner's problem. Even if we could show the planner's necessary Euler equations (as derived by Dávila et al. (2012)) imply monotone policy functions, an optimum may not exist. In particular, if we restrict the space of feasible functions for the aggregate planner's problem to monotone functions and find a optimum for the restricted space, we still need to explicitly show the optimum we have found dominates all other sequences in the unrestricted original problem: for the constrained planner's problem, I have been unable to show monotone policy functions with aggregate capital bounded below dominate all other policy functions. (Notwithstanding, the approach taken by Cao (2016) does resemble intuitively the boundedness condition on variance of Assumption 3.4.)

However, once existence has been established, whether or not it is possible to iterate on the Bellman Equation for computation is a promising immediate area of research.

5 Conclusion

The paper presented a proof of existence for constrained optima in a standard Aiyagari (1994) model. While the constrained planner's problem is a natural way to understand optimal policy in heterogeneous agent models, mathematical difficulties with the standard recursive problem, arising from the infinite dimensional state and action space, means we cannot apply existing dynamic programming theory to ensure the planner's problem is well-defined. The paper addressed the mathematical challenges by first shifting to a sequential characterisation of the constrained planner's problem and showing sequential solutions are also recursive solutions. We then used dynamic optimisation results for non-compact state-spaces from Shanker (2017b) to show existence of a solution to the sequential problem.

An immediate path for further work is to develop computational methods which are known to converge to the true constrained optima. It is hoped the results here can contribute to the endeavour. As detailed in the introduction, further research should also examine constrained optima in models belonged to the wider class of heterogeneous agent models.

Appendix A: Mathematical Preliminaries

5.1 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 will have a **closed graph** if $\operatorname{Gr}\Gamma$ 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 hemi-continuous correspondences need not be compact valued or have closed graph. Closed graph correspondences also need not be upper hemi-continuous. However, one useful result is that the image of a compact set under a compact valued and upper hemi-continuous correspondence will be compact (see Lemma 17.8 and related results in chapter 17 of Aliprantis and Border (2005)).

5.2 Semicontinuity

Let (X, τ) be a topological vector space. A function $f \colon X \to \bar{\mathbb{R}}$ is **upper semi-continuous** if the upper contour sets

$$UC_f(\epsilon) \colon = \{x \in X \,|\, f(x) \geq \epsilon\}$$

are sequentially closed for all $\epsilon \in \mathbb{R}$. Equivalently, f will be upper-semicontinuous if for any $x^n \to x$, with $x \in X$,

$$\limsup_{n \to \infty} f(x^n) \leq f(x)$$

5.3 Probability and Conditional Expectation

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

Definition 5.1. 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.

Fact 5.1. *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). □

Fact 5.2. *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 $B \in \sigma(\mathcal{A} \cup \mathcal{B} \cup \mathcal{C})$. □

Fact 5.3. ***(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). □

Fact 5.4. ***(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).

5.4 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

$$e_{t+1} = G(e_t, \eta_{t+1}), \qquad t \in \mathbb{N} \tag{30}$$

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

$$Q(x, B) = \mathbb{P}\{G(x, \eta_{t+1}) \in B\} = \mathbb{E}\mathbb{1}_B G(x, \eta_{t+1}) = \int \mathbb{1}_B G(x, z)\phi(dz) \tag{31}$$

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

Appendix B: Proofs for Sections 2.1 and 2

The first proof defines the distributions of $(x_t^i)_{t=0}^{\infty}$ and $(e_t^i)_{t=0}^{\infty}$ under a sequence of policy functions $(h_t)_{t=0}^{\infty}$. Consider the setting described in section (2.1).

Claim 5.1. *Let Assumptions 2.1 and 2.1 hold. If $(x_t^i)_{t=0}^{\infty}$ is defined by the recursion in Equation* (3) *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* (4).

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

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

The first equality is given by the standard definition of expectations. The second equality follows from Equation (3), the recursive characterisation of the Markov process (Equation (30)) 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 (31).

The above argument shows μ_1 and μ_0 satisfy the recursion (4) and hence the claim holds for $t = 0$.

Now make the inductive assumption that the claim holds for arbitrary t, that is, $\{x^i_{t+1}, e^i_{t+1}\} \sim \mu_{t+1}$ and $\{x^i_t, e^i_t\} \sim \mu_t$ where μ_t and μ_{t+1} satisfy the recursion (4) for each i. To see the claim holds for $t+1$,

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

□

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

Lemma 5.1. *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 5.3 and showing

$$\int_B \mathbb{E}(y_{j+1}|\sigma(x_t, e_t, \dots, e_j))\, \mathrm{d}P = \int_B y_{j+1}\, \mathrm{d}P \tag{32}$$

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

We begin by verifying

$$\mathbb{E}(y_{j+1}|\sigma(x_t, e_t, \dots, e_j)) = \mathbb{E}(y_{j+1}|\sigma(x_t, e_t, \dots, e_j, \eta_{j+1})) \tag{33}$$

By construction of the Markov process at Equation (30), $e_{i+1} = G(e_i, \eta_{i+1})$ 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, 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 $\sigma(x_0, e_0, \eta_1, \ldots, \eta_j)$ are independent, $\sigma(\eta_{j+1})$ and $\sigma(x_t, e_t, \ldots, e_j)$ will also be independent.

Now, using Fact 5.2 allows us 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 5.1, Equation (33) 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

$$\begin{aligned}
\sigma(x_t, e_t, \ldots, e_{j+1}) &\subset \sigma\left(\sigma(x_t) \cup \sigma(e_t) \cup \cdots \cup \sigma(e_j) \cup \sigma(e_j, \eta_{j+1})\right) \\
&\subset \sigma\left(\sigma(x_t) \cup \sigma(e_t) \cup \cdots \cup \sigma(e_j) \cup \sigma(e_j) \cup \sigma(\eta_{j+1})\right) \\
&= \sigma\left(\sigma(x_t) \cup \sigma(e_t) \cup \cdots \cup \sigma(e_j) \cup \sigma(\eta_{j+1})\right)
\end{aligned} \tag{34}$$

where the second inclusion follows from Fact 5.2.

To complete the proof by showing (32), take any $B \in \sigma(x_t, e_t, \ldots, e_{j+1})$. Recalling independence of $\sigma(\eta_{j+1})$ and Equation (33), we write

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

where the final equality comes from the definition of conditional expectation and since, by (34), B will satisfy $B \in \sigma(x_t, e_t, \ldots, e_j, \eta_{j+1})$.

□

Claim 5.2. *Consider the setting of Section 2.3 and consider a sequence $(x_t^i)_{t=0}^{\infty}$ generated by a sequence of measurable policy functions $(h_t)_{t=0}^{\infty}$ satisfying Equation* (3) *and* (8)*. If Assumption 2.1 holds, then x_t^i has finite variance for each t.*

Proof. We will proceed by induction and show that if x_t^i has finite variance for some t, then x_{t+1}^i will have finite variance. Observe

$$x_{t+1}^{i\,+} \leq (1 + r(\mu_t))x_t^i + w(\mu_t)e_t \implies (x_{t+1}^{i\,+})^2 \leq ((1 + r(\mu_t))x_t^i + w(\mu_t)e_t)^2$$

and

$$x_{t+1}^{i\,-} \geq \underline{a} \implies (x_{t+1}^{i\,-})^2 \leq \underline{a}^2$$

Use $x_{t+1}^i = x_{t+1}^{i\,+} - x_{t+1}^{i\,-}$ to write

$$\begin{aligned}
(x_{t+1}^i)^2 &\leq (x_{t+1}^{i\,+})^2 + (x_{t+1}^{i\,-})^2 \\
&\leq ((1 + r(\mu_t))x_t^i + w(\mu_t)e_t)^2 + \underline{a}^2
\end{aligned}$$

The integral of the right hand side above will be finite since x_t^i and e_t have finite variance. As such, integrating across both sides reveals

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

To conclude, x_1^i will have finite variance since x_0^i has finite variance by assumption. Moreover, if x_t^i have finite variance, then x_{t+1}^i will have finite variance. By the principle of induction x_t^i will have finite variance for each t. □

Proof of Claim 2.1. We will prove the claim by using the definition of conditional expectation from section 5.3 and 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.

Next, by Lemma 5.1, 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 $B \in \sigma(x_t, e_t)$. We have $B \in \sigma(x_{t-1}, e_{t-1}, e_t)$, thus

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

as was to be shown to prove the claim.

□

Proof of Theorem 2.1. Let x_0 be a random variable on $(Z, \mathscr{B}(Z), P)$ and let μ_0 be the distribution of x_0. Let $(x_t)_{t=0}^{\infty}$ denote a solution to the sequential problem (Definition 2.2) that 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 (24) and note by definition that μ_t will the distribution of $\{x_t, e_t\}$ for each t.

The proof proceeds in two steps. To show $(\mu_t)_{t=0}^{\infty}$ and $(h_t)_{t=0}^{\infty}$ solve the recursive problem, we will first show feasibility in part 1 and then in part 2 show that the sum of discounted pay-offs from $(\mu_t)_{t=0}^{\infty}$ and $(h_t)_{t=0}^{\infty}$ dominate the sum of discounted pay-offs from any other feasible sequence of distributions and policy functions.

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

Our first task is to show $(\mu_t, h_t)_{t=0}^{\infty}$ satisfies Equations (10) and (11) for each $t \in \mathbb{N}$. Fix any $t \in \mathbb{N}$, to show (10), we have to 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 will show

$$\mu_t\{a, e \in S \mid h_t(a, e) > (1 + r(\mu_t))a + w(\mu_t)e\} = 0$$

The condition says the policy function h_t satisfies agents' budget constraints μ_t - almost everywhere. Using the definition of μ_t by Equation (24), we have

$$\begin{aligned}
&\mu_t\{a, e \in S \mid h_t(a, e) > (1 + r(\mu_t))a + w(\mu_t)e\} \\
&= P\{\omega \in Z \mid h_t(x_t(\omega), e_t(\omega)) \\
&\quad > (1 + \tilde{r}(x_t))x_t(\omega) + \tilde{w}(x_t)e_t(\omega)\} \\
&= P\{\omega \in Z \mid x_{t+1}(\omega) \\
&\quad > (1 + \tilde{r}(x_t))x_t(\omega) + \tilde{w}(x_t)e_t(\omega)\} \\
&= 0
\end{aligned} \tag{35}$$

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

$$\int\int x\mu_t(dx, de) = \int x_t \, \mathrm{d}P > 0 \tag{36}$$

whence,

$$r(\mu_t) = F_1\left(\int\int x\mu_t(dx, de), L\right) = F_1\left(\int x_t \, \mathrm{d}P, L\right) = \tilde{r}(x_t) \tag{37}$$

An identical arguments shows $\tilde{w}(x_t) = w(\mu_t)$. The second equality in (35) follows from the assumption of the proposition that $x_{t+1} = h_t(x_t, e_t)$. The final equality is true because $\int x_t \, \mathrm{d}P > 0$ and because the sequence $(x_t)_{t=0}^{\infty}$ satisfies (17) for each t.

Now suppose $\int\int x\mu_t(dx, de) = 0$, we have,

$$\int x \int \mu_t(dx, de) = \int x_t \, \mathrm{d}P = 0 \tag{38}$$

Since $(x_t)_{t=0}^{\infty}$ satisfies (17), $\int x_{t+1} \, \mathrm{d}P = 0$. Furthermore,

$$0 = \int x_{t+1} \, \mathrm{d}P = \int h_t(x_t, e_t) \, \mathrm{d}P = \int\int h_t(x, e) \, \mu_t(dx, de) \tag{39}$$

Thus satisfying the requirement of (10) when $\int\int x\mu_t(dx,de) = 0$.

Now we turn to the condition of Equation (11). Let $B \in \mathscr{B}(A \times E)$, where $B = B_A \times B_E$ for some $B \in \mathscr{B}(A)$ and $B_E \in \mathscr{B}(E)$. Use the definition of μ_{t+1} to write

$$\begin{aligned}
\mu_{t+1}(B) &= P\{x_{t+1} \in B_A, e_{t+1} \in B_E\} \\
&= P\{h_t(x_t, e_t) \in B_A, G(e_t, \eta_{t+1}) \in B_E\} \\
&= \mathbb{E}\left[\mathbb{1}_B\{h_t(x_t, e_t), G(e_t, \eta_{t+1})\}\right] \\
&= \mathbb{E}\left[\mathbb{1}_{B_A}\{h_t(x_t, e_t)\} \times \mathbb{1}_{B_E}\{G(e_t, \eta_{t+1})\}\right]
\end{aligned}$$

We have used the recursive formulation of the Markov process in the second line, Equation (30).

Note the joint distribution of e_t and x_t is μ_t and the distribution of the shock η_{t+1} is ψ. Furthermore, since η_{t+1} is independent of e_t and x_t,

$$\begin{aligned}
\mu_{t+1}(B) &= \mathbb{E}\left[\mathbb{1}_{B_A}\{h_t(x_t, e_t)\} \times \mathbb{1}_{B_E}\{G(e_t, \eta_{t+1})\}\right] \\
&= \int\int\int \mathbb{1}_{B_A}\{h_t(a, e)\} \\
&\quad \times \mathbb{1}_{B_E}\{G(e, \eta)\}\mu_t(da, de)\psi(d\eta) \\
&= \int\int \mathbb{1}_{B_A}\{h_t(a, e)\} \\
&\quad \times \left[\int \mathbb{1}_{B_E}\{G(e, \eta)\}\psi(d\eta)\right]\mu_t(da, de) \\
&= \int\int \mathbb{1}_{B_A}\{h_t(a, e)\}Q(e, B_E)\mu_t(da, de)
\end{aligned}$$

the final equality is the RHS of Equation (11). The second equality uses the recursive formulation of the Markov process (Equation (30)).

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

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

$$\sum_{t=0}^{\infty} \beta^t u(\mu_t, h_t) \geq \sum_{t=0}^{\infty} \beta^t u(\tilde{\mu}_t, \tilde{h}_t) \tag{40}$$

holds for any other sequence of feasible Borel probability measures and policy functions $(\tilde{\mu}_t, \tilde{h}_t)_{t=0}^{\infty}$.

As such, let $(\tilde{\mu}_t, \tilde{h}_t)_{t=0}^{\infty}$ be any other sequence of Borel probability measures on S and measurable policy functions $h_t \colon S \to A$ satisfying $\tilde{\mu}_0 = \mu_0$ and Equations (10) and (11). 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

$(Z, \mathscr{B}(Z), P)$. By Claim 5.1, $\tilde{\mu}_t$ is the distribution of $\{\tilde{x}_t, e_t\}$ for each t. Moreover, using an analogous argument to Claim 5.2, each $\tilde{x}_t$ will have finite variance and hence $\tilde{x}_t \in L^2(Z, P)$ for each t.

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

To check $(\tilde{x}_t)_{t=0}^{\infty}$ is feasible for the sequential problem, first, let us check whether $(\tilde{x}_t)_{t=0}^{\infty}$ is adapted to the filtration $(\mathscr{F}_t)_{t=0}^{\infty}$. We 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 5.3), $\tilde{x}_1$ will be $\sigma(x_0, e_0)$ measurable. Now suppose x_t is $\sigma(x_0, e_0, \ldots, e_{t-1})$ measurable. Consider

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

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 $\sigma(x_0, e_0, \ldots, e_t)$ measurable. By the principle of induction, each $\tilde{x}_{t+1}$ is then $\sigma(\tilde{x}_0, e_0, e_1, \ldots, e_t)$ measurable and $(\tilde{x}_t)_{t=0}^{\infty}$ is adapted to the filtration $(\mathscr{F}_t)_{t=0}^{\infty}$.

Now we turn to show the sequence $(\tilde{x}_t)_{t=0}^{\infty}$ satisfies the feasibility condition (17). Fix any $t \in \mathbb{N}$ and suppose $\int \tilde{x}_t \, \mathrm{d}P > 0$. We have

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

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 (10) for $\tilde{\mu}_t$ - almost everywhere.

On the other hand, suppose $\int \tilde{x}_t \, \mathrm{d}P = 0$. We have $\int \tilde{x}_t \, \mathrm{d}P = \int \int x \tilde{\mu}_t(dx, de) = 0$. Since $(\tilde{\mu}_t, \tilde{h}_t)_{t=0}^{\infty}$ satisfies (10), $\int \int \tilde{h}_t(x, e) \mu_t(dx, de) = 0$ will hold. We can now infer

$$\int \tilde{x}_{t+1} \, \mathrm{d}P = \int \tilde{h}_t(\tilde{x}_t, e_t) \, \mathrm{d}P = \int \tilde{h}_t(x, e) \, \tilde{\mu}_t(dx, de) = 0 \tag{41}$$

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, $(\tilde{x}_t)_{t=0}^{\infty}$ satisfies (17) for each t.

To complete the proof, note for each t

$$
\begin{aligned}
u(\mu_t, h_t) &= \int v((1+r(\mu_t))a + w(\mu_t)e - h_t(a,e))\mu_t(da, de) \\
&= \int v((1+\tilde{r}(x_t))x_t + \tilde{w}(x_t)e_t - h_t(x_t, e_t)) \,\mathrm{d}P \\
&= \int v((1+\tilde{r}(x_t))x_t + \tilde{w}(x_t)e_t - x_{t+1}) \,\mathrm{d}P \\
&= \rho_t(x_t, x_{t+1})
\end{aligned}
\tag{42}
$$

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

$$
\sum_{t=0}^{\infty} \beta^t u(\mu_t, h_t) = \sum_{t=0}^{\infty} \beta^t \rho_t(x_t, x_{t+1}) \geq \sum_{t=0}^{\infty} \beta^t \rho_t(\tilde{x}_t, \tilde{x}_{t+1}) = \sum_{t=0}^{\infty} \beta^t u(\tilde{\mu}_t, \tilde{h}_t) \tag{43}
$$

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, note the above implies $V(\mu_0) = \sum_{t=0}^{\infty} \beta^t u(\mu_t, h_t)$, and since, by (42), $\tilde{V}(x_0) = \sum_{t=0}^{\infty} \beta^t u(\mu_t, h_t)$, we can conclude $\tilde{V}(x_0) = V(\mu_0)$.

□

Appendix C: Proofs for section 3

Recall point-wise inequalities in $\mathbb{X}$ hold P - almost everywhere and convergence of (x^n) with $x^n \in \mathbb{X}$ for each n will be with respect to the weak topology.

Remark 5.1. The inequality $x \leq y$ holds P-almost everywhere if and only if for every $B \in \mathscr{B}(Z)$ such that $P(B) > 0$ we have $\int \mathbb{1}_B x \,\mathrm{d}P \leq \int \mathbb{1}_B y \,\mathrm{d}P$. To verify the claim, if $x \leq y$ holds almost everywhere, then for all $i \in B$ such that $P(B) > 0$ we have $x(i) \leq y(i)$, which will imply $\int \mathbb{1}_B x \,\mathrm{d}P \leq \int \mathbb{1}_B y \,\mathrm{d}P$. Conversely, let $\int \mathbb{1}_B x \,\mathrm{d}P \leq \int \mathbb{1}_B y \,\mathrm{d}P$ for all B such that $P(B) > 0$. Suppose by contradiction that $x(i) > y(i)$ for an B such that $P(B) > 0$. A contradiction since $\int \mathbb{1}_B x \,\mathrm{d}P > \int \mathbb{1}_B y \,\mathrm{d}P$.

Proof of proposition 3.3. Set t, and suppose $(x^n, y^n)_{n=0}^{\infty}$ satisfies $y^n \in \Gamma_t(x^n)$ for each n. Suppose $(x^n)_{n=0}^{\infty}$ converges to x and $(y^n)_{n=0}^{\infty}$ converges to y. We will show $y \in \Gamma_t(x)$.

Observe that $x^n \in \mathbb{S}_t$ for each n. It is straight forward to show $\mathbb{S}_t$ is sequentially closed, thus x must satisfy $\int x\, P \geq 0$. We have either $\int x \,\mathrm{d}P = 0$ or $\int x \,\mathrm{d}P > 0$. First let $\int x \,\mathrm{d}P > 0$ and note $\tilde{r}(x^n) = F_1(\tilde{K}(x^n), L)$ will converge, which means

$$
(1+\tilde{r}(x^n)) \int x^n h \,\mathrm{d}P + \tilde{w}(x^n) \int h e_t \,\mathrm{d}P \to (1+\tilde{r}(x)) \int x h \,\mathrm{d}P + \tilde{w}(x) \int h e_t \,\mathrm{d}P \tag{44}
$$

for any function h satisfying $h \in L^2(Z, P)$. In particular, for any B satisfying $B \in \mathscr{B}(Z)$ and $P(B) > 0$, letting $h = \mathbb{1}_B$, and noting $y^n \in \Gamma_t(x^n)$, by (17), we will have

$$\int \mathbb{1}_B y^n \, \mathrm{d}P \leq (1 + \tilde{r}(x^n)) \int \mathbb{1}_B x^n \, \mathrm{d}P + \tilde{w}(x^n) \int \mathbb{1}_B e_t \, \mathrm{d}P$$

Accordingly

$$\int \mathbb{1}_B y \, \mathrm{d}P \leq (1 + \tilde{r}(x)) \int \mathbb{1}_B x \, \mathrm{d}P + \tilde{w}(x) \int \mathbb{1}_B e_t \, \mathrm{d}P$$

establishing, by remark 5.1,

$$y \leq (1 + \tilde{r}(x))x + \tilde{w}(x)e_t$$

By an analogous argument, y will satisfy $y \geq \underline{a}$, allowing us to conclude $y \in \Gamma_t(x)$.

Now suppose $\int x^n \, \mathrm{d}P \to 0$. Note

$$\int y^n \, \mathrm{d}P \leq \int (1 + \tilde{r}(x^n))x^n + \tilde{w}(x^n)e_t \, \mathrm{d}P = F(\tilde{K}(x^n), L) + (1 - \delta)\tilde{K}(x^n)$$

The above equality follows from homogeneity of degree one of the production function (see Equation (45) below). Since $\tilde{K}(x^n) = \int x^n \, \mathrm{d}P \to 0$, we have $F(\tilde{K}(x^n), L) \to 0$ by Assumption 2.4, and

$$0 = \lim_{n \to \infty} \int y^n = \int y \, \mathrm{d}P$$

Allowing us to conclude $y \in \Gamma_t(x)$.

□

Lemma 5.2. *Consider the setting and notation of the sequential planner's problem in section 3.2. Fix a set C with $C \subset \mathbb{S}_0$, an $\epsilon > 0$ and $t \in \mathbb{N}$. If assumptions 2.1 - 2.4 hold, then there exists $\bar{r} \in \mathbb{R}_+$ such that for any $(x_i)_{i=0}^{\infty} \in \mathcal{G}^{\infty}(C)$ satisfying $\rho_t(x_t, x_{t+1}) \geq \epsilon$, we have $\tilde{r}(x_i) \leq \bar{r}$ for each $i \leq t$.*

Proof. Fix any $C \subset \mathbb{S}_0$, $\epsilon > 0$ and $t \in \mathbb{N}$ and select $(x_i)_{i=0}^{\infty}$ satisfying $(x_i)_{i=0}^{\infty} \in \mathcal{G}^{\infty}(C)$ and $\rho_t(x_t, x_{t+1}) \geq \epsilon$.

Since $x_i \in \Gamma_{i-1}(x_{i-1})$, by the feasibility correspondence (Equation (17)) and homogeneity of degree one (Assumption 2.4) of the production function F, we have

$$\begin{aligned} \tilde{K}(x_i) = \int x_i \, \mathrm{d}P &\leq \int (1 + \tilde{r}(x_{i-1}))x_{i-1} + \tilde{w}(x_{i-1})e_{i-1} \, \mathrm{d}P \\ &= (1 + F_1(\tilde{K}(x_{i-1}), L) - \delta)\tilde{K}(x_{i-1}) + F_2(\tilde{K}(x_{i-1}), L)L \\ &= F(\tilde{K}(x_{i-1}), L) + (1 - \delta)\tilde{K}(x_{i-1}) \end{aligned} \tag{45}$$

for each $i \in \mathbb{N}$.

Define $\hat{F}(K) := F(K, L) + (1-\delta)K$. By (45), we can write

$$\tilde{K}(x_i) \leq \hat{F}(\tilde{K}(x_{i-1})), \qquad i \in \mathbb{N} \tag{46}$$

Since $(\rho_t)_{t=0}^{\infty}$ is concave, from Jensen's inequality,

$$\begin{aligned} \epsilon \leq \rho_t(x_t, x_{t+1}) &= \int \nu\left((1+\tilde{r}(x_t))x_t + \tilde{w}(x_t)e_t - x_{t+1}\right) \, \mathrm{d}P \\ &\leq \nu\left(\int (1+\tilde{r}(x_t))x_t + \tilde{w}(x_t)e_t \, dP\right) \\ &= \nu(\hat{F}(\tilde{K}_t(x_t))) \end{aligned} \tag{47}$$

Note the inverse of v, v^{-1}, is also increasing since v is increasing. (The inverse of ν exists by Assumption 2.3). From Equation (47), $v^{-1}(\epsilon) \leq \hat{F}(\tilde{K}(x_t))$. And, by (46), since $\hat{F}$ is increasing, we arrive at

$$v^{-1}(\epsilon) \leq \hat{F}(\tilde{K}(x_t)) \leq \hat{F}^2(\tilde{K}(x_{t-1}))$$

Moreover, for any i satisfying $i \leq t$,

$$\nu^{-1}(\epsilon) \leq \hat{F}(\tilde{K}(x_t)) \leq F^{t-i+1}(\tilde{K}(x_i)) \tag{48}$$

Next, Let G^j denote the inverse of $\hat{F}^j$. Since $\hat{F}$ is strictly increasing, by (48), we have $\tilde{K}(x_i) \geq G^{t-i+1}(\nu^{-1}(\epsilon))$ for each $i \leq t$. Define $\underline{K} := \min_{i \leq t}\{G^{t-i+1}(\nu^{-1}(\epsilon))\}$, so $\tilde{K}(x_i) \geq \underline{K}$ for each $i \leq t$.

Finally, define $\bar{r} := F_1(\underline{K}, L) - \delta$. Note $F_1(K, L)$ is decreasing in the first argument since F is concave. We can conclude

$$\tilde{r}(x_i) = F_1(\tilde{K}(x_i), L) - \delta \leq F_1(\underline{K}, L) - \delta := \bar{r}, \qquad \forall\, i \leq t$$

Since $\bar{r}$ depends only on t and ϵ, the above will hold for any $(x_i)_{i=0}^{\infty} \in \mathcal{G}^{\infty}(C)$ satisfying $\rho_t(x_t, x_{t+1}) \geq \epsilon$. $\square$

Proof of proposition 3.4*.* Fix any sequentially compact set C satisfying $C \subset \mathbb{S}_0$, $\epsilon > 0$ and t. Suppose $(x_i)_{i=0}^{\infty} \in \mathcal{G}^{\infty}(C)$ satisfies $\rho(x_t, x_{t+1}) \geq \epsilon$. By Lemma 5.2, for any $i \leq t$, we know $r(x_i) \leq \bar{r}$. Since aggregate capital will be bounded from above, the maximum possible wage rate will be bounded above by a constant, which we now denote as $\bar{w}$. For all $i \leq t+1$, we have

$$\begin{aligned} \underline{a} \leq x_i &\leq (1+\bar{r})x_{i-1} + \bar{w}e_i \\ &\leq (1+\bar{r})^2 x_{i-2} + \bar{w}e_i + (1+\bar{r})\bar{w}e_{i-1} \\ &\leq (1+\bar{r})^i x_0 + \bar{w}\sum_{j=1}^{i}(1+\bar{r})^{i-j}e_{i-j} \end{aligned}$$

Let $W\colon = \bar{w}\sum_{j=1}^{i}(1+\bar{r})^{i-j}e_{i-j}$ to simplify notation and note $\|W\|$ will be finite. Next,

$$x_i^+ \leq (1+\bar{r})^i x_0 + W \implies (x_i^+)^2 \leq \Big((1+\bar{r})^i x_0 + W\Big)^2 \tag{49}$$

and

$$x_i^- \leq \underline{a} \implies (x_i^-)^2 \leq \underline{a}^2 \tag{50}$$

Since $x_0 \in C$ and C is compact, $\|x_0\| \leq M$ for some finite real number M. Now use $x_i = x_i^+ - x_i^-$ (see Stachurski (2009), section 7.2.3) and (49) and (50) to write

$$\begin{aligned}
\|x_i\| \leq \|x_i^+\| + \|x_i^{-1}\| &\leq \|(1+\bar{r})^i x_0 + W\| + \underline{a} \\
&\leq (1+\bar{r})^i \|x_0\| + \|W\| + \underline{a} \\
&\leq (1+\bar{r})^i M + \|W\| + \underline{a} \\
&: = \hat{M}_i \in \mathbb{R}
\end{aligned}$$

There will be such a $\hat{M}_i$ for each $i \leq t+1$. Moreover, the constants M_i will depend only on C, t and ϵ and so the above bound will hold for any feasible sequence $(x_i)_{i=0}^{\infty} \in \mathcal{G}(C)$ satisfying $\rho(x_t, x_{t+1}) \geq \epsilon$. □

Appendix D: Further Discussion on Dynamic Programming Limitations

I now consider the following topologies:

- The weak topology if we let $\mathbb{Y} = L^2(S)$ (weakly closed and norm-bounded sub-sets are compact)
- The weak topology if we let $\mathbb{Y} = L^1(S)$ (order intervals are weakly compact)
- The weak-star topology if we let $\mathbb{Y} = L^\infty(S)$ (weak-star closed and norm-bounded sub-sets will are compact)
- The weak topology if we let $\mathbb{Y} = \mathscr{C}b(S)$
- The weak- star topology inherited from $L^\infty(S)$ if we let $\mathbb{Y} = \mathscr{C}b(S)$ (weak-star closed and norm-bounded sub-sets will be compact)

Consider the weak topology on $L^2(S)$. If we let $\mathbb{M}$ be the space of Borel probability measures on S, then Φ will not be defined. To see why, simplify to a special case where we are only keeping track of a distribution on A and the policy function h maps A to A (this could be, for example, if the shocks $(e_t)_{t=0}^{\infty}$ are I.I.D.). We have

$$\mu_{t+1}(B) = \int \mathbf{1}_B\{h(a)\}\mu_t(da) \tag{51}$$

Now suppose $\mu_t = \delta_x$ is the Dirac delta measure which puts all weight on a point $x \in S$. We can then write

$$\mu_{t+1}(B) = \mathbf{1}_B\{h(x)\}$$

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

$$\mu_{t+1}(\mathbb{R}_{++}) + \mu_{t+1}(\mathbb{R}_{-}) = 0 \neq \mu_{t+1}(\mathbb{R}) = 1 \tag{52}$$

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 thus maps to values outside of $\mathbb{M}$.

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

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)).

There could be promise if we let $\mathbb{Y} = \mathscr{C}b(S) \subset L^\infty(S)$, where $L^\infty(S)$ has the weak-star topology. The map ϕ will be defined, and $\Gamma_t(\mu)$ could be compact-valued since it is weak-star closed and bounded. However, the map Φ will not be continuous. In particular, if we let $X = \mathcal{M}(S)$ with the topology of set-wise convergence, continuity will require

$$\int \mathbf{1}_B(h^n(a))\mu^n(da) \to \int \mathbf{1}_B(h(a))\mu(da), \qquad B \in \mathcal{B}(S) \tag{53}$$

To construct a counter-example, consider μ^n defined by $\mu^n(B) = \int_B sin(nx)dx$ for $B \in \mathcal{B}(S)$, and let $h^n(x) = sin(nx)$. We have that $h^n \to 0$ in the weak-star topology and $\mu^n \to \mu$ where $\mu(B) = 0$ for all $B \in \mathcal{B}(S)$. As such we have $\mu \circ h^{-1}(B) = 0$ for all $B \in \mathcal{B}(S)$ and $\int g \, d\mu \circ h^{-1} = 0$ for any $g \in \mathcal{C}(S)$. But, letting $g = 1$, we will have

$$\langle g, \mu_{t+1}^n \rangle = \int_S sin(nx)\mu^n(dx) = \int_S sin^2(nx)dx > 0 \tag{54}$$

We then conclude μ_{t+1}^n does not converge in the weak-star topology to μ_{t+1}. Since the weak-star topology is weaker than the topology of set-wise convergence, set-wise convergence is also ruled out.

References

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

Acikgoz, O., 2015. On the Existence and Uniqueness of Stationary Equilibrium in Bewley Economies with Production. Working Paper.
URL https://mpra.ub.uni-muenchen.de/71066/

Aiyagari, S. R., 1994. Uninsured Idiosyncratic Risk and Aggregate Saving. The Quarterly Journal of Economics 109 (3), 659–684.

Aiyagari, S. R., 1995. Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting. Journal of Political Economy 103 (6), 1158–1175.

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

Balder, E., 1987. Necessary and Sufficient Conditions for L1-Strong-Weak Lower Semicontinuity of Integral Functionals. Nonlinear Analysis, Theory, Methods & Applications 11 (12), 1399 – 1404.

Benhabib, J., Bisin, A., Zhu, S., 2015. The Wealth Distribution in Bewley Economies with Capital Income Risk. Journal of Economic Theory 159, 489–515.

Berger, D., Vavra, J., 2015. Consumption Dynamics During Recessions. Econometrica 83 (1), 101–154.

Bergin, J., Bernhardt, D., 1992. Anonymous Sequential Games with Aggregate Uncertainty. Journal of Mathematical Economics 21 (6), 543–562.

Berkovitz, L. D., 1974. Lower Semicontinuity of Integral Functionals. Transactions of the American Mathematical Society 192, 51–57.

Bewley, T., 1977. The Permanent Income Hypothesis: A Theoretical Formulation. Journal of Economic Theory 16, 252–292.

Bhandari, A., Evans, D., Sargent, T. J., 2017. Optimal Fiscal-Monetary Policy with Redistribution. Working Paper.
URL http://www.tomsargent.com/research/begs2.pdf

Boucekkine, R., Camacho, C., Zou, B., jan 2009. Bridging the Gap Between Growth Theory and the New Economic Geography: the Spatial Ramsey Model. Macroeconomic Dynamics 13 (01), 20.

Brock, W. A., Xepapadeas, A., Yannacopoulos, A. N., aug 2014. Optimal agglomerations in dynamic economics. Journal of Mathematical Economics 53, 1–15.

Brunnermeier, M. K., Sannikov, Y., 2016. On The Optimal Inflation Rate. American Economic Review 106 (5), 484–489.

Buera, B. F. J., Moll, B., 2015. Aggregate Implications of a Credit Crunch: The Importance of Heterogeneity. American Economic Journal. Macroeconomics 7 (3), 1–42.

Cao, D., 2016. Existence of Generalized Recursive Equilibrium in Krusell and Smith (1998). Working Paper.
URL `http://faculty.georgetown.edu/dc448/KSExistenceV17.pdf`

Chen, Y., Yang, C., 2017. Aiyagari Meets Ramsey: Optimal Capital Taxation with Incomplete Markets. Working Paper.
URL `https://dx.doi.org/10.20955/wp.2017.003`

Dávila, J., Hong, J. H., Krusell, P., Ríos-Rull, J.-V., 2012. Constrained Efficiency in the Neoclassical Growth Model With Uninsurable Idiosyncratic Shocks. Econometrica 80 (6), 2431–2467.

Fabbri, G., Faggian, S., Freni, G., 2015. On the mitra-wan forest management problem in continuous time. Journal of Economic Theory 157, 1001–1040.

Feinberg, E. A., Kasyanov, P. O., Zadoianchuk, N. V., 2012. Average Cost Markov Decision Processes with Weakly Continuous Transition Probabilities. Mathematics of Operations Research 37 (4), 591–607.

Feinberg, E. A., Kasyanovb, P. O., Zadoianchukb, N. V., 2013. Berge's Theorem for Non-Compact Image Sets. Journal of Mathematical Analysis and Applications 397 (1), 255–259.

Genicot, G., Ray, D., 2015. Aspirations and Inequality 85 (2), 1–35.

Hernández-Lerma, O., 1989. Adaptive Markov Control Processes. Springer-Verlag New York.

Hernández-Lerma, O., Lasserre, J. B., 1996. Discrete-Time Markov Control Processes. Springer-Verlag New York.

Hopenhayn, H. A., 1992. Entry, Exit, and firm Dynamics in Long Run Equilibrium. Econometrica 60 (5), 1127–1150.

Huggett, M., 1993. The Risk-Free Rate in Heterogeneous-Agent Incomplete-Insurance Economies. Journal of Economic Dynamics and Control 17, 953–969.

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

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

Kaplan, G., Moll, B., Violante, G. L., 2016. Monetary Policy According to HANK. Working Paper, 1–61.

Kaplan, G., Violante, G. L., 2010. How Much Consumption Insurance Beyond Self-Insurance? American Economic Journal: Macroeconomics 2 (2), 53–87.

Khan, A., Thomas, J. K., 2013. Credit Shocks and Aggregate Fluctuations in an Economy with Production Heterogeneity. Journal of Political Economy 121 (6), 1055–1107.

Krusell, P., Mukoyama, T., Smith, A. A., 2011. Asset prices in a Huggett economy. Journal of Economic Theory 146 (3), 812–844.

Krusell, P., Smith, A. A., 1998. Income and Wealth Heterogeneity in the Macroeconomy. Journal of Political Economy 106 (5), 867–896.

Kuhn, M., 2013. Recursive Equilibriua in an Aiyagari-Style Economy with Permanent Income Shocks. International Economic Review 54 (3), 807–835.

Ljungqvist, L., Sargent, T. J., 2004. Recursive Macroeconomic Theory. MIT Press.
URL `https://books.google.co.uk/books?id=Xx-j-tYaPQUC`

Mckay, A., Reis, R., 2016. The Role of Automatic Stabilizers in the U.S. Business Cycle. Econometrica 84 (1), 141–194.

Melitz, M. J., 2003. The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity. Econometrica 71 (6), 1695–1725.

Miao, J., 2002. Stationary Equilibria of Economies with a Continuum of Heterogeneous Consumers. Working Paper.
URL `http://people.bu.edu/miaoj/shockid48.pdf`

Miao, J., 2006. Competitive equilibria of economies with a continuum of consumers and aggregate shocks. Journal of Economic Theory 128 (1), 274–298.

Nuño, G., Moll, B., 2015. Controlling a Distribution of Heterogeneous Agents. Working Paper at http://www.princeton.edu/˜moll/research.htm, 1–37.

Nuno, G., Thomas, C., 2017. Monetary Policy with Heterogeneous Agents. Working Paper.
URL `http://ideas.repec.org/p/red/sed013/356.html`

Ray, D., Robson, A., 2012. Status, Intertemporal Choice, and Risk-Taking. Econometrica 80 (4), 1505–1531.

Sampson, T., 2016. Dynamic Selection and the New Gains from Trade with Heterogeneous Firms. Quarterly Journal of Economics 131 (1), 315–380.

Shanker, A., 2017a. Controlling a Distribution of Heterogenous Agents. Ph.D. thesis.

Shanker, A., 2017b. Maximum Theorem for Non-Compact Dynamic Optimisation. Working Paper.

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

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

Uhlig, H., 1996. A Law of Large Numbers for Large Economies. Economic Theory 8 (1), 41–50.

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