Bewley-Huggett-Aiyagari Models: Computation and Simulation

Robert Kirkby*

March 27, 2015

Abstract

I provide conditions under which an algorithm for computation and simulation of Bewley-Huggett-Aiyagari models based on state-space discretization will converge to the true solution. These conditions are shown to be satisfied in two standard examples: Aiyagari (1994) and Pijoan-Mas (2006). The algorithm is implemented for the two models and used to create solutions whose convergence to the true solution is known. Bewley-Huggett-Aiyagari models are general equilibrium models with incomplete markets and idiosyncratic, but no aggregate, shocks. The algorithm itself is based on discretization while the theory importantly allows for the fact that the simulations will be made using the approximate computational solution of the value function problem, rather than the true model solution. Numerical results from applying the algorithm to both models are given and investigated, both in terms of replication and looking at inequality. Matlab codes implementing the algorithm for both models using parallelization on the GPU are provided.

Keywords: Bewley-Huggett-Aiyagari models, Numerical methods, Convergence.

JEL Classification: E00; C68; C63; C62; C15

---

*Kirkby: Victoria University of Wellington. We acknowledge the useful comments of seminar participants at the Universidad Carlos III de Madrid and Australian National University. We thank Manuel Santos and John Stachurski for helpful discussion. Please address the correspondence about this article to Robert Kirkby at <robertdkirkby@gmail.com>, robertdkirkby.com.

Contents

1 Introduction

Recent concerns about inequality has increased interest in models with many different households, known as heterogeneous agent models. Bewley-Huggett-Aiyagari models are one of the main categories of such heterogeneous agent models; they are characterized by idiosyncratic (but no aggregate) shocks, incomplete markets, and general equilibrium. With the Income and Wealth shares of the Top 1% being a topic of particular public interest an important aspect of these models is their able to capture this inequality (Castaneda et al., 2003). Being able to model inequality is an key step to understanding it's possible causes (Benhabib et al., 2011) and the effects to taxation aimed at reducing this inequality (Kinderman and Krueger, 2014).

Solving Bewley-Huggett-Aiyagari models and using them to evaluate the effects of policies is almost always done numerically on the computer. This reflects both an interest in quantitative answers, and the difficulty of establishing any analytical results without strong assumptions on the parameterization of distributions, shocks, and functional forms. An important question therefore becomes: Do we know that the computer is giving us the correct solution? This paper introduces an algorithm based on discretization and provides theory showing that under certain assumptions this algorithm will converge to the true solution of the model. It is then shown that these conditions are satisfied by two workhorse models from the literature, Aiyagari (1994) and Pijoan-Mas (2006). Quantitative results from applying the algorithm to these two models are then given.

The Bewley-Huggett-Aiyagari class of models has its beginnings in Bewley (1983) with early quantitative explorations being Huggett (1993) and Aiyagari (1994). Heterogeneous agent models of this class have been used to give quantitative answers to questions on topics as varied as: progressive taxation (Conesa and Krueger, 2006), capital taxation (Domeij and Heathcote, 2004; Conesa, Kitao, and Krueger, 2009), inequality (Castaneda, Díaz-Giménez, and Ríos-Rull, 2003), entrepreneurship (Quadrini, 2000), and working longer hours (Pijoan-Mas, 2006). For more on the applications of heterogeneous agent models, both of the Bewley-Huggett-Aiyagari class and other more-complicated classes, see Ríos-Rull (1995, 2001); Heathcote, Storesletten, and Violante (2009). There are few qualitative results on these models largely because doing so involves imposing strong assumptions either on shock distributions (Benhabib et al., 2011) or on functional forms (Huggett, 2003). Part of the appeal of using Bewley-Huggett-Aiyagari models to look at inequality comes from their naturally incorporating some of the multi-dimensional aspects of inequality, such as the interrelations between earnings, income, and wealth (Díaz-Giménez et al., 1997, 2011).

Both the algorithm and the related theory for solving Bewley-Huggett-Aiyagari models invole a number of steps: (i) a value function problem must be solved to get the optimal policy function; (ii) the optimal policy function, together with the exogenous shock process, are used to find the steady-state agents distribution; (iii) some moments of the steady-state agent distribution are calculated; and lastly (iv) steps (i)-(iii) must be repeated for different prices until a general equilibrium is found.

The main challenges in proving that a certain algorithm will converge is that the simulations themselves are typically made using a computational approximation to the solution to the value function problem. So numerical errors that occur while computing the solution to the value function may cause problems with the convergence of the simulations themselves. This issue was addressed in Santos and Peralta-Alva (2005) based on the Feller property. But applying it to the simulations depended on a Stochastic Contraction Condition, rather than the Monotone Mixing Condition that underlies the theory on the Bewley-Huggett-Aiyagari class of models. The theory developed here takes the approach of variously using both the Feller property (but not Stochastic Contractions) and the Montone Mixing Condition depending on the model property being investigated at the time. This combination of the two approaches is key to deriving the theoretical results.

When deriving the theory emphasis is placed on ensuring that the conditions on which it is based can be proven to apply to Bewley-Huggett-Aiyagari models. The importance of using conditions that economic models can be proven to satisfy has been an important component of much recent theoretical work (Hopenhayn and Prescott, 1992; Santos and Peralta-Alva, 2005; Kamihigashi and Stachurski, 2014, 2015). The conditions underlying the theory of this paper are shown to be satisfied for two workhorse models of the the Bewley-Huggett-Aiyagari class, namely those of Aiyagari (1994) and Pijoan-Mas (2006). Having shown that the theory applies we know that the algorithm will converge to the true solution. Some quantitative results from applying the algorithm to these models are then given and discussed.

Having a numerical method which is known to converge to the true solution has further uses beyond simply using it to solve Bewley-Huggett-Aiyagari models. Interest in inequality and the Top 1% is a major application of these and similar models (Castaneda et al., 2003; Kinderman and Krueger, 2014; Badel and Huggett, 2014; Guner et al., 2014). Given that models statistics about the the Top 1% are likely to be highly sensitive to numerical approximation errors it becomes even more important to know that the algorithms being used to solve the models are converging to the true solution. Beyond the ambit of this paper it also suggests that solving such models using methods which minimize the degree of numerical approximation, even at the cost of slower run times, may be important to getting reliable numbers on, eg., the income share of the top percentiles. A partial summary of the literature on numerical errors can be found in Peralta-Alva and Santos (2014), while a practical example of how numerical errors can be problematic for certain model statistics is given in Hatchondo, Martinez, and Sapriza (2010). Another advantage of having a numerical method which is known to converge to the true solution is that this can then be used to evaluate alternative numerical methods for speed and accuracy. Since analytical solutions to Bewley-Huggett-Aiyagari models are not possible, having a numerical solution created by algorithms known to converge to the true solution and that we might treat as if it were the true solution is key to judging the accuracy of alternative numerical methods (Aruoba et al., 2006; Caldara et al., 2012).

Section 2 provides a general description of the model. Section 3 describes the algorithm that is used. Section 4 provides theory addressing numerical errors that occur during the computational solution and simulation of the model. Section 5 shows that this theory is directly applicable to the models of Aiyagari (1994) and Pijoan-Mas (2006). An investigation of the numerical results of implementing the algorithm are then given in Section 6.

2 Bewley-Huggett-Aiygari Models

In this section I give a formal description of the class of models being considered here; namely general equilibrium heterogeneous agent models with incomplete markets and idiosyncratic, but no aggregate, uncertainty. Notation is loosely based on Stokey, Lucas, and Prescott (1989, henceforth SLP); loosely as SLP do not treat heterogeneous agent models of this type.

The models are those which can be expressed as follows: Let $X \subseteq \mathbb{R}^l$ be the endogenous state variable, $Y = Y_1 \times Y_2 \subseteq X \times \mathbb{R}^c$ be the choice variable, and $Z \subseteq \mathbb{R}^k$ be the exogenous state variable. Let $\Theta \subseteq \mathbb{R}^{q_1} \times \mathbb{R}^{q_2}$ be a parameter space; The state of a agent is then a pair (x, z). A value function maps $V_p : X \times Z \to \mathbb{R}$. A policy function maps $g_p : X \times Z \to Y$. Let $S = X \times Z$, and let $\mathcal{S}$ be it's Borel σ-field. The measure of agents μ_p is a probability distribution over $(S, \mathcal{S})$. The return function maps $F_p : X \times Y \times Z \to \mathbb{R}$, and the discount factor is $0 < \beta < 1$.

Aggregate variables are $A \in \mathbb{A} \subseteq \mathbb{R}^a$. A price vector is $p \in \mathbb{P} \subseteq \mathbb{R}^p$. The exogenous shock follows a Markov-chain with transition function Q mapping from Z to Z. The aggregation function maps $\mathcal{A} : \mathcal{M}(S, \mathcal{S}) \to \mathbb{R}^a$, where $\mathcal{M}(S, \mathcal{S})$ is the space of probability measures on $(S, \mathcal{S})$. The market clearance function maps $\lambda : \mathbb{R}^a \times \mathbb{R}^p \to \mathbb{R}$.

Definition 1. *A Competitive Equilibrium is an agents value function V_p; agents policy function g; vector of prices p; measure of agents μ_p; such that*

1. *Given prices p, the agents value function V_p and policy function g_p solve the agents problem*

$$V_p(x, z) = \max_{y=(y_1,y_2)\in Y} \left\{ F_p(x, y, z) + \beta \int V_p(y_1', z') Q(z, dz') \right\} \tag{1}$$

2. *Aggregates are determined by individual actions: $A_p = \mathcal{A}(\mu_p)$.*

3. *Markets clear (in terms of prices): $\lambda(A_p, p) = 0$.*

4. *The measure of agents is invariant:*

$$\mu_p(x, z) = \int \int \left[\int 1_{x=g_p^{y_1}(\hat{x},z)} \mu_p(\hat{x}, z) Q(z, dz') \right] d\hat{x} dz \tag{2}$$

Models fitting this definition include Huggett (1993) and Aiyagari (1994), as well as numerous extensions endogenizing labour supply, introducing taxation, and modeling dynasties. The aggregates in point two generally correspond to the household variables (such as aggregate capital) but in some models may also be aggregates of functions thereof (such as tax revenue). The third point, that prices clear markets, involves rewriting market clearance equations in terms of prices, rather than quantities.

So for example in Aiyagari (1994) — a model with which many readers will be familiar, and a full description of which is given in Section 5 — the requirement that aggregates are determined by individual actions is that aggregate capital is the sum of individuals capital holdings, $K = \int k'(x,z)d\mu$. While the market clearance condition is that the interest rate is equal to the marginal product of capital, $\lambda(K,r) = r - \alpha K^{1-\alpha} - \delta = 0$, where K depends on individual behaivour.

3 The Algorithm

The results and theory regarding the solution of Bewley-Huggett-Aiyagari models, in particular those of Aiyagari (1994) and Pijoan-Mas (2006), are all based on an algorithm implementing pure discretization of the state-space. This algorithm, while not as sophisticated as many others, is both widely used and has the strengths of being both highly robust and easy to implement. Discretization also provides a good benchmark against which other algorithms can be compared (Aruoba, Fernandez-Villaverde, and Rubio-Ramirez, 2006) and has the advantage of being easily parallelized, including on the graphics processor (Aldrich, Fernandez-Villaverde, Gallant, and Rubio-Ramirez, 2011). The solution of Bewley-Huggett-Aiyagari models involves three steps: (i) computation of the optimal policy function, (ii) simulation of the steady-state distribution of the agents (or at least some of it's moments), and (iii) finding the general equilibrium. I now briefly describe how pure discretization of the state-space works in each of these three steps.

(i) Discretized Value Function Iteration and the Optimal Policy Function:
The algorithm for discretized value function iteration is largely standard. A version incorporating Howards improvement is now given.[1]

Declare initial value V_0.
Declare iteration count $n = 0$.
while $||V_{n+1} - V_n||$ **do**
    Increment n. Let $V_{old} = V_{n-1}$.
    for $x = 1, ..., n_x$ **do**
        for $z = 1, ..., n_z$ **do**
            Calculate $E[V_{old}|z]$ ▷ Using quadrature method such as Tauschen method.
            Calculate $V_n(x,z) = \max_{y=1,...,n_y} F(x,y,z) + \beta E[V_{old}(y,z')|z]$

[1] The convergence results used later specifically allow for the use of Howards improvement.

Calculate $g(x,z) = \arg\max_{y=1,...,n_y} F(x,y,z) + \beta E[V_{old}(y,z')|z]$
end for
end for
for $i = 1, ..., H$ **do** ▷ Howards Improvement Algorithm (do H updates)
$V_{old} = V_n$
for $x = 1, ..., n_x$ **do**
for $z = 1, ..., n_z$ **do**
Calculate $V_n(x,z) = F(x, g(x,z), z) + \beta E[V_{old}(g(x,z), z')|z]$
end for
end for
end for
end while
for $x = 1, ..., n_x$ **do**
for $z = 1, ..., n_z$ **do**
Calculate $g_n(x,z) = \arg\max_{y=1,...,n_y} F(x,y,z) + \beta E[V_n(y,z')|z]$
end for
end for

This algorithm implements (pure) discretized Value Function Iteration.

Many alternatives to this algorithm exist for solving value function iteration, involving various degrees of discretization. For example here z is fully discretized (using, eg., quadrature methods such as the Tauchen method (Tauchen, 1986)); an alternative would be to use Monte-carlo integration methods (eg. Pál and Stachurski (2013)). The choice of next periods state is also discretized, but partial discretization would be an alternative (Aruoba, Fernandez-Villaverde, and Rubio-Ramirez, 2006; Santos and Vigo-Aguiar, 1998). Instead of discretizing the value function itself one can used fitted value function methods (Cai and Judd, 2014; Benitez-Silva, Hall, Hitsch, Pauletto, and Rust, 2005). Other related algorithms for value function iteration are the endogenous grid method (Carroll, 2006; Barillas and Fernandez-Villaverde, 2007; Fella, 2014), and envelope condition method (Maliar and Maliar, 2013).

(ii) Iteration on Discretized Agents Distribution:

The algorithm for iterating on the discretized agent distribution is now given:

Declare initial distribution μ_0. ▷ A matrix the elements of which sum to one.
Declare iteration count $n = 0$.
while $||\mu_{n+1} - \mu_n||$ **do**
Increment n. $\mu_n = 0$ ▷ A matrix of zeros on x-by-z
for $x = 1, ..., n_x$ **do**
for $z = 1, ..., n_z$ **do**
for $z' = 1, ..., n_z$ **do** $\mu_n(x(g(x,z)), z') = \mu_n(x(g(x,z)), z') + (g(x,z) * Q(z,z')) * \mu_{n-1}(x,z)$

end for

end for

end for

end while

The algorithm consists of a while-loop for convergence of the agents distributions, together with iterative updating of the agents distribution based on the formula $\mu_n(x,z) = \int\int\left[\int 1_{x=g(\hat{x},z)}\mu_{n-1}(\hat{x},z)Q(z,dz')\right]d\hat{x}dz$. This algorithm is simple to implement as in practice it the agents distribution is just a matrix of numbers (probability masses) at each point on the grid. Since only values on the grid can be taken it helps later in developing our convergence results when combined with the optimal policy function which, computed by discretized VFI, also only takes values on the grid. In practice it is often a good idea to create the starting point for the agents distribution by simulation methods.

The main alternative to iterating on the agents distribution is to use simulated paths to generate an approximation of the distribution. Both Santos and Peralta-Alva (2005) and Kamihigashi and Stachurski (2015) provide relevant results on the convergence of simulated path methods, but the first involves a condition that would be difficult to show for the models at hand[2] while the second addresses the issue of convergence of simulated paths using the true optimal policy function (rather than the optimal policy function we get from computational solution of the value function iteration). While generally not as accurate as iterating on the discretized agents distribution these alternatives are more widely used as they are both faster and less memory intensive.[3]

(iii) Calculate Aggregate Variables and Evaluate Market Clearance:
Calculating each of aggregate variables is trivial and involves evaluating a function on the grid and then taking weighted sum using the agents distribution as the weights. Evaluating market clearance is then a single function evaluation.[4]

Evaluate $A = \sum_{x=1,...,n_x,\, z=1,...,n_z} A(x,z)\mu_n(x,z)$
Evaluate $\lambda(A,p)$

(iv) Find a/the General Equilibrium using a Grid on Prices:
The algorithm is,

Declare grid for prices $p = 1,...,n_p$
for $p = 1,...,n_p$ **do**
Given p, solve Step (i) to get value fn and optimal policy function.
Given p, solve Step (ii) to get agents distribution.

[2]Namely, that the combination of the optimal policy with the exogenous shocks gives a Stochastic Contraction; their Condition C.

[3]Based on personal experience my practical recommendation would be the use of iterating on the discretized agents distribution when you care about certain sensitive statistics such as the share of wealth of the top percentile (eg. Castaneda, Díaz-Giménez, and Ríos-Rull (2003)), but using simulated path methods when the interest is just in first and second moments.

[4]In a model with more than one market the function will be vector valued.

Given p, solve Step (iii) to get aggregate variables and the market clearance.

end for

Evaluate every pair of two consecutive grid points on the price grid, if the market clearance conditions switch sign (eg. from negative to positive), then whichever of those two prices has to lowest absolute value of the market clearance condition is an equilibrium price. [For simplicity this has been written assuming price is one-dimensional, for an n-dimensional price the consecutive grid points would be a consecutive 2^n neighbourhood of grid points.]

Conceptually, this step is about finding which price(s) satisfy the general equilibrium requirement. From a theoretical perspective we know that there is a least one equilibrium, and we will make an assumption that ensures there are at most a finite number of equilibria.[5] Algorithmically this step will be done by placing a grid on prices and simply evaluating the model at each point on the grid. While this is slower than the more commonly used algorithms (function minimization algorithms, such as Nelson-Raphson, or simplex-search methods) it allows us to prove that the algorithm will converge for any and all equilibria.

4 Theory on Computational Solution and Simulation

In this section we address the issue of numerical errors that might arise during the computation and simulation of the model. There are a variety of possible sources of error and of dimensions of convergence to be considered. First, we have solving the value function and optimal policy function (for a given price), here we can simply appeal to existing results on uniform convergence of the optimal policy function under discretized value function iteration. Second, we have the agents distribution, there are two aspects to this: that iterating on the agents distribution will lead it to converge to the true agents distribution (for a given price), and further that this remains true even though we are using an approximation of the optimal policy (that we got from discretized value function iteration) rather than the true optimal policy. Third, we calculate the aggregate variables and assess the market clearance equations, this step is largely trivial. Fourth, we have the price dimension where we need to show that the algorithm will converge to any and all equilibria of the model.

Results are provided, firstly to show that numerical errors in the value function and optimal policy function are bounded, and that they will go to zero as the distance between points on the grids used to approximate the value function and optimal policy function go to zero. It is then shown that if, as was shown, the numerical errors in the optimal policy function are bounded, then numerical errors in the steady-state distribution will also be bounded. It follows almost trivially that numerical errors in the moments of the steady-state distribution are bounded. Furthermore, as numerical errors in the optimal policy function go to zero, so will those in the steady-state

[5] This assumption is the only assumption that cannot be proved from first principles. See Section 4.4.

distribution and it's moments. Together these results ensure that the any role played by numerical errors will go to zero as the distance between points on the grid used in the approximation go to zero.

In what follows we consider in turn: computation of the value function and optimal policy function; simulation of the steady state distribution; computation of moments of the steady state distribution and the market clearance; finding the equilibria. At every stage emphasis is placed on deriving convergence results that are based on the specific algorithm that will be actually used to compute and solve the model.

4.1 Computing the Value Function and Optimal Policy Function

The first step in solving these models is to compute the (value function and the) optimal policy fn. A standard approach, especially in more complicated examples, is value function iteration.[6] We will concentrate here on pure discretized value function iteration.[7]

Various theorems bounding the numerical errors in value function iteration exist: for example Santos and Vigo-Aguiar (1998) provide results based on partial discretization, while Stachurski (2008) provides results for a variety of fitted value function iteration methods. Here we will use the results of Kirkby (2015) which specifically cover discretized value function iteration allowing for non-interior choices, and provide results on the optimal policy function, both of which are important in the present application.

Our results are based on the Case 1 value function problem,

$$V(x,z) = \sup_{y=(y_1,y_2)\in\Gamma(x,z)} \left\{ F(x,y,z) + \beta \int V(y_1,z')Q(z,dz') \right\} \tag{3}$$

Indirectly we are interested in the convergence of the solution to the this problem, V, and the numerical solution to the discretized problem, V_N^G; the later is the solution after N iterations (once a standard convergence criterion is met) of value function iteration on the discrete grid (hence the G). More specifically though our interest is in covergence of the discretized optimal policy function g_N^G — the optimal policy associated with V_N^G and taking values only on the grid — to the true optimal policy function g. Our interest in the convergence of the optimal policy reflects that it is the optimal policy that then used to find the stationary agents distribution.

[6] Other methods such as projection methods can be used, but value function iteration has the advantage of being applicable to more complex examples such as those incorporating fixed-costs of adjustment, discrete choices, and even non-concave return functions (needed, eg., to model that US payroll taxes are only paid up to a threshold level).

[7] While more advanced methods, such as fitted value function (Cai and Judd, 2014) are available, pure discretization remains widely used due to it's ease of implementation and robustness. The use of graphics processors has granted discretized VFI a new lease of life in practical applications (Aldrich, Fernandez-Villaverde, Gallant, and Rubio-Ramirez, 2011). Discretized VFI also be useful in creating a starting point for other methods based on value function iteration that require a decent initial guess, such as the Envelope Condition Method (Maliar and Maliar, 2013).

We assume that the spaces for the endogenous state X, the control variables Y, and the exogenous state Z are all compact, that the return function F is monotone, continuous and bounded[8], the discount factor β is less than one, and that the transition function Q has the Feller property.

Then we have the following result (Kirkby (2015), Propostion 1),

Lemma 1. *For the value function defined in 3. Let X, Y, Z be compact, the return function F be continuous, monotone, and bounded, the discount factor β be less than one, and the transition function Q has the Feller property. Let V_N^G be the numerical solution to the discretized value function problem (discretizing X, Y and Z, using Tauchen method to discretize Q). Then, the numerical errors in the value function, $||V - V_N^G||$, and in the optimal policy function, $||g - g_N^G||$, converge to zero as the distance between grid points in the dimensions being discretized go to zero.*

This result gives us uniform convergence of both the value function and the optimal policy function (for a given price). The assumptions of 1 are satisfied by a wide range of models and it will be easy to show that they are satisfied in our applications.

4.2 Simulating the Steady-State Agents Distribution

We turn now to the steady-state agents distribution. Focus is on the issue of uniform convergence of the computed steady-state to the true steady-state; for a given price. A number of theoretical issues must be addressed and we will need to call on a number of results from the literature. Our first step will be establishing that, for a given price, there exists an stationary distribution and that it is unique; note that this is not related to it's being or not a general equilibrium which would further require market clearance. For this we will have recourse to results from Hopenhayn and Prescott (1992). We then turn to the computation of the steady-state distribution which has two aspects, firstly that the computed steady-state distribution converges to the true steady-state distribution — for which we use further results from Bhattacharya and Lee (1988) and Bhattacharya and Majumdar (2001) — and secondly that this will continue to hold even though the computed steady-state distribution is calculated based on the computed optimal policy, rather than the true optimal policy — for which we will use results from Santos and Peralta-Alva (2005). In doing this we will be appealing in turn to two of the properties of markov chains most useful in Macroeconomic models, monotone order-mixing properties and the Feller property.[9]

Following SLP we use P to denote the transition function on $(S, \mathcal{S}) = (X \times Z, \mathcal{X} \times \mathcal{Z})$ resulting

[8]Since F is a continous function defined on a compact space, it will therefore also be bounded.

[9]Specific use is made of the Monotone Mixing Condition of Hopenhayn and Prescott (1992). Related alternatives include splitting conditions (Bhattacharya and Lee, 1988; Bhattacharya and Majumdar, 2001). Kamihigashi and Stachurski (2014) provide a monotone order-mixing condition which incorporates all of the aforementioned as subcases. The decision to use the Monotone Mixing Condition, rather than it's generalizations the main innovation of which are to allow for non-compact state spaces is in part based on the fact that our algorithm's use of (pure) discretization of the state-space means that bounding numerical errors — and hence uniform convergence — for both the optimal policy function and for the agent's distribution anyway necessitate the existence of a compact state-space.

from the combination of the (optimal) policy function[10] $g : X \times Z \to X$, and the transition function Q on $(Z, \mathcal{Z})$. Our interest is then in the existence and uniqueness of a probability distribution μ over $S = \mathcal{X} \times \mathcal{Z}$ which is a stationary distribution for P, ie. $P\mu = \mu$, and on convergence to this distribution. We begin with some results from Hopenhayn and Prescott (1992) which can be used for proving the existence of a unique equilibria. Let $\mathcal{M} = \mathcal{P}(S)$, the set of probability measures on S.

4.2.1 Existence, Uniqueness of, and Convergence to, an Invariant Distribution

The monotone mixing condition is introduced, which, when combined with a requirement trivally satisfied by compact subspaces of $\mathcal{R}^n$ needed for the existence of invariant distributions, gives us results for uniqueness and global convergence. This result is Theorem 2 of Hopenhayn and Prescott (1992)

Theorem 1. *Suppose P is increasing, S contains a lower bound (which we will denote by a) and an upper bound (which we will denote by b) and the following condition is satisfied: Monotone Mixing Condition (MMC): there exists $s^* \in S$,$m \in \mathbb{Z}$, $\beta_P > 0$ such that $P^m(b, [a, s^*]) \geq 1 - \beta_P$ and $P^m(a, [s^*, b]) \geq 1 - \beta_P$. Then there is a unique stationary distribution μ^* for process P, and for any initial measure μ_0, $\mu_n \equiv T^{*n}\mu_0 = \int P^n(s, \cdot)\mu_0(ds)$ converges to μ^*.*

The following result, adapted from Bhattacharya and Lee (1988; Lemma 2.3) and Bhattacharya and Majumdar (2001) shows how fast the sequence $\mu_n = T^{*n}\mu_0$ converges to it's invariant distribution μ^*. First we define a distance d_1 stronger than that we have been using till now[11]. For $a \geq 0$, let $\mathcal{G}_a$ denote the class of all real-valued Borel measurable nondecreasing functions f on S satisfying $0 \leq f(s) \leq a, \forall s \in S$. Define

$$d_a(\mu, \upsilon) = \sup\{|\int f d\mu - \int f d\upsilon| : f \in \mathcal{G}_a\}, \quad \mu, \upsilon \in \mathcal{P}(S)$$

Proposition 1. *Under the assumptions of Theorem 1 we have further that*

$$d_1(T^{*n}\mu, T^{*n}\upsilon) \leq \beta_P^{[n/m]} d_1(\mu, \upsilon), \quad \forall \mu, \upsilon \in \mathcal{P}(S)$$

and

$$d(T^{*n}\mu_0, \mu^*) \leq \beta_P^{[n/m]} \quad \forall \mu_0 \in \mathcal{P}(S) \tag{4}$$

Remark: The later result of this proposition is expressed in the Kolmogorov metric and is essentially just the first result of the propostition rewritten to take advantage of the fact that

[10]Or more accurately the restriction g_{y_1} of the optimal policy $g : X \times Z \to Y = (Y_1, Y_2)$ onto $Y_1 = X$; $g_{y_1} : X \times Z \to Y_1 = X$.

[11]Till now we used the Kolmogorov measure $d(\mu, \upsilon) = \sup\{|F_\mu(A) - F_\upsilon(A)| : A \in \mathcal{S}\}$, $\mu, \upsilon \in \mathcal{P}(S)$, where F_μ is the distribution function of μ. It is noted that convergence in the Kolmogorov metric implies weak convergence in the power set of S.

$d(\mu, v) \leq 1$ follows directely from the definition of Kolmogorov metric as based on the conditional distribution functions.

Bounding the distance between the iterated agents distribution and the discretized agents distribution is then just a well-known property of contraction mappings,

Corollary 1. *Under the assumptions of Theorem 1. Let $\{\mu_n\}$ be the sequence defined by repeated application of P, ie. $\mu_n = P\mu_{n-1}$, $\mu_1 = P\mu_0$, from an arbirary starting probability measure $\mu_0 \in \mathcal{S}$. Let $d(\mu_N, \mu_{N+m}) < \epsilon$, then*

$$d(\mu_N, \mu^*) \equiv\leq \frac{1}{1-\beta_P}\epsilon$$

Remark: Convergence in the d_1 metric is stronger than convergence in moments. Since in heterogeneous agents what we often care about is the aggregate value (which is just the first moment) we could use analogues of this result for convergence in moments.

4.2.2 The Agents Distribution with the Approximation of the Optimal Policy

In the previous subsection we had a result covering convergence of the agents distribution to the true steady-state agents distribution if we were to calculate it using the true optimal policy function. We now turn to address the issue of convergence even though we are iterating on the agents distribution not with the true optimal policy function but instead using the computated optimal policy function. For this we turn some results from Santos and Peralta-Alva (2005) based on the Feller property: loosely speaking, their Theorem 2 shows that the invariant distribution is continuous in the policy function (technically, that the invariant distribution correspondence is upper semicontinuous).

Following Santos and Peralta-Alva (2005) and Stenflo (2001) the theory of this section is developed using the iterated function systems notation. Up till now the solution to the individuals problem has been given by the optimal policy function[12] $g(x, z) : X \times Z \to X$, which is then combined with the transition function Q on $(Z, \mathcal{Z})$ to give $P = g{\cdot}Q$, with $P((x, z), A{\times}B) = 1_{\{g(x,z)\in A\}}Q(z, B)$, where $(x, z) \in X \times Z$ and $A \times B \in \mathcal{X} \times \mathcal{Z}$. In iterated function systems notation we instead consider a function from the state space $S = (X, Z)$ into itself whose value depends on shocks coming from Ω, in this way we can makes the shock process an iid variable without losing any of the richness of the environment. Let $\varphi : X \times Z \times W \to X \times Z$ and let $\Omega : (W, \mathcal{W}) \to \mathbb{R}$ be an iid random variable. Then, for any process which can be represented by $P = g \cdot Q$ can also be represented as $P = \varphi \cdot \Omega$, with $P(s, C) = \Omega(\{w|\varphi(s, w) \in C\})$, where $s \in S = (X, Z)$ and $C \in \mathcal{S} = \mathcal{X} \times \mathcal{Z}$ (cf Stenflo (2001)).

Let us start by laying out the necessary assumptions

Assumption 1. *S is a compact set.*

[12] Again, more accurately, the restricion of the optimal policy function onto Y_1, see footnote 10.

Assumption 2. *(Feller property) Let* T^* *be the adjoint-operator of* $P = g \cdot Q = \varphi \cdot \Omega$ *(in our standard, and the iterated function systems notations respectively). Moreover, for every continuous function* $f : S \to \mathbb{R}$,

$$\int f(\varphi(s_j, \epsilon))\Omega(d\epsilon) \to_j \int f(\varphi(s, \epsilon))\Omega(d\epsilon) \text{ as } s_j \to_j s$$

These assumptions give us Theorem 2 of Santos and Peralta-Alva (2005), namely

Theorem 2. *Let* $\{g_j\}$ *be a sequence of policy functions that converge to* g*. Let* $\{\mu_j^*\}$ *be a sequence of probabilities on* $\mathbb{S}$ *such that* $\mu_j^* = T_j^* \mu_j^*$ *for each* j*. Under assumptions 1 and 2, if* μ^* *is a weak limit point of* μ_j^**, then* $\mu^* = T^* \mu^*$*.*

While we have already established uniqueness of the steady-state distribution for BHA models using the monotone mixing condition, this result of Santos and Peralta-Alva (2005) does not assume uniqueness of the invariant distribution, only existence of an invariant distribution.[13,14]

This results establishes that as the approximation of the optimal policy function converges to the optimal policy function, the invariant distribution it implies will converge to the invariant distribution of the true optimal policy function. It establishes this based on an assumption that P has the Feller property.

4.2.3 Bounding the distance to the true invariant distribution

In this section we present a theorem providing a bound on the distance between the invariant distribution associated with the true optimal policy function and that associated with the approximation of the optimal policy function. An related result is given by Theorem 5 of Santos and Peralta-Alva (2005) based on an assumption that φ is a stochastic contraction[15], S is compact, and L is a Lipschitz function.

As part of this we will need an assumption on the way in which the transition function of the exogenous shocks is discretized. Assume that the discretization process for the transition function satisfies

Assumption 3. Q^G *is defined such that* $Q^G(z, z_i) = Q(z, Z_i)$*,* $i = 1, ..., n_z$

Notice that (in the case where z has been previously discretized) this assumption will be satisfied by the Tauchen Method (see Tauchen (1986)).

[13] This result provides a more general and easily applied version of that in Theorem 12.13 of SLP.

[14] This version of Theorem 2 of Santos and Peralta-Alva (2005) is worded slightly differently from the original and takes advantage of the fact that as g_j converges to g we will also have that φ_j converges to φ.

[15] φ being a stochastic contraction in turn implies that P will have the Feller property. The focus here is instead on directly appealing to the properties of the Feller property and being a contraction mapping, the later of which we establish using the monotone mixing condition.

We start with the definition of a Lipschitz function: f is a Lipschitz function with constant L if, $|f(s) - f(s')| \leq L||s - s'||, \forall s, s' \in S$.

Theorem 3. *Let f be a Lipshitz function with constant L. Let $d(P^m, \hat{P}^m) \leq \delta$ for some $\delta > 0$. Assume that T^{*m} is a contraction mapping of modulus β_P. Then, under assumptions 1 & 2,*

$$|\int f(s)\mu^*(ds) - \int f(s)\hat{\mu}^*(ds)| \leq \frac{L\delta}{1-\beta_P}$$

Proof.

$$\begin{aligned}
|\int f(s)\mu^*(ds) - \int f(s)\hat{\mu}^*(ds)| &= |\int f(s)P^m \cdot \mu^*(ds) - \int f(s)\hat{P}^m \cdot \hat{\mu}^*(ds)| \\
&\leq |\int f(s)P^m \cdot \mu^*(ds) - \int f(s)\hat{P}^m \cdot \hat{\mu}^*(ds)| \\
&\leq |\int f(s)P^m \cdot \mu^*(ds) - \int f(s)P^m \cdot \hat{\mu}^*(ds)| \\
&\quad + |\int f(s)P^m \cdot \hat{\mu}^*(ds) - \int f(s)\hat{P}^m \cdot \hat{\mu}^*(ds)| \\
&\leq |\int f(s)P^m \cdot \mu^*(ds) - \int f(s)P^m \cdot \hat{\mu}^*(ds)| + Ld(P^m, \hat{P}^m) \\
&\leq \beta_P|\int f(s)\mu^*(ds) - \int f(s)\hat{\mu}^*(ds)| + Ld(P^m, \hat{P}^m) \\
&\leq \beta_P|\int f(s)\mu^*(ds) - \int f(s)\hat{\mu}^*(ds)| + L\delta
\end{aligned}$$

first line as invariant distributions, second by triangle inequality, third by definition of d and since f is Lipschitz, fourth as T^* is a contraction mapping of modulus β_P, fifth as $d(P^m, \hat{P}^m) \leq \delta$. The theorem follows from simply rearranging the terms. *Q.E.D.*

Corollary 2. *Let $d(P^m, \hat{P}^m) \leq \delta$ for some $\delta > 0$. Assume that T^{*m} is a contraction mapping of modulus β_P. Then, under assumptions 1 & 2,*

$$||\mu^* - \hat{\mu}^*|| \leq \frac{\delta}{1-\beta_P}$$

Proof. Follows from the equivalence between $||\mu^* - \hat{\mu}^*||$ and $\sup\{|\int f(s)\mu^*(ds) - \int f(s)\hat{\mu}^*(ds)| : f \in (\textit{set of bounded continuous functions on } S)$, and $||f|| \leq 1\}$ (See, eg. SLP section 11.3). Let $d = \sup_{s,s'} ||s - s'||$, which exists and is finite since S is compact. So $||f|| \leq 1$ implies that f is Lipshitz with $L = 1/d$. Then apply Theorem 3. *Q.E.D.*

To make this result useful it remains for us to get it out of the iterated function systems notation in which it is currently expressed (recall the definition of $d(P^m, \hat{P}^m)$). For this we use the following result

Lemma 2. *Let $g, \hat{g} : (X, Z) \to X$ be two policy functions satisfying $||g - \hat{g}|| \leq \delta_g$. Define $P = g \cdot Q$ and $P^G = \hat{g} \cdot Q^G$, where Q, Q^G are stochastic transition matrices on $(Z, \mathcal{Z})$, and Q^G satisfies assumption 3. Then $d(P^m, P^{Gm}) \leq m\delta_g$.*

Proof. Define $\hat{P} = \hat{g} \cdot Q$.
First, observe that

$$
\begin{aligned}
d(P^m, \hat{P}^m) &= \max_{s\in S}\left[\int ||P^m(s,w) - \hat{P}^m(s,w)||\Omega(dw)\right] \\
&\leq m \max_{s\in S}\left[\int ||P(s,w) - \hat{P}(s,w)||\Omega(dw)\right] \\
&= m \max_{s\in S}\left[\int ||(g(s), Q(z, dz(dw)) - (\hat{g}(s), Q(z, dz(dw))||\Omega(dw)\right] \\
&\leq m \max_{s\in S}\left[\int ||g(s) - \hat{g}(s)||\Omega(dw)\right] \\
&\leq m \max_{s\in S}\left[\int \delta_g \Omega(dw)\right] \\
&= m \max_{s\in S} \delta_g \\
&= m\delta_g
\end{aligned}
$$

where the first step is by the definition of $d(P^m, \hat{P}^m)$. The third from the our ability to use the standard notation in place of the iterated function systems representation (cf Stenflo (2001)). The remainder since the Qs are the same, the assumption that $||g - \hat{g}|| \leq \delta_g$, and since δ_g is a constant and Ω a distribution function.
Similarly, observe that

$$
\begin{aligned}
d(\hat{P}^m, P^{Gm}) &= \max_{s\in S}\left[\int ||\hat{P}^m(s,w) - P^{Gm}(s,w)||\Omega(dw)\right] \\
&\leq m \max_{s\in S}\left[\int ||\hat{P}(s,w) - P^{G}(s,w)||\Omega(dw)\right] \\
&= m \max_{s\in S}\left[\int ||(\hat{g}(s), Q(z, dz(dw)) - (\hat{g}(s), Q^G(z, dz(dw))||\Omega(dw)\right] \\
&= m \max_{s\in S}\left[\sum_{\substack{i=1,\ldots n_x,\\ j=1,\ldots,n_j}} \int_{X_i\times Z_j} ||(\hat{g}(s), Q(z, dz(dw)) - (\hat{g}(s), Q^G(z, dz(dw))||\Omega(dw)\right] \\
&= m \max_{s\in S}\left[\sum_{\substack{i=1,\ldots n_x,\\ j=1,\ldots,n_j}} \int_{X_i\times Z_j} 0\Omega(dw)\right] \\
&= 0
\end{aligned}
$$

where the fourth step is because the $X_i \times Z_j$, $i = 1, \ldots n_x$, $j = 1, \ldots, n_j$ form a partition of $X \times Z$, and the fifth as $\hat{g}$ is piecewise constant on the partition and by the assumption 3 on Q^G.
Combining these two by the triangle inequality we get,

$$d(P^m, P^{Gm}) \leq d(P^m, \hat{P}^m) + d(\hat{P}^m, P^{Gm}) \leq m\delta_g + 0 = m\delta_g$$

Q.E.D.

The interpretation of the just presented Lemma 2 is that due to the nature of the discretization of the optimal policy and the transition matrix for the exogenous shocks, and in particular due to considering the optimal policy function as a piecewise constant extension on the discretization grid, the discretization of the steady state distribution introduces no new errors. Thus the only errors in the steady state distribution are those which we have already bounded in terms of errors in the optimal policy function, and those coming from stopping after a finite number of iterations. The intuition for why the discretization of the steady state distribution does not create any further errors comes from noting that the approximate transition function $P^G = \hat{g} \cdot Q^G$ is piecewise constant on the partition imposed by discretization; a property it inherits via $\hat{g}$ and Q^G.

Thus, our result becomes

Corollary 3. *Let f be a Lipshitz function with constant L. Let $||g - \hat{g}|| \leq \delta_g$ for some $\delta_g > 0$. Assume that T^{*m} is a contraction mapping of modulus β_P. Then, under assumptions 1 & 2,*

$$|\int f(s)\mu^*(ds) - \int f(s)\hat{\mu}^*(ds)| \leq \frac{Lm\delta_g}{1-\beta_P}$$

and furthermore

$$||\mu^* - \hat{\mu}^*|| \leq \frac{m\delta_g}{1-\beta_P}$$

4.2.4 Combining our results on the Agents Distribution

Combining our results on the discretization of the agents distribution, and on the errors caused by only having an approximation of the optimal policy we get

Proposition 2. *Let μ_g^* be the true distribution. Let $\mu_{\hat{g},G}^N$ be the distribution obtained by iterating on the discretization using the approximate optimal policy until the convergence criterion $||\mu_{\hat{g},G}^{N+m} - \mu_{g,G}^N|| \leq \epsilon_\mu$ is reached. Let the approximation of the optimal policy function be sufficently accurate, in the sense that $||g - \hat{g}|| \leq \delta_g$. Assume that the state-space S is compact and that the discretization of Q satifies Assumption 3. Assume that the transition function $P = g\dot{Q}$ satisfies both the Feller property and the Monotone Mixing Condition. Then*

$$||\mu_g^* - \mu_{\hat{g},G}^N|| \leq \delta_\mu \equiv \frac{1}{1-\beta_P^G}(m\delta_g + \beta_P^G \epsilon_\mu) \tag{5}$$

Proof. By the triangle inequality,

$$||\mu_g^* - \mu_{\hat{g},G}^N|| \leq ||\mu_g^* - \mu_{\hat{g}}^*|| + ||\mu_{\hat{g}}^* - \mu_{\hat{g},G}^N|| \tag{6}$$

and then simply apply Corollary 3 to the first term, and Corollary 1 to the second. *Q.E.D.*

This result tells us that by iterating on the discretized steady-state agents distribution using the computed optimal policy function until we satisfy a convergence criterion (as done in the algorithm in Section 3) we will get a computational solution for the steady-state agent distribution which converges to the true solution as distance between the grid points goes to zero (as the grid size goes to infinity).

4.3 Aggregate Variables and Market Clearance

Since in many BHA models, in particular both the applications we will consider, the aggregate variables of interest are first moments of the steady-state agents distribution we have in fact already addressed their convergence. The first moments of the steady-state agents distribution can be expressed as the integral of a Lipshitz function with respect to the steady-state agents distribution. Thus convergence to the true aggregate variables of the aggregate variables computed using the computational solution of the steady-state agents distribution (computed in turn with computed optimal policy function) follows trivially from the assumption that all aggregate variables of interest can be expressed as integrals of Lipshitz functions. [16]

Similarly, since evaluating the market clearance condition, for a given price, just involves evaluating a continuously differentiable function at a point in a compact set it also introduces no further numerical errors and uniform convergence between the value of the true market clearance condition and the computational market clearance condition (using computed optimal policy function, steady-state agent distribution, and aggregate variable) follows directly from that of the aggregate variables.

4.4 General Equilibrium

So far we have concentrated on the question of convergence between the computational solution and the true solution for the optimal policy function, steady-state agent distribution, aggregate variables, and market clearance condition; for a given price. However this leaves an important part of solving a BHA model: finding the general equilibrium price. It is the computation of the general equilbrium price to which we now turn.

How should we think about finding a general equilibrium? Consider the mapping from a price — through an optimal policy function, a steady-state agent distribution, aggregate variables, and the market clearance condition — to another price. A general equilibrium will be a fixed point of this mapping, when the original price and the price implied by the mapping are the same. Equivalently we can think of this as finding the zeros of a function $F : \mathbb{R}^p \to \mathbb{R}$ which takes the value of the price minus the price implied by the aforementioned mapping. This function F, which takes the

[16] No further numerical errors are introduced by the integral as thanks to the discretization of the steady-state agents distribution computation of the integral is simply the evaluation of a function at a finite number of grid points, and then taking a weighted sum of these.

same values as the market clearance condition λ but is defined on a different space,[17] will form the basis of our theory. Looking for the general equilibria thus involves finding the zeros of F.

We begin with the following assumption,

Assumption 4. *The BHA model has a finite number of equilibria, that is F has a finite number of zeros, and F is locally strictly monotone in the local neighbourhoods of these zeros.*

Bewley (1984) proved the existence of a (general) equilbrium for BHA models, but did not prove uniqueness. In fact it remains an open issue just to prove that there are only a finite number of equilibria. Huggett (2003) provides some conditions under which the function F would be globally strictly monotone and thus the equilibrium would be unique, but these are not widely applicable. That the number of equilibria is finite would follow directly from Sard's Theorem if it could be shown that the mapping from the price — via the optimal policy, steady-state agent distribution, aggregate variables, market clearance condition — to another price is continuously differentiable; which would imply that F is continuously differentiable.[18] However the step of showing that the steady-state agent distribution is differentiable (that the steady-state of a Markov process is continuously differentiable in the parameters of the Markov transition function) remains beyond the limits of existing theory.[19] Given that neither can be proved the decision to work with Assumption 4 rather than with an assumption of continous differentiability is arbitrary and taken on the grounds that it seems like the weaker assumption.

While such continuous differentiability results do not exist we will be able to prove convergence of the algorithm in Section 3 based on assumption 4 together with continuity of the mapping from the price — via the optimal policy, steady-state agent distribution, aggregate variables, market clearance condition — to another price (and that F, defined below, changes sign). This continuity property can easily be established, in fact we have already provided conditions for the part relating to the steady-state agent distribution in Theorem 2 (Theorem 2 of Santos and Peralta-Alva (2005)[20]). That we can prove continuity but not continuous differentiability provides an important motivation for the decision, in the algorithm, to use a discrete grid on prices rather than the approach, more common in practical implementation, of finding the general equilibrium (which can be formulated as a fixed-point problem on prices) using shooting-algorithms (such as Newton-Raphson); proving the convergence of such an algorithm would require both uniqueness of the equilibrium as well as continuous differentiability.

[17] F is the convolution of the market clearance condition with the mapping from prices through to aggregate variables implied by the optimal policy function and associated steady-state agent distribution.

[18] Since, by the monotone mixing condition, we know that there is a unique steady-state agent distribution for any given price it would also eliminate the need for the local strict monotoncity assumption on F in Proposition 3.

[19] It would require some kind of Implicit Function Theorem for dynamic stochastic systems which could be applied to a function that takes the value zero when it's two arguments are the transition function and the corresponding steady-state distribution of a Markov chain.

[20] A related extension to non-compact state spaces is provided by Le Van and Stachurski (2007).

Proving the continuity[21] is quite trivial as we have already done the main steps. Assuming that the market clearance condition is continuous and that the aggregate variables are a continuous function of the steady-state agent distribution gives us that prices are a continuous function of the steady-state agent distribution. We already saw in Theorem 2 conditions under which the steady-state agent distribution is (upper-semi)continuous in the optimal policy function. That the optimal policy function is continuous in the parameters of the value function problem can be established under the assumptions of SLP Theorem 9.8 (most of these were already assumed as part of the results on the optimal policy function, but a few additional ones are involved). Since all of the relevant spaces are compact these continuities can all be considered as based in uniform convergences meaning that interchanging their order/combining them is no problem; for more see Le Van and Stachurski (2007). Note that this continuity does not require any additional assumptions beyond those already made previously.

We now have that the mapping from a price — through an optimal policy function, a steady-state agent distribution, aggregate variables, and the market clearance condition — to another price has finitely many zeros and is continuous. Convergence of the algorithm in Section 3 is thus given by the following result; interpretation is that F is the mapping from prices to prices, while $\hat{F}_i$ represents a computational approximation of the mapping (given, for each price, by the original price minus the price computated with the market clearnce condition).

Proposition 3. *Let $F : S \to \mathbb{R}$ be a continuous function with finitely many zeros, and that F changes sign around each zero (ie. Assumption 4. Assume $S \subseteq \mathbb{R}^n$ is a compact set. Let G be a finite grid on S, and define d_G to be the minimum distance between any point in S and the nearest point on the grid G. Let $\{\hat{F}_i\}$ be a sequence of approximations to F on G that converges uniformly to F on the grid. We will call a 'zero' of the function F_i to the grid point with the least absolute value out of any 2^n neighbouring grid points for which $\hat{F}_i$ is positively valued for a least one of these 2^n neighbouring grid points and negatively valued for at least one of these 2^n neighbouring grid points. Then as first i goes to infinity and second d_G goes to zero, the 'zeros' of the sequence of the set of 'zeros' of F_i converges to the set of zeros of F.*

Proof. Proof proceeds in two parts. First that the set of 'zeros' chosen by F_i converges to a subset of zeros of F. Second that for any zero of F there will be (asymptotically) a corresponding zero in F_i that converges to it.

Begin with showing that the set of 'zeros' chosen by F_i converges to a subset of zeros of F. First it is shown that for any 2^n neighbouring grid points, asymptotically (in i) there can only be at least one positively valued and one negatively valued point if there is a zero of F inbetween them. Let $s_{G,1}, s_{G,2}, ..., s_{G,n}$ be 2^n neighbouring grid points at least one of which is postively valued and one negatively valued; without loss of generality let $F_i(s_{G,1})$ be positive and $F_i(s_{G,2})$ be negative.

[21] This continuity was apparently proved in Bewley (1984). I say apparently as I have never seen this document. I assume given that Google is unable to find it that it does not exist in digital form.

Assume that there does not exist a zero of F in between them. Then by the continuity of F either $F(s_{G,1})$ and $F(s_{G,2})$ are both positive or they are both negative; without loss of generality let both be positive. Thus as i goes to infinity the convergence of F_i to F implies that there exists an I such that $F_i(s_{G,1})$ and $F_i(s_{G,2})$ are both positive for all $i > I$ — contradiction.
Thus there must be zero of F inbetween any such 2^n neighbouring grid points for all $I > i$. That the 'zero' of F_i choosen from these 2^n neighbouring grid points on the basis that it is the point of least absolute value converges to this zero follows trivially as all members of this neighbourhood will converge towards the zero of F as the grid spacing d_G goes to zero. Since there are a finite number of zeros of F, each with an associated I we can simply take the maximum of these and thus get the result that the set of 'zeros' chosen by F_i converges to a subset of zeros of F.

Now we turn to showing that for any zero of F there will be a corresponding zero in F_i that converges to it. Let s be a zero of F, then since F changes sign around each zero and by the continuity of F we have that exists some $\underline{\mathrm{d}}_G$ such that for all $d_G < \underline{\mathrm{d}}_G$ there exist 2^n neighbouring points $s_{G,1}, s_{G,2}, ..., s_{G,n}$ which will form part of the grid for d_G and for which at least one grid point F will be positively valued and for at least one grid point F will be negatively valued; without loss of generality let $F(s_{G,1}) > 0$ and $F(s_{G,2}) < 0$. Thus, by convergence of F_i to F there exists an I such that for all $i > I$, $F_i(s_{G,1}) > 0$ and $F_i(s_{G,2}) < 0$. So one member of this neighbourhood will be a 'zero' of F_i for all $i > I$ (which member may be changing in i). Since all members of the neighbourhood converge to s as d_G goes to zero we have that the sequence of 'zero's of F_i generated in this manner will converge s. Since s was an arbitrary zero of F we have shown that for any zero of F there will be a corresponding zero in F_i that converges to it. *Q.E.D.*

When the price is one-dimensional, as in many applications, "2^n neighbouring grid points" simply means two consecutive grid points. Observe that the convergence of the sequence of approximations $\{\hat{F}_i\}$ to F was established for a given price (grid point) by the results of the previous subsections; that there are a finite number of grid points is enough to make the convergence uniform on the grid. The order of convergence, first i, then d_G does matter. The assumption about F changing sign around it's zeros is not proven in the applications but would immediately follow from local continuous differentiability as not changing sign would mean it had two (or more) zeros in the same place, but we already know from the uniqueness of the steady-state agent distribution for a given price (implied by the monotone mixing condition) that this is not possible. It also seems reasonable on the grounds that if F did not change sign then an ϵ shift in any of the parameter values would increase/decrease the number of equilibria.

4.5 Summary of Theoretical Results

The following proposition simply lists all of the assumptions made until now about BHA models which together provide our final result: that the algorithm described in Section 3, based on discretization of the state space, will converge uniformly to all equilibria of a BHA model.

Proposition 4. *For any BHA model, assume that X, Y, Z are compact; X is convex; the return function F be continuous, monotone, strictly concave, and bounded; the correspondence Γ is non-empty, compact-valued, convex, and continuous; the discount factor β be between zero and one; the transition function on exogenous shocks, Q, has the Feller property; that Q is discretized using the Tauchen method (ie. satisfies Assumption 3); that the transition function P, implied by the optimal policy function together with the exogenous shock process, satisfies both the Feller property and the Monotone Mixing Condition; that the aggregate variables can be expressed as integrals of Lipschitz functions of the agents distribution; the market clearance condition is continuously differentiable; that the model has a finite number of equilibria. Then the computational solution of the BHA model by discretization using the algorithm in Section 3 will converge uniformly to all solutions as the distance between grid points goes to zero (first in the state-space, then in the price-space).*

Importantly, these assumptions are easy to establish in applications, and this will now be done for [22]

5 Applicability of the Theory to Two Examples

I turn now to the application of the theory on computational solution of BHA models to two well-known examples: Aiyagari (1994) and Pijoan-Mas (2006). Every assumption is proven to hold for both examples with the sole exception of Assumption 4 which, as previously discussed, remains an open issue in the literature. Thus, following Proposition 4 solving the models using full discretization as implemented by the algorithm of Section 3 will converge, as the grid size goes to infinity, to the true solution (for all equilibria) of these models. Results from implementing the algorithm for these two models are presented in Section 6.

5.1 The Example Models

In the model of Aiyagari (1994) infinitely lived households face a stochastic income — due to exogenous stochastic labour supply — and make consumption-savings decisions; given an interest rate. So the exogenous shock (z) is the labour supply h, the endogeneous state (x) is capital holdings k, and the decision variable (y) is next periods capital k'. The state of a household is their current capital holding and their exogenous labour supply shock, (k, h). Individual household capital holdings, given an interest rate, aggregate to give aggregate capital holdings. The market clearance condition is that the interest rate will be determined by perfect competition in the goods market together with a representative firm with Cobb-Douglas production function. A general

[22] As previously discussed, the last of these assumptions — a finite number of equilibria — will not be proven in the applications as demonstrating it remains beyond existing economic theory.

equilibrium is that an interest rate, which determines household capital holdings, which in turn by the market clearance condition determine an interest rate, and that this later interest rate is the same as the first. In short,

Definition 2. *A Competitive Equilibrium is an agents value function $V(k,h)$; agents policy function $k' = g(k,h)$; an interest rate r and wage w; aggregate capital K and labour H; and a measure of agents $\mu(k,h)$; such that*

1. *Given prices r & w, the agents value function $V(k,h)$ and policy function $k' = g(k,h)$ solve the agents problem*

$$\begin{aligned} V(k,h) = & \max_{k'} \left\{ u(c) + \beta \int V(k',h')Q(h,dh') \right\} \\ s.t. \quad & c + k' = wh + (1+r)k \\ & c \geq 0,\ k' \geq \underline{k} \end{aligned}$$

2. *Aggregates are determined by individual actions: $K = \int k d\mu(k,h)$, and $H = \int h d\mu(k,h)$*
3. *Markets clear (in terms of prices): $r - (\alpha K^{\alpha-1} H^{1-\alpha} - \delta) = 0$.*
4. *The measure of agents is invariant:*

$$\mu(k,h) = \int \int \left[\int 1_{k=g(\hat{k},h)} \mu(\hat{k},h) Q(h,dh') \right] d\hat{k} dh \tag{7}$$

where h is the labour supply shock which takes values in $Z = \{h_1, ..., h_{n_h}\}$ and evolves according to Markov transition function $Q(h,h')$. Note that the wage is residualy determined by r.[23] The market clearance condition is more commonly expressed as $r = \alpha K^{\alpha-1} H^{1-\alpha} - \delta$, that the interest equals the marginal product of capital (minus the depreciation rate). Since $H = E(h) = 1$ the Cobb-Douglas production function is really only based on aggregate capital (in the sense that H is a fixed constant).

The model of Pijoan-Mas (2006) endogenizes the labour supply, or equivalently leisure.[24] Infinitely lived households face a stochastic labour productivity shock. They make consumption-leisure and consumption-savings decisions; given an interest rate. So the exogenous shock (z) is the labour productivity shock z, the endogeneous state (x) is capital holdings k, and the decision variable (y) is next periods capital k' and leisure l (=1−h). The state of a household is their current capital holding and their exogenous labour productivity shock, (k,z). Individual household capital holdings, given an interest rate, aggregate to give aggregate capital holdings. The market clearance condition is that the interest rate will be determined by perfect competition in the goods market

[23]The wage, which is given by the derivative of the Cobb-Douglas production with respect to labour, can be rewritten as a function of the interest rate and the parameters of the production function.

[24]Since the theoretical results, are based on F being increasing we will use leisure as the endogeneous decision variable, rather than labour supply.

together with a representative firm with Cobb-Douglas production function. A general equilibrium is that an interest rate, which determines household capital holdings, which in turn by the market clearance condition determine an interest rate, and that this later interest rate is the same as the first. In short,

Definition 3. *A Competitive Equilibrium is an agents value function $V(k,z)$; agents policy function $(k',l) = g(k,z)$; an interest rate r and wage w; aggregate capital K and aggregate labour (efficiency-units) L; and a measure of agents $\mu(k,z)$; such that*

1. *Given prices r & w, the agents value function $V(k,z)$ and policy function $(k',l) = g(k,z)$ solve the agents problem*

$$V(k,z) = \max_{k'} \left\{ u(c) + n(l) + \beta \int V(k',z')Q(z,dz') \right\}$$
$$s.t. \quad c + k' = w(1-l)z + (1+r)k$$
$$c \geq 0,\ 0 \leq l \leq 1,\ k' \geq \underline{k}$$

2. *Aggregates are determined by individual actions: $K = \int k d\mu(k,z)$, and $L = \int z(1-l)d\mu(k,z)$*

3. *Markets clear (in terms of prices): $r - (\alpha K^{\alpha-1}L^{1-\alpha} - \delta) = 0$.*

4. *The measure of agents is invariant:*

$$\mu(k,z) = \int\int \left[\int 1_{k=g^1(\hat{k},z)} \mu(\hat{k},z)Q(z,dz') \right] d\hat{k}dz \tag{8}$$

where z is the labour productivity shock which takes values in $Z = \{z_1, ..., z_{n_z}\}$ and evolves according to Markov transition function $Q(z,z')$. The wage is again residualy determined by r.[25]

5.2 Proving the Assumptions Hold

The first couple of assumptions all trivially hold as they are mostly assumptions explicitly made about the primatives of the model while setting up the models themselves. Thus we have that X, Y, Z are compact; X is convex; the return function F be continuous, monotone, strictly concave, and bounded; the correspondence Γ is non-empty, compact-valued, convex, and continuous; the discount factor β be between zero and one; and the transition function on exogenous shocks, Q, has the Feller property.[26] Also trivially holding are the assumptions that the aggregate variables can

[25] Note that L is the integral of $z(1-l)$, it is *not* the aggregate value of l. This notation is used, rather than using $h = 1-l$, because applying the theory requires that the return function be increasing in it's arguments. It also happens to coincide with the notation of Pijoan-Mas (2006).

[26] The only one of these which is not entirely trivial is that X and Y are compact. It is however easy to show that since Z is bounded (by assumption) it follows that X, and hence Y must be. For example Aiyagari (1993) shows that Z being bounded implies bounds on X given by $X = [\underline{x}, \bar{x}]$ where $\underline{x} = \lim_{t\to\infty} x_t^{low}$ and $\bar{x} = \lim_{t\to\infty} x_t^{high}$; where $x_t^{low} = g(x_{t-1}^{low}, \underline{z})$ and $x_t^{high} = g(x_{t-1}^{high}, \bar{z})$. While Aiyagari (1993) only treats the case of an iid exogenous shock an extension to the general Markov shock case can be found in Miao (2006).

be expressed as integrals of Lipschitz functions of the agents distribution (they are first moments); and that the market clearance condition is continuously differentiable.

This leaves us two assumptions which remain to be proven: that the transition function P, implied by the optimal policy function together with the exogenous shock process, satisfies both the Feller property and the Monotone Mixing Condition.

Proving the Feller property holds for both models is relatively easy. Both models satisfy Theorem 9.8 of SLP, which gives conditions under which the optimal policy function is continuous. Theorem 9.14 of SLP then implies that both models satisfy the Feller property (under the additional assumptions that X is convex, Z is compact, and the transtion function on exogenous shocks, Q, itself satisfies the Feller property).

While the literature using BHA models makes frequent reference to the monotone mixing condition of Hopenhayn and Prescott (1992) (see, eg. Marcet, Obiols-Homs, and Weil (2007) and Ríos-Rull (2001)) actual proofs of it's applicability are few. Huggett (1993) provides a proof that his model satisfies the monotone mixing condition, but this is very much the exception that proves the rule. Proofs are now given that the models of Aiyagari (1994) and Pijoan-Mas (2006) do satisfy the monotone mixing condition.

5.2.1 MMC in Aiyagari Model

For the model of Aiyagari (1994) the monotone mixing condition is established by the

Lemma 3. *The model of Aiyagari (1994) satisfies that P is increasing.*

Proof. Optimal policies are increasing (eg. Aiyagari (1993), or Miao (2006)), and Q is increasing. Thus $P = g \cdot Q$ is increasing. *Q.E.D.*

Proposition 5. *The model of Aiyagari (1994) satisfies the monotone mixing condition.*

Proof. Let $k^* = (\underline{k} + \bar{k})/2$. Define a sequence $a_1 = \underline{k}$, $a_2 = g(a_1, \bar{h})$, $a_3 = g(a_2, \bar{h}), ...$ and a sequence $b_1 = \bar{k}$, $b_2 = g(b_1, \underline{h})$, $b_3 = g(b_2, \underline{h}), ...$ Note that $\{a_i\} \to \bar{k}$ monotonically and $\{b_i\} \to \underline{k}$ monotonically. Therefore, there exists an N_1 such that an agent goes from $\underline{k}$ to $\{k \in X : k \geq k^*\}$ with positive probability in N_1 or less steps, and likewise an N_2 such that an agent goes from $\bar{k}$ to $\{k \in X : k \leq k^*\}$ with positive probability in N_2 or less steps. Choose $N = max\{N_1, N_2\}$ in the mixing condition. *Q.E.D.*

5.2.2 MMC in Pijoan-Mas (2006) Model

Proving that the model of Pijoan-Mas (2006) also satisfies the monotone mixing condition involves the same general steps, but each is now more complicated. We proceed through the steps in the following three Lemmas and a Proposition. First though we will establish a more convenient notation.

We denote the optimal policies that solve this problem by $g^{k'}(k,z)$, $g^c(k,z)$, & $g^l(k,z)$. To simplify notation we define $S = X \times Z$, where $X = [\underline{\text{k}}, \infty)$ is the set of possible asset choices (we will later show that in fact we can bound X from above). That the solution to this problem will display the following properties; $V(k,z)$ is strictly increasing in k and increasing in z, $V(k,z)$ is strictly concave in k. V is continuously differentiable, follows from some standard results.[27]. For the purpose of all the following theory we simply assume that the value function to be solved is of the more general form

$$V(k,z') = \max_{l,k'} F(k,k',l,z) + \beta E[V(k',z')|z]$$

where z takes values in a compact set with minimum z_1 and maximum z_{n_z}; F is str. increasing and str. concave in k & l, inc. and concave in z, str. decreasing and str. concave in k'; F_l is independent of k, k' and z. The model of Pijoan-Mas (2006) fits this general form.

We begin our proofs by showing that, for all points where the optimal policies are interior, the optimal polices for both assets and leisure are strictly increasing (we cannot just invoke the standard Theorems of SLP because of the presence of the leisure choice, which is not a state).

Lemma 4. *Let (k,z) be such that $g^{k'}(k,z) > \underline{k}$ and $g^l(k,z) < 1$, then $g^{k'}(k,z)$ and $g^l(k,z)$ are both strictly increasing in k.*

Proof. For optimality, the first-order conditions imply that

$$-F_{k'}(k, g^{k'}(k,z), g^l(k,z), z) = \beta E[V_k(g^{k'}(k,z), z')|z] \tag{9}$$

and

$$-F_{k'}(k, g^{k'}(k,z), g^l(k,z), z) = F_l(k, g^{k'}(k,z), g^l(k,z), z) \tag{10}$$

(we know that V can be derivised even when the borrowing constraint binds from Rincón-Zapatero and Santos (2009)).
Let $k^2 > k^1$, and $g^{k'}(k^1,z) > \underline{\text{k}}$.
Assume $g^{k'}(k^2,z) \leq g^{k'}(k^1,z)$ and $F_{k'}$ is independent of l (a sufficient condition for this would be that utility is seperable in consumption and leisure, as is the case here).

[27] Marcet, Obiols-Homs, and Weil (2007) cite Thm 9.6 of SLP for existence of a solution. In fact the required Theorems for their model are the extensions of Theorem 9.6 to the Case 2 value function (SLP's categorization) in Exercise 9.7 of SLP. For the rest of our results we use the extensions of Thm 9.7 and 9.11 contained in Exercise 9.7 of SLP. The 'strictly' in Thm 9.11 (V strictly inc. in z) must be dropped as the return function is only increasing in z (because of possibility that labour supply equals zero/leisure equals one). That V is differentiable, even in presence of the borrowing constraint, is proved in Rincón-Zapatero and Santos (2009)

Then,

$$\begin{aligned}
-F_{k'}(k^1, g^{k'}(k^1,z), g^l(k^1,z), z) &= \beta E[V_k(g^{k'}(k^1,z), z')|z] \\
&\leq \beta E[V_k(g^{k'}(k^2,z), z')|z] \\
&= -F_{k'}(k^2, g^{k'}(k^2,z), g^l(k^2,z), z) \\
&< -F_{k'}(k^1, g^{k'}(k^2,z), g^l(k^2,z), z) \\
&\leq -F_{k'}(k^1, g^{k'}(k^1,z), g^l(k^2,z), z) \\
&= -F_{k'}(k^1, g^{k'}(k^1,z), g^l(k^1,z), z)
\end{aligned}$$

where, steps are by FOC; by strict concavity of V in k; by FOC; as $k^2 > k^1$ & F is increasing and strictly concave in k; as $g^{k'}(k^2,z) \leq g^{k'}(k^1,z)$, F is dec. and concave in k'; as utility is seperable in c & l.
– Contradiction.
Thus, $g^{k'}(k^2,z) \geq g^{k'}(k^1,z)$.

So $g^{k'}$ is strictly increasing in k for (k,z) s.t. $g^{k'}(k,z) > \underline{\mathrm{k}}$.

That $g^l(k,z)$ must then also be str. inc. in l for (k,z) s.t. $g^{k'}(k,z) > \underline{\mathrm{k}}$ & $g^l(k,z) < 1$, then follows immediately from one of the FOCs together with the envelope condition. The envelope condition, together with that V is str. inc. in k and that $g^{k'}(k,z)$ is str. inc. in k implies that the LHS of 10 is str. dec. in k; which implies that the RHS of 10 is str. dec. in k; which implies that g^l is str. inc. in k (observe again that this argument uses the seperability of the utility fn in c & l).

Trivially, these results for strictly increasing $g^{k'}$ & g^l on the 'interior' (choice-wise) could be extended to a result of increasing for all (k,z). *Q.E.D.*

We now turn to two lemmas that will be used to show the mixing condition.

Lemma 5. *$g^{k'}(k,z_1) < k, \quad \forall (k,z_1)$ s.t. $g^l(k,z_1) < 1$, $k > \underline{k}$.*

Proof. Observe that, $V_k(k,z_1) > V_k(k,z), \forall z > z_1, \quad \forall (k,z_1)$ s.t. $g^l(k,z_1) < 1$.
$\implies V_k(k,z_1) > \beta E[V_k(k,z')|z_1], \quad \forall (k,z_1)$ s.t. $g^l(k,z_1) < 1$.
$\implies g^{k'}(k,z_1) < k, \quad \forall (k,z_1)$ s.t. $g^l(k,z_1) < 1$, $k > \underline{\mathrm{k}}$.
where this last step follows from the envelope condition, FOC, and that V is str. inc. and str. concave in k by the following reasoning,
Env. Condn $\implies V_k(k,z_1) = -F_{k'}(k, g^{k'}(k,z_1), g^l(k,z_1), z_1)$
FOC $\implies \quad = \beta E[V_k(g^{k'}(k,z_1), z')|z_1]$
Therefore, $E[V_k(g^{k'}(k,z_1), z')|z_1] > E[V_k(k,z')|z_1]$ which in turn implies $g^{k'}(k,z) < k$. *Q.E.D.*

Notice that if $g^l(k,z_1) = 1$, we would get $g^{k'}(k,z_1) \leq k$.

Lemma 6. *There exists k s.t. $g^{k'}(k, z_{n_z}) = k$.*

Proof. Suppose not. Then $g^{k'}(k, z_{n_z}) > k$, $\forall k$.
Since $l \in [0, 1]$ a compact set, & F is inc. and concave in l it follows that there exists a $\min_{l \in [0,1]} F_l$ (because of seperable utility assumption; observe also that it is given by $F_l(\cdot, \cdot, 1, \cdot)$). Now, by assumption $F_{k'} \to 0$ as $k' \to \infty$, and we have seen that $g^{k'}$ is str inc. for $g^l < 1$, and inc. everywhere. Thus, there exists some k^* s.t $-F_{k'}(k^*, g^{k'}(k^*, z), g^l(k^*, z), z) < \min_{l \in [0,1]} F_l$. Now, since for interior solution the FOCs require that $-F_{k'} = F_l$, it follows that the optimal choice at (k^*, z_{n_z}) cannot be interior, so $g^l(k^*, z_{n_z}) = 1$.
Rest of proof follows as g^k is inc. in z[28], so then for all $k > k^*$, the choice of leisure equals ones means the the budget constraint becomes $c + k' \leq (1 + r)k$ and $u(c)$ is str. inc. and str. concave, so it is well known that there is some maximum optimal choice of k (the maximum k may be below k^*, but this doesn't matter, as we still get a bound). This is a contradiction. *Q.E.D.*

Now we use the proceeding three lemmas to show that the Monotone Mixing Condition holds. Denote $\bar{k}$ as $\min_k g^{k'}(k, z_{n_z}) = k$.

Proposition 6. *The model of Pijoan-Mas (2006) satisfies the Monotone Mixing Condition.*

Proof. That P is monotone follows immediately from monotonicity of Q (which was assumed by model), and monotoncity of $g^{k'}(\cdot, z)$, $\forall z$ (Lemma 4).
The mixing condition: Choose $s^* = ((\underline{k} + \bar{k})/2, z_{n_z/2})$. Define a sequence $a_1 = \underline{k}$, $a_2 = g^{k'}(a_1, z_{n_z})$, $a_3 = g^{k'}(x_2, z_{n_z})$,... and a sequence $b_1 = \bar{k}$, $b_2 = g^{k'}(b_1, z_1)$, $b_3 = g^{k'}(b_2, z_1)$,.... Note that $\{a_n\} \to \bar{k}$ monotonically and $\{b_n\} \to \underline{k}$ montonically. Therefore, there is an N_1 such that an agent goes from $(\underline{k}, z_1)$ to $\{s \in S : s \geq s^*\}$ with positive probability in N_1 or greater steps and there is an N_2 such that an agent goes from $(\bar{k}, z_{n_z})$ to $\{s \in S : s \leq s^*\}$ with positive probability in N_2 or greater steps. Let $N = \{N_1, N_2\}$ in the mixing condition. *Q.E.D.*

6 Results for the Two Examples

CURRENT RESULTS USE SMALL GRIDS AND ARE INTENDED AS ILLUSTRATIVE, NOT FINAL.

In presenting the results of applying the algorithm to the models of Aiyagari (1994) and Pijoan-Mas (2006) focus will be on to two aspects. The first is replication of the results of the original papers, this means replication of the main tables and numerical results of the papers. The second relates current interest in the work of Piketty (2014) and is the relationship between what Piketty calls $r > g$ — that interest rates are greater than the economic growth rate — and inequality. Piketty posits a positive relationship between $r > g$ and inequality. Such a relationship is present

[28] A simple implication of observation $z^2 > z^1$ implies $V_k(k, z^1) \geq V_k(k, z^2)$, which by envelope condn implies $-F_{k'}(k, g^{k'}(k, z^1), g^l(k, z^1), z^1) = -F_{k'}(k, g^{k'}(k, z^2), g^l(k, z^2), z^2)$. From there it is just a matter of following the later steps of the proof of Lemma 4.

as a steady-state relationship in a number of models Jones (2015); although as it is only one of many forces driving inequality we would not expect the relationship to be unconditional (Piketty, 2015), an expectation borne out empirically in cross-country data (Acemoglu and Robinson, 2015). While there is no growth in these models, so $g = 0$, we will however look at the relationship between r and inequality.

We start with the results for the model of Aiyagari (1994). The functional forms and calibrated parameter values are as follows. The utility function is parameterized as $u(c) = \frac{c^{1-\mu}}{1-\mu}$. The shock process is $z' = \rho z + \epsilon$, where $\epsilon \sim N(0, \sigma^2)$. The discount rate is $\beta = 0.96$, captial-share of production is $\alpha = 0.36$, and the depreciation rate is $\delta = 0.08$. Following Aiyagari (1994) we look at varying the parameters $\mu \in \{1, 3, 5\}$, $\rho \in \{0, 0.3, 0.6, 0.9\}$ and $\sigma \in \{0.2, 0.4\}$. The grid on the exogenous shocks is given by the Tauchen method with $n_z = 21$. The grid on assets is $n_k = 128$ points, one-third evenly spaced on the interval $[0,K_{ss}]$, one-third evenly spaced on the interval $(K_{ss}, 3K_{ss}]$, and the final third evenly spaced on the interval $(3K_{ss},15K_{ss}]$; where $K_{ss} = (\frac{r_{ss}+\delta}{\alpha})^{\frac{1}{\alpha-1}}$ and $r_{ss} = 1/\beta - 1$ are the steady-state capital stock and interest rate of the corresponding complete markets representative agent economy. The grid on prices is $n_p = 151$ points, one-third evenly spaced on the interval $[-\delta,0)$, and the two-thirds evenly spaced on the interval $[0,r_{ss}]$.

Table 1 shows that, at least as measured by the first-order autocorrelation and variance of the process, the Tauchen method was accurate in discretizing the exogenous process. Although in my experience the weakness of the Tauchen method tends to be related to the choice of the parameter q, an issue not addressed by this Table.[29] In any case, these results confirm those of Aiyagari (1994).

A comparison of Table 2 with the corresponding Table 2 of Aiyagari (1994) shows that the original quantitative results that he gives for the equilibrium interest rates display the correct qualitative behaviour; decreasing both in risk-aversion (μ) and in the riskiness of earnings (σ & ρ). Quantitatively however the results of Aiyagari (1994) are quite innaccurate due to the roughness of the numerical approximations used.

Suggests a more subtle relationship between r (or $r - g$) and inequality, as measured by the Gini coefficient. There is a positive relationship between r and Gini coefficients is the increase in r is due to a decrease in μ (risk-aversion). But if the increase in r comes about due to a decrease in earnings risk (decreasing σ or ρ) then this will in fact be associated with Gini coefficients. Decreasing earnings risks leads to a decrease in precautionary savings, which means a decreased capital stock and an increase in the equilibrium interest rate.

We now turn to the results for the model of Pijoan-Mas (2006). The functional forms and

[29] The grid used by the Tauchen Method for discretizing the AR(1) process $y_t = \rho y_{t-1} + \epsilon_t$, where $\epsilon \sim N(0, \sigma^2)$ ranges from $-q\sigma/\sqrt{1-\rho}$ to $q\sigma/\sqrt{1-\rho}$. The (arbitrary) choice of q therefore determines the biggest and smallest shocks, which are strongly related to the 'risk' of the process and therefore is an important choice quantitatively in Economic models of behaviour under risk. I follow Footnote 33 of Aiyagari (1993), pg 25, which implicitly says that he uses $q = 3$.

Table 1: Accuracy of the Tauchen Method in Aiyagari (1994)
Markov Chain Approximation to the Labour Endowment Shock
Markov Chain σ/Markov Chain ρ

σ/ρ	0.00	0.30	0.60	0.90
0.2	0.20/0.01	0.20/0.30	0.20/0.60	0.20/0.90
0.4	0.40/0.01	0.40/0.31	0.40/0.59	0.40/0.90

Replication of Table 1 of Aiyagari (1994) using grid sizes $n_k = 256$, $n_z = 21$, $n_p = 151$

Table 2: General Equilibrium Interest Rates in Aiyagari (1994)

A. Net Return to Capital in %/Aggregate savings rate in % ($\sigma = 0.2$)			
ρ/μ	1	3	5
0.0	4.1667/31.04	4.1250/23.82	4.0417/23.63
0.3	4.1250/22.84	4.0417/24.10	3.9583/24.43
0.6	4.0833/22.59	3.8750/23.55	3.6667/25.15
0.9	4.0000/24.45	3.6250/24.71	3.1250/26.23
B. Net Return to Capital in %/Aggregate savings rate in % ($\sigma = 0.4$)			
ρ/μ	1	3	5
0.0	4.0833/22.77	3.9167/24.11	3.7083/24.55
0.3	4.0000/22.92	3.6250/25.08	3.1250/25.94
0.6	3.8750/24.51	3.1250/26.07	2.3333/28.02
0.9	3.6250/24.92	2.2083/28.38	0.7500/32.95

Replication of Table 2 of Aiyagari (1994) using grid sizes $n_k = 256$, $n_z = 21$, $n_p = 151$

calibrated parameter values are as follows. The utility function is parameterized as $u(c) = \frac{c^{1-\sigma_1}}{1-\sigma_1} + \chi\frac{l^{1-\sigma_2}}{1-\sigma_2}$. The shock process is $z' = \rho z + \epsilon$, where $\epsilon \sim N(0, \sigma^2)$. Four different parameterizations are considered and the parameter values for each appear in Table 4; except those for ρ and σ which are 0.92 and 0.21 for the first three calibrations, and 0.977 and 0.12 for the fourth ($HighPersist$) calibration. The grid on the exogenous shocks is given by the Tauchen method with $n_z = 15$. The grid on hours worked is $n_l = 21$ points evenly spaced on the interval [0,1]. The grid on assets is $n_k = 128$ points, one-third evenly spaced on the interval [$\underline{a}$,K_{ss}], one-third evenly spaced on the interval $(K_{ss}, 3K_{ss}]$, and the final third evenly spaced on the interval ($3K_{ss}$,$15K_{ss}$]; where $K_{ss} = (\frac{r_{ss}+\delta}{\alpha})^{\frac{1}{\alpha-1}}$ and $r_{ss} = 1/\beta - 1$ are the steady-state capital stock and interest rate of the corresponding complete markets representative agent economy with labour supply fixed at one (ie. the same as in Aiyagari (1994) model). The grid on prices is $n_p = 151$ points, one-third evenly spaced on the interval [$-\delta$,0), and the two-thirds evenly spaced on the interval [0,$2r_{ss}$].

Table 3: Interest Rates and Inequality in Aiyagari (1994)
Gini Coefficients for Earnings, Income, and Wealth

A. Earnings Gini/Income Gini/Wealth Gini ($\sigma = 0.2$)			
ρ/μ	1	3	5
0.0	0.11/0.25/0.29	0.11/0.21/0.26	0.11/0.21/0.26
0.3	0.11/0.29/0.35	0.11/0.27/0.32	0.11/0.25/0.30
0.6	0.11/0.34/0.41	0.11/0.31/0.37	0.11/0.29/0.35
0.9	0.11/0.46/0.55	0.11/0.40/0.47	0.11/0.35/0.41
B. Earnings Gini/Income Gini/Wealth Gini ($\sigma = 0.4$)			
ρ/μ	1	3	5
0.0	0.22/0.29/0.35	0.22/0.27/0.32	0.22/0.26/0.31
0.3	0.22/0.31/0.37	0.22/0.29/0.34	0.22/0.27/0.32
0.6	0.22/0.35/0.41	0.22/0.31/0.36	0.22/0.29/0.33
0.9	0.23/0.47/0.55	0.23/0.41/0.47	0.23/0.37/0.41

Aiyagari (1994) reports a few Gini coefficients, but no Table.
Uses grid sizes $n_k = 256$, $n_z = 21$, $n_p = 151$

7 Conclusion

Theory is developed showing that solving Bewley-Huggett-Aiyagari models with an algorithm based on discretization will give us the true solution as grid spacing goes to zero. The theory is shown to apply to the models of Aiyagari (1994) and Pijoan-Mas (2006) and then used to provide replication of the results of those papers. It was seen that the results of Aiyagari (1994), while qualitatively correct, are not so accurate quantitatively.

The solutions to Bewley-Huggett-Aiyagari models using this algorithm provide both a way to ensure accuracy — especially important for applications like inequality and the Top 1% — and a known benchmark against which other numerical methods can be compared. When implemented on a GPU the algorithm is also quite fast thanks to the ability to parallelize many steps, in particular that of value function iteration.

References

Daron Acemoglu and James Robinson. The rise and decline of general laws of capitalism. Journal of Economic Perspectives, 29(1):3–28, 2015.

S. Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving, 1993. Working Paper 502.

S. Rao Aiyagari. Uninsured idiosyncratic risk and aggregate saving. Quarterly Journal of Economics, 109(3):659–684, 1994.

Eric Aldrich, Jesus Fernandez-Villaverde, Ronald Gallant, and Juan Rubio-Ramirez. Tapping the

Table 4: Calibrated Parameters and Targets in Pijoan-Mas (2006)
Parameter Values and the Accuracy of the Calibration Targets

Parameter	Target	Calibrated Value			
		Model E0	$\underline{a} = -Y$	$\sigma_2 = 8.0$	$HighPersist$
σ_1	corr(h,z)=0.02	1.458	1.514	1.412	0.992
σ_2	cv(h)=0.22	2.833	3.459	8.000	1.290
χ	H=0.33	0.856	0.674	0.110	1.545
β	K/Y=3.00	0.945	0.948	0.948	0.955
α	wL/Y=0.64	0.640	0.640	0.640	0.640
δ	I/Y=0.25	0.083	0.083	0.083	0.083
Parameter	**Target**	**Model Moments**			
		Model E0	$\underline{a} = -Y$	$\upsilon = 8.0$	$HighPersist$
σ_1	corr(h,z)=0.02	0.102	0.000	0.089	0.378
σ_2	cv(h)=0.22	0.233	0.000	0.089	0.167
χ	H=0.33	0.250	0.010	0.296	0.315
β	K/Y=3.00	9.474	-12.983	8.394	7.935
α	wL/Y=0.64	0.107	0.000	0.113	0.113
δ	I/Y=0.25	0.786	-1.078	0.697	0.659

Replication of Table 8 of Pijoan-Mas (2006) using grid sizes $n_l = 21$, $n_k = 64$, $n_z = 15$, $n_p = 71$
My σ_1, σ_2, χ, & α are denoted σ, υ, λ, & θ by Pijoan-Mas (2006).
Pijoan-Mas (2006) does not report second half of this Table. But does note that "For the economy with $\sigma_2 = 8.0$ the $cv(h)$ is kept free and turns out to be 0.09. For all other economies the model parameters deliver the same 6 statistics".

supercomputer under your desk: Solving dynamic equilibrium models with graphics processors. Journal of Economic Dynamics and Control, 35(3):386–393, 2011.

S. Aruoba, Jesus Fernandez-Villaverde, and Juan Rubio-Ramirez. Comparing solution methods for dynamic equilibrium economies. Journal of Economic Dynamics and Control, 30(12):2477–2508, 2006.

Alejandro Badel and Mark Huggett. Taxing top earners: a human capital perspective. Working Papers of the Fed Reserve of St Louis, 2014-017, 2014.

Fransisco Barillas and Jesus Fernandez-Villaverde. A generalization of the endogenous grid method. Journal of Economic Dynamics and Control, 31:2698–2712, 2007.

Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The distribution of wealth and fiscal policy in economies with finitely lived agents. Econometrica, 79(1):123–157, 2011.

Hugo Benitez-Silva, George Hall, Gunter Hitsch, Giorgio Pauletto, and John Rust. A comparison of discrete and parametric approximation methos for solving dynamic programming problems in economics. Working Paper, n.a., 2005.

Truman Bewley. A difficulty with the optimum quantity of money. Econometrica, 51:1485–1504, 1983.

Table 5: Some Model Statistics for Pijoan-Mas (2006)
Parameter Values and the Accuracy of the Calibration Targets

Model Statistic	Model			
	E0	$\underline{a} = -Y$	$\sigma_2 = 8.0$	$HighPersist$
Y	1.048	0.034	1.148	1.287
C	0.224	0.236	0.348	0.439
K	9.928	-2.433	9.638	10.212
L	0.296	0.012	0.347	0.401
H	0.250	0.010	0.296	0.315
L/H	1.183	1.167	1.173	1.272
Y/H	4.193	3.397	3.880	4.079
corr(h,z)	0.102	0.000	0.089	0.378
corr(k,z)	-0.537	-0.000	-0.420	-0.189

Replication of parts of Tables 3, 4, 6, and 9 of Pijoan-Mas (2006) using grid sizes $n_l = 21$, $n_k = 64$, $n_z = 15$, $n_p = 71$

Truman Bewley. Notes on stationary equilibrium with a continuum of independently fluctuating consumers, 1984.

Rabi Bhattacharya and Oesook Lee. Asymmptotics of a class of markov processes which are not in general irreducible. The Annals of Probability, 16(3):1333–1347, 1988. A correction to this article was published in the same journal in 1997, Vol. 25(3), pg. 1541-1543.

Rabi Bhattacharya and Mukal Majumdar. On a class of stable random dynamical systems: Theory and applications. Journal of Economic Theory, 96:208–229, 2001.

Yongyang Cai and Kenneth L. Judd. Advances in numerical dynamic programming and new applications. In Karl Schmedders and Kenneth L. Judd, editors, Handbook of Computational Economics, volume 3, chapter 8. Elesevier, 2014.

Dario Caldara, Jesus Fernandez-Villaverde, Juan Rubio-Ramirez, and Wen Yao. Computing dsge models with recursive preferences and stochastic volatility. Review of Economic Dynamics, 15(2): 188–206, 2012.

Christopher Carroll. The method of endogenous gridpoints for solving dynamic stochastic optimization problems. Economics Letters, 91:312–320, 2006.

Ana Castaneda, Javier Díaz-Giménez, and Jose Victor Ríos-Rull. Accounting for the u.s. earnings and wealth inequality. Journal of Political Economy, 111(4):818–857, 2003.

Juan Carlos Conesa and Dirk Krueger. On the optimal progressivity of the income tax code. Journal of Monetary Economics, 53(7):1425–1450, 2006.

Juan Carlos Conesa, Sagiri Kitao, and Dirk Krueger. Taxing capital? not a bad idea after all! American Economic Review, 99(1):25–48, 2009.

Table 6: Inequality in Pijoan-Mas (2006)
Distributional Statistics

Variable	cv	Gini	q1	q2	q3	q4	q5
Hours							
Earnings							
Model E0	0.60	0.33	0.06	0.12	0.17	0.25	0.39
Model $\underline{a} = -Y$	0.59	0.30	0.08	0.13	0.17	0.23	0.38
Model $\sigma_2 = 8.0$	0.58	0.31	0.08	0.13	0.17	0.24	0.38
Model $HighPersist$	0.75	0.38	0.06	0.11	0.16	0.23	0.45
Income							
Model E0	0.50	0.28	0.04	0.14	0.23	0.28	0.30
Model $\underline{a} = -Y$	-2.03	-0.53	0.31	0.31	0.31	0.31	-0.26
Model $\sigma_2 = 8.0$	0.51	0.29	0.04	0.14	0.22	0.29	0.31
Model $HighPersist$	0.55	0.29	0.02	0.14	0.25	0.29	0.30
Wealth							
Model E0	0.50	0.28	0.04	0.14	0.23	0.29	0.30
Model $\underline{a} = -Y$	-2.03	-0.53	0.31	0.31	0.31	0.31	-0.26
Model $\sigma_2 = 8.0$	0.51	0.29	0.04	0.14	0.22	0.29	0.31
Model $HighPersist$	0.55	0.29	0.02	0.14	0.26	0.29	0.29

Replication of Tables 2 of Pijoan-Mas (2006) using grid sizes $n_l = 21$, $n_k = 64$, $n_z = 15$, $n_p = 71$
The equilibrium interest rates for the four calibrations are $r_{E0} = 5.8201$, $r_{\underline{a}=-Y} = -8.2900$, $r_{\sigma_2=8.0} = 5.4852$, $r_{HighPersist} = 4.7120$.
cv refers to coefficient of variance. $q1, ..., q5$ refer, for earnings, income, and wealth, to the share held by all people in the corresponding quintile with respect to the total. However, for hours it is the average number of hours worked by people in the corresponding quintile.

Javier Díaz-Giménez, Vicente Quadrini, and Jose Victor Ríos-Rull. Dimensions of inequality: facts on the u.s. distributions of earnings, income, and wealth. Quarterly Review of the Federal Reserve of Minneapolis, Spring:3–21, 1997.

Javier Díaz-Giménez, Andy Glover, and Jose Victor Ríos-Rull. Facts on the distributions of earnings, income, and wealth in the united states: 2007 update. Quarterly Review of the Federal Reserve of Minneapolis, 34(1):2–31, 2011.

David Domeij and Jonathan Heathcote. On the distributional effects of decreasing capital taxes. International Economic Review, 45(2):523–544, 2004.

Giulio Fella. A generalized endogenous grid method for non-smooth and non-concave problems. Review of Economic Dynamics, 17(2):329–344, 2014.

Nezih Guner, Martin Lopez-Daneri, and Gustavo Ventura. Heterogeneity and government revenues: Higher taxes at the top? IZA Discussion Papers, 8335, 2014.

Juan Carlos Hatchondo, Leonardo Martinez, and Horacio Sapriza. Quantitative properties of sovereign default models: Solution method. Review of Economic Dynamics, 13(4):919–933, 2010.

Jonathan Heathcote, Kjetil Storesletten, and Giovanni Violante. Quantitative macroeconomics with heterogeneous households. Annual Review of Economics, 1(5):319–354, 2009.

Hugo Hopenhayn and Edward C. Prescott. Stochastic monotonicity and stationary distributions for dynamic economies. Econometrica, 60(6):1387–1406, 1992.

Mark Huggett. The risk-free rate in heterogenous-agent incomplete-insurance economies. Journal of Economic Dynamics and Control, 17:953–969, 1993.

Mark Huggett. When are comparative dynamics monotone? Review of Economic Dynamics, 6(1): 1–11, 2003.

Charles Jones. Pareto and piketty: The macroeconomics of top income and wealth inequality. Journal of Economic Perspectives, 29(1):29–46, 2015.

Takashi Kamihigashi and John Stachurski. Stochastic stability in monotone economies. Theoretical Economics, 9(2):383–407, 2014.

Takashi Kamihigashi and John Stachurski. Seeking ergodicity in dynamic economies. Working paper, 2015.

Fabian Kinderman and Dirk Krueger. High marginal tax rates on the top 1 CFS Working Paper Series, 473, 2014.

Robert Kirkby. Convergence of discretized value function iteration. Working Paper, 2015.

C. Le Van and John Stachurski. Parametric continuity of stationary distributions. Economic Theory, 33:333–348, 2007.

Lilia Maliar and Serguei Maliar. Envelope condition method versus endogenous grid method for solving dynamic programming problems. Economics Letters, 120:262–266, 2013.

Albert Marcet, Francesc Obiols-Homs, and Philippe Weil. Incomplete markets, labor supply and capital accumulation. Journal of Monetary Economics, 54(8):2621–2635, 2007.

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

Jeno Pál and John Stachurski. Fitted value function iteration with probability one contractions. Journal of Economic Dynamics and Control, 37:251–264, 2013.

Adrian Peralta-Alva and Manuel Santos. Analysis of numerical errors. In Karl Schmedders and Kenneth L. Judd, editors, Handbook of Computational Economics, volume 3, chapter 9. Elesevier, 2014.

Josep Pijoan-Mas. Precautionary savings or working longer hours? Review of Economic Dynamics, 9(2):326–352, 2006.

Thomas Piketty. Capital in the Twenty-First Century. Harvard University Press, first (english) edition, 2014.

Thomas Piketty. Putting distribution back at the center of economics: Reflections on capital in the twenty-first century. Journal of Economic Perspectives, 29(1):67–88, 2015.

Vicenzo Quadrini. Entrepreneurship, saving and social mobility. Review of Economic Dynamics, 3(1):1–40, 2000.

Juan Pablo Rincón-Zapatero and Manuel Santos. Differentiability of the value function without interiority assumptions. Journal of Economic Theory, 144(5):1948–1964, 2009.

José Víctor Ríos-Rull. Models with heterogenous agents. In T. Cooley, editor, Frontiers in Business Cycle Research, chapter 4. Princeton University Press, 1995.

José Víctor Ríos-Rull. Computation of equilibria in heterogenous agent models. In R. Márimon and A. Scott, editors, Computational Methods for the Study of Dynamic Economies, chapter 11. Oxford University Press, 2 edition, 2001.

Manuel Santos and Adrian Peralta-Alva. Accuracy of simulations for stochastic dynamic models. Econometrica, 66:409–426, 2005.

Manuel Santos and Jesus Vigo-Aguiar. Analysis of a numerical dynamic programming algorithm applied to economic models. Econometrica, 66:409–426, 1998.

John Stachurski. Continuous state dynamic programming via nonexpansive approximations. Computational Economics, 31(2):141–160, 2008.

O. Stenflo. Ergodic theorems for markov chains represented by iterated function systems. Bulletin of the Polish Academy of Sciences: Mathematics, 49:27–43, 2001.

Nancy Stokey, Robert E. Lucas, and Edward C. Prescott. Recursive Methods in Economic Dynamics. Harvard University Press, 1989.

George Tauchen. Finite state markov-chain approximations to univariate and vector autoregressions. Economics Letters, 20:177–181, 1986.

A Pareto Coefficients for Aiyagari (1994) and Pijoan-Mas (2006) models.

Some of the discussion of inequality prefers to focus on Inverse Pareto coefficients, rather than Gini coefficients. Table 7 presents the same results as Table 3 on inequality in the Aiyagari (1994) model in terms of Inverse Pareto coefficients.[30]

Table 7: Interest Rates and Inequality in Aiyagari (1994)
(Inverse) Pareto Coefficients for Earnings, Income, and Wealth

A. Earnings Pareto Coeff/Income Pareto Coeff/Wealth Pareto Coeff ($\sigma = 0.2$)			
ρ/μ	1	3	5
0.0	4.93/2.47/2.24	4.93/2.83/2.44	4.93/2.85/2.46
0.3	4.90/2.23/1.93	4.89/2.34/2.04	4.94/2.48/2.16
0.6	4.91/1.97/1.71	4.93/2.11/1.83	4.93/2.20/1.93
0.9	4.87/1.58/1.41	4.87/1.76/1.57	4.87/1.93/1.72
B. Earnings Pareto Coeff/Income Pareto Coeff/Wealth Pareto Coeff ($\sigma = 0.4$)			
ρ/μ	1	3	5
0.0	2.75/2.22/1.92	2.75/2.34/2.05	2.75/2.44/2.14
0.3	2.75/2.10/1.84	2.75/2.22/1.97	2.75/2.34/2.08
0.6	2.74/1.93/1.72	2.74/2.09/1.87	2.74/2.25/2.03
0.9	2.72/1.56/1.42	2.72/1.71/1.57	2.72/1.84/1.71

Aiyagari (1994) does not report Inverse Pareto coefficients.
Inverse Pareto coefficients are equal to (1+1/Gini)/2.
Uses grid sizes $n_k = 256$, $n_z = 21$, $n_p = 151$

[30] I do not estimate the Inverse Pareto coefficients directly from the steady-state agents distribution. Instead I use the formula that the Inverse Pareto coefficient=(1+1/Gini)/2, to convert the Gini coefficients. This formula is true under the assumption that the distribution of agents is Pareto, but is not the standard way of estimating the Inverse Pareto coefficient from data.