



**HAL**  
open science

## Solving heterogeneous-agent models with parameterized cross-sectional distribution

Yann Algan, Olivier Allais, Wouter J. den Haan

► **To cite this version:**

Yann Algan, Olivier Allais, Wouter J. den Haan. Solving heterogeneous-agent models with parameterized cross-sectional distribution. *Journal of Economic Dynamics and Control*, 2008, 32 (3), pp.875-908. 10.1016/j.jedc.2007.03.007 . halshs-00754295

**HAL Id: halshs-00754295**

**<https://pjse.hal.science/halshs-00754295>**

Submitted on 31 May 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# SOLVING HETEROGENEOUS-AGENT MODELS WITH PARAMETERIZED CROSS-SECTIONAL DISTRIBUTIONS

Yann ALGAN, Oliver ALLAIS, and Wouter J. DEN HAAN\*

March 16, 2007

## Abstract

A new algorithm is developed to solve models with heterogeneous agents and aggregate uncertainty. Projection methods are the main building blocks of the algorithm and—in contrast to the most popular solution procedure—simulations only play a very minor role. The paper also develops a new simulation procedure that not only avoids cross-sectional sampling variation but is 10 (66) times faster than simulating an economy with 10,000 (100,000) agents. Because it avoids cross-sectional sampling variation, it can generate an accurate representation of the whole cross-sectional distribution. Finally, the paper outlines a set of accuracy tests.

*Key Words:* Incomplete markets, numerical solutions, projection methods, simulations

*JEL Classification:* C63, D52

---

\*Algan: Paris School of Economics, University of Marne La Vallée; yann.algan@ens.fr. Allais: INRA, Corela; Olivier.Allais@ivry.inra.fr. Den Haan: University of Amsterdam, London Business School, and CEPR; wdenhaan@uva.nl. The authors would like to thank Chris Carroll, Ken Judd, Michel Juillard, Michael Reiter, and two anonymous referees for useful comments.

# 1 Introduction

Models with heterogeneous agents and aggregate uncertainty are becoming increasingly important. They not only improve the predictions of representative agent models, they also make it possible to study the behavior of sub groups in a general equilibrium framework. Solving such models is difficult, because the set of state variables contains the cross-sectional distribution of agents' characteristics, which is a time-varying infinite dimensional object in the presence of aggregate uncertainty.

The most commonly used algorithm summarizes the cross-sectional distribution with a finite set of moments and calculates the transition law for these state variables using a simulation procedure.<sup>1</sup> This algorithm is relatively easy to implement and has been used to solve a variety of models. Nevertheless, it is important to have alternative algorithms. First, if an alternative algorithm generates a very similar numerical solution then this builds confidence in the generated results. Accuracy checks are of course also helpful in this respect but accuracy tests have its limits. It is especially difficult to test in all possible dimensions if the model has many dimensions. For complex models it is not uncommon that the results of some tests are not that great. If they occur in dimensions that are deemed not that important for the main properties of the model, we typically argue that it is not worth the extra computing time to fix the problem and sometimes it may not even be feasible. The problem is that it is very hard to determine whether inaccuracies in some parts of the model do not spill over to other more important parts or whether tiny but systematic inaccuracies do not accumulate to something more important. It is therefore important if one can replicate results using algorithms that uses different methodologies. The second reason to develop multiple algorithms is that different algorithms have different strenghts and weaknesses and so one type of algorithm may be more suitable for a particular problem.

The algorithm developed in this paper to solve models with heterogenous agents and aggregate uncertainty is quite different from the popular simulations-based algorithm in that simulations only play a minor role. In particular, simulation procedures are not used to obtain the aggregate law of motion. Like Den Haan (1997) and Reiter (2002), we parameterize the cross-sectional distribution, which makes it possible to obtain a numerical solution using standard quadrature and projection techniques. Our algorithm has important efficiency gains over the algorithms developed in these two papers. In particular, we follow Reiter (2002) in using reference moments to get a better characterization of the cross-sectional distribution without increasing the number of state variables, but we do so in a way that is much more

---

<sup>1</sup>Descriptions of this algorithm can be found in Den Haan (1996), Krusell and Smith (1998), and Rios-Rull (1997).

tractable.

Two elements of the algorithm are likely to be useful in other applications as well. First, to obtain reference moments we need to simulate the economy and we develop a simulation procedure that avoids cross-sectional sampling uncertainty. Second, we propose a particular class of parameterizing densities that makes the problem of finding the coefficients that correspond to a set of moments a convex optimization problem. Using this class of functions avoids the need for good starting values, whereas for other functional forms we found this to be a major concern. Linking moments with a parameterized density is part of our simulation procedure and also plays a key role in the solution algorithm itself.

The standard simulation procedure constructs a panel of  $N_T$  observations and  $N_N$  agents. By parameterizing the cross-sectional distribution and using quadrature integration, however, our procedure can generate an accurate simulation with a continuum of agents. Note that models with a large number of heterogeneous agents almost always assume a continuum of agents, so that the law of large numbers ensures that idiosyncratic risk is averaged out. In fact, the assumption of a continuum of agents plays a key role, not only in the specification of the state variables and the definition of the equilibrium, but also in the construction of most algorithms.<sup>2</sup> Thus, by simulating with a continuum instead of a finite number of agents, we stay much closer to the actual model being solved.

Our simulation procedure not only avoids cross-sectional sampling variation, it is also much cheaper. We found that simulating an economy with 10,000 agents for 1000 periods took 13 times as long as simulating the same economy with a continuum of agents. In our own algorithm, the simulation procedure only plays a very minor role. This is, of course, no reason not to use a more accurate simulation procedure. For algorithms that use a simulation procedure to calculate the aggregate law of motion, however, the improved simulation procedure will have bigger benefits.

The rest of this paper is organized as follows. The next section describes the production economy of Krusell and Smith (1998). Section 3 briefly discusses existing algorithms and summarizes the contributions of this paper. Section 4 describes the algorithm in detail and Section 5 describes the simulation procedure. Section 6 discusses how to check for accuracy and reports the results. The last section concludes.

---

<sup>2</sup>In particular, a crucial property being used is that conditional on realizations of the aggregate shock and this period's cross-sectional distribution, next period's cross-sectional distribution is known with certainty.

## 2 The production economy

The economy is a production economy with aggregate shocks in which agents face different employment histories and partially insure themselves through (dis)saving in capital. For more details see Krusell and Smith (1998).

**Problem for the individual agent.** The economy consists of a unit mass of ex ante identical households. Each period, agents face an idiosyncratic shock  $\varepsilon$  that determines whether they are employed,  $\varepsilon = 1$ , or unemployed,  $\varepsilon = 0$ . An employed agent earns a wage rate of  $\varpi_t$ . An employed agent earns an after-tax wage rate of  $(1 - \tau_t)\varpi_t$  and an unemployed agent receives unemployment benefits  $\mu\varpi_t$ . Markets are incomplete and the only investment available is capital accumulation. The net rate of return on this investment is equal to  $r_t - \delta$ , where  $r_t$  is the rental rate and  $\delta$  is the depreciation rate. Agent's  $i$  maximization problem is as follows:

$$\begin{aligned} \max_{\{c_t^i, k_{t+1}^i\}_{t=0}^{\infty}} E \sum_{t=0}^{\infty} \beta^t \frac{(c_t^i)^{1-\gamma} - 1}{1-\gamma} \\ \text{s.t. } c_t^i + k_{t+1}^i = r_t k_t^i + (1 - \tau_t)\varpi_t \bar{l} \varepsilon_t^i + \mu\varpi_t(1 - \varepsilon_t^i) + (1 - \delta)k_t^i \\ k_{t+1}^i \geq 0 \end{aligned} \tag{1}$$

Here  $c_t^i$  is the individual level of consumption,  $k_t^i$  is the agent's beginning-of-period capital, and  $\bar{l}$  is the time endowment.

**Firm problem.** Markets are competitive and the production technology of the firm is characterized by a Cobb-Douglas production function. Consequently, firm heterogeneity is not an issue. Let  $K_t$  and  $L_t$  stand for per capita capital and the employment rate, respectively. Per capita output is given by

$$Y_t = a_t K_t^\alpha (\bar{l} L_t)^{1-\alpha} \tag{2}$$

and prices by

$$\varpi_t = (1 - \alpha) a_t \left(\frac{K_t}{\bar{l} L_t}\right)^\alpha \tag{3}$$

$$r_t = \alpha a_t \left(\frac{K_t}{\bar{l} L_t}\right)^{\alpha-1} \tag{4}$$

Aggregate productivity,  $a_t$ , is an exogenous stochastic process that can take on two values,  $1 - \Delta^a$  and  $1 + \Delta^a$ .

**Government** The only role of the government is to tax employed agents and to redistribute funds to the unemployed. We assume that the government’s budget is balanced each period. This implies that the tax rate is equal to

$$\tau_t = \frac{\mu u_t}{\bar{l} L_t}. \quad (5)$$

where  $u_t = 1 - L_t$  denotes the unemployment rate in period  $t$ .

**Exogenous driving processes.** There are two stochastic driving processes. The first is aggregate productivity and the second is the employment status. Both are assumed to be first-order Markov processes. We let  $\pi_{aa'\varepsilon\varepsilon'}$  stand for the probability that  $a_{t+1} = a'$  and  $\varepsilon_{t+1}^i = \varepsilon'$  when  $a_t = a$  and  $\varepsilon_t^i = \varepsilon'$ . These transition probabilities are chosen such that the unemployment rate can take on only two values. That is,  $u_t = u^b$  when  $a_t = a^b$  and  $u_t = u^g$  when  $a_t = a^g$  with  $u^b > u^g$ .

**Equilibrium.** Krusell and Smith (1998) consider recursive equilibria in which the policy functions of the agent depend on his employment status,  $\varepsilon^i$ , his beginning-of-period capital holdings,  $k^i$ , aggregate productivity,  $a$ , and the cross-sectional distribution of capital holdings. An equilibrium consists of the following.

- Individual policy functions that solve the agent’s maximization problem.
- A wage and a rental rate that are determined by 3 and 4, respectively.
- A transition law for the cross-sectional distribution of capital, that is consistent with the investment policy function.

### 3 Relation to existing algorithms

In this paper we give a brief overview of existing algorithms and then highlight the contributions of this paper.

#### 3.1 Existing algorithms

A standard aspect of numerical algorithms that solve models with heterogeneous agents is to summarize the infinite-dimensional cross-sectional distribution of agents’ characteristics by a finite set of moments,  $m$ . The transition law is then a mapping that generates next period’s moments,  $m'$ , given the values of the moments in the current period,  $m$ , and the realization of the aggregate shock. Den Haan (1996),

Krusell and Smith (1998), and Rios-Rull (1997) propose to calculate this transition law as follows. First, construct a time series for the cross-sectional moments by simulating an economy with a large but finite number of agents. Second, regress the simulated moments on the set of state variables. The first disadvantage of the simulation procedure is that moments are calculated using Monte Carlo integration, which is known to be an inefficient numerical integration procedure.<sup>3</sup> The second disadvantage is that the observations are clustered around the mean, since they are taken from a simulated series. An efficient projection procedure, however, requires the explanatory variables to be spread out, for example, by using Chebyshev nodes.<sup>4</sup> Consequently, the use of simulation procedures may make it expensive (in terms of computing time) to obtain an accurate solution, especially if one wants the solution to be accurate across the whole state space.

Den Haan (1997) parameterizes the cross-sectional distribution with a flexible functional form,  $P(k; \rho)$ , which makes it possible to use quadrature techniques to do the numerical integration. In addition, his algorithm uses Chebyshev nodes to construct a grid of explanatory variables for the projection step. The coefficients of the approximating density,  $\rho$ , are pinned down by the set of moments used,  $m$ . The disadvantage of Den Haan (1997) is that the shape of the distribution is completely pinned down by the moments used as state variables and the class of flexible functional forms used. Consequently, a large number of state variables may be needed to get the shape of the cross-sectional distribution right even when just a few moments actually matter for agents' behavior. Another drawback of Den Haan (1997) is that an inefficient procedure is used to find the coefficients of the approximating density. Reiter (2002) improves upon the algorithm of Den Haan (1997) in an ingenious way by letting the shape of the distribution depend not only on the moments used as state variables,  $m$ , but also on a set of reference moments that are obtained by a simulation procedure.

Promising recent alternatives to the standard algorithm have been developed in Preston and Roca (2006) and Reiter (2006). Reiter (2006) first solves a model without aggregate uncertainty using standard projection procedures. Next, by replacing the endogenous variables in the equations of the model with the parameterized numerical solution, he obtains a difference equation in the numerical coefficients. Then he uses perturbation techniques to solve for the sensitivity of the numerical solution to aggregate shocks. This is quite a different approach than the procedure used here, which is good for the profession, because the more variety among available approaches used the better. Preston and Roca (2006) use a "pure" perturbation

---

<sup>3</sup>See Judd (1998).

<sup>4</sup>Note that in the classic regression problem the standard errors,  $\sigma^2(X'X)^{-1}$ , are also lower when the explanatory values are more spread out.

method to solve the model.<sup>5</sup> Perturbation methods are likely to work well when the distribution needs to be characterized by many statistics, because dealing with many arguments is the strength of perturbation methods.

### 3.2 The contributions of this paper

The main contributions of this paper are the following.

**Calculating the transition law of the cross-sectional distribution.** The disadvantage of Reiter (2002) is that the particular implementation of the idea of reference moments is very cumbersome.<sup>6</sup> As in Den Haan (1997) and Reiter (2002), this paper develops a procedure to calculate this transition law without relying on simulation procedures to calculate moments and to carry out the projection step. As in Reiter (2002), it uses reference moments, but the modifications introduced make the procedure much more straightforward to implement. In other words, the algorithm is an important improvement over earlier attempts that use projection methods to solve models with heterogeneous agents. Moreover, because the building blocks are so different from the simulation procedures, it provides a constructive alternative.

**Calculating the approximating density for given moments.** The algorithm links a set of moments with a parameterized density. Consequently, an important part of the algorithm is the mapping between the set of moments and the coefficients of the density. One possibility would be to use an equation solver that chooses the set of coefficients so that the moments of the parameterized density are equal to the specified moments. We found this procedure to be slow and intermittently breaking down. By choosing a particular choice for the basis functions of the approximating

---

<sup>5</sup>They replace the inequality constraint with a penalty function, since perturbation methods cannot deal well with the kind of inequality constraint used here. This also accomplishes that agents do not have negative capital holdings and try to stay away from low capital stocks.

<sup>6</sup>Reiter (2002) constructs a reference density  $G(m)$ , which relates the shape of the distribution to the set of moments that serve as state variables. It is a weighted average of distributions from a simulated economy, where distributions with moments closer to  $m$  get more weight. Step functions are used to construct a reference distribution, which has the advantage of being very flexible but has the disadvantage of using a lot of parameters. One problem of the approach in Reiter (2002) is that the moments of the reference density may not be equal to  $m$ . This means that one first has to apply operations to ensure that one obtains a new reference function  $\tilde{G}(m)$  for which this is not the case. But even if  $m$  contains only first and second moments, then this problem entails more than a linear transformation, since  $\tilde{G}(m)$  has to be a step function that conforms to the specified grid and cannot violate the constraints on the support of the distribution, such as the constraint that  $k \geq 0$ .

polynomial we transform this problem into a convex optimization problem, for which reliable convergence algorithms exist. This procedure is likely to be useful outside the literature of numerical solution techniques, since characterizing a cross-sectional distribution with a CDF from a class of flexible functional forms is a common problem in econometrics.

**Simulating a panel without cross-sectional sampling variation.** This paper develops a procedure to simulate an economy without cross-sectional sampling variation. Standard procedure is to simulate data using a finite number of agents and a finite number of time periods, which means that the outcome depends on the particular random draw used. Sampling variation disappears at a slow rate and could be especially problematic if the number of a particular type of agent is small relative to the total number of agents.

Existing models with a large number of heterogeneous agents typically assume that there is a continuum of agents. This implies that *conditional on the realization of the aggregate shock* there is *no* cross-sectional sampling variation, a property that plays a key role in the definition of the set of state variables and the definition of the recursive equilibrium. The simulation procedure developed in this paper sticks to the assumption of the model and uses a continuum of agents. Moreover, because the procedure avoids cross-sectional sampling variation, it can obtain an accurate description of aspects of the distribution such as behavior in the tails much more easily than simulation methods.

**Accuracy tests.** It is never trivial to check the accuracy of a numerical solution for dynamic stochastic models, since the true solution is not known. Checking for accuracy is made especially difficult, because there are many aspects to the solution of this type of model. In this paper, we discuss several tests to evaluate the accuracy of the solution for a model with heterogeneous agents.

## 4 The algorithm

In this section, we discuss the different steps of the algorithm. In Section 4.1, we start with a discussion of the state variables used, followed by an overview of the algorithm in Section 4.2. The remaining sections describe the steps of the algorithm.

## 4.1 State variables and transition laws

Krusell and Smith (1998) consider a recursive equilibrium in which the policy functions of the agent depend on his employment status,  $\varepsilon^i$ , his beginning-of-period capital holdings,  $k^i$ , aggregate productivity,  $a$ , and the cross-sectional distribution of capital holdings.<sup>7</sup> Let  $\overleftarrow{f}^w(k)$  be the cross-sectional distribution of beginning-of-period capital holdings for agents with employment status  $w \in \{e, u\}$  and  $k \geq 0$ . The arrow pointing left indicates that the cross-sectional distribution refers to the distribution at the beginning of the period (but after all shocks are observed, that is once individual employment shocks and aggregate shocks had taken place). Similarly,  $\overrightarrow{f}^w(k)$  refers to the distribution at the end of the period. The following two steps determine the transition law that links the current-period distribution,  $\overleftarrow{f}^w(k)$ , with next period's distribution,  $\overleftarrow{f}^{w'}(k)$ .

- The end-of-period distribution is determined by  $a$ ,  $\overleftarrow{f}^e$ ,  $\overleftarrow{f}^u$ , and the individual investment function. That is,  $\overrightarrow{f}^e = \overrightarrow{\Upsilon}^e(a, \overleftarrow{f}^e, \overleftarrow{f}^u)$  and  $\overrightarrow{f}^u = \overrightarrow{\Upsilon}^u(a, \overleftarrow{f}^e, \overleftarrow{f}^u)$ .
- Next period's beginning-of-period distribution,  $\overleftarrow{f}^{w'}(k)$ , is determined by the end-of-period distribution and the employment-status flows corresponding to the values of  $a$  and  $a'$ . Thus,  $\overleftarrow{f}^{e'} = \overleftarrow{\Upsilon}^e(a, a', \overrightarrow{f}^e, \overrightarrow{f}^u)$  and  $\overleftarrow{f}^{u'} = \overleftarrow{\Upsilon}^u(a, a', \overrightarrow{f}^e, \overrightarrow{f}^u)$ .  $\overleftarrow{\Upsilon}^e(\cdot)$  and  $\overleftarrow{\Upsilon}^u(\cdot)$  are simple functions that are determined directly by the transition probabilities.<sup>8</sup>

An alternative to using the cross-sectional distribution of employment and capital holdings is to use all past realizations of the aggregate shocks.<sup>9</sup> For the model considered here, we found that a large number of lags is needed. Nevertheless, if one doesn't have a complete description of the cross-sectional distribution, it still may be worthwhile to add some lagged values of  $a$ .<sup>10</sup> In our algorithm we, therefore, add the lagged value of  $a$  as a state variable. But there is another reason, which will become evident in the remainder of this section.

---

<sup>7</sup>Miao (2006) shows the existence of a recursive equilibrium, but also uses expected payoffs as state variables. It is not clear whether a recursive equilibrium exists when the smaller set of state variables is used. For a numerical solution this is less important in the sense that approximation typically entails not using all information.

<sup>8</sup>Details are given in A.1.

<sup>9</sup>This is the approach used in Veracierto (2002). He solves a model with irreversible investment in which the cross-sectional distribution matters because the investment decision is of the (S,s) variety. Instead of keeping track of the cross-sectional distribution, he keeps track of a history of lower and upper threshold levels.

<sup>10</sup>The numerical cost is fairly low, since  $a$  can take on only two values.

To deal with the infinite dimension of the cross-sectional distribution, we follow Den Haan (1996, 1997), Krusell and Smith (1998), and Rios-Rull (1997) and describe the cross-sectional distribution with a finite set of moments. The remainder of this section discusses in detail which moments we use.

In this model, agents face a borrowing constraint,  $k \geq 0$ . We, therefore, include the fraction of agents of each type that start the period with zero capital holdings,  $\overleftarrow{m}^{e,c}$  and  $\overleftarrow{m}^{u,c}$ . Employed agents never choose a zero capital stock. This means that the density  $\overrightarrow{f}^e$  does not have mass at zero. In contrast,  $\overleftarrow{f}^e$  does have mass at zero, because some of the agents that are employed in the current period were unemployed in the last period and chose a zero capital stock. Both  $\overleftarrow{m}^{e,c}$  and  $\overleftarrow{m}^{u,c}$  can be easily calculated from  $\overrightarrow{m}_{-1}^{u,\hat{c}}$  and the employment-status flows corresponding to the values of  $a$  and  $a'$ . Thus, instead of using  $[a, \overleftarrow{m}^{e,c}, \overleftarrow{m}^{u,c}]$  we also can use  $[a, a_{-1}, \overrightarrow{m}_{-1}^{u,\hat{c}}]$ . We prefer to use  $[a, a_{-1}, \overrightarrow{m}_{-1}^{u,\hat{c}}]$ , because  $a_{-1}$  can take on only two values and is, thus, computationally an inexpensive state variable. Moreover, as explained above  $a_{-1}$  could have predictive value that goes beyond the ability to determine  $\overleftarrow{m}^{e,c}$  and  $\overleftarrow{m}^{u,c}$ .

In addition, the algorithm uses centralized moments of the distributions of strictly positive capital holdings. The set of moments that are used as state variables are stored in the following vector

$$m = \left[ \overrightarrow{m}_{-1}^{u,\hat{c}}, \overleftarrow{m}^{e,1}, \dots, \overleftarrow{m}^{e,N_M}, \overleftarrow{m}^{u,1}, \dots, \overleftarrow{m}^{u,N_M} \right],$$

where  $\overleftarrow{m}^{w,j}$  is the  $j^{\text{th}}$ -order centralized moment for workers with employment status  $w$  and strictly positive capital holdings. The dimension of the vector is  $N_M^* = 2N_M + 1$ .

The aggregate state is thus given by  $s = [a, a_{-1}, m]$ . Since we only use a limited set of moments as state variables, the transition law only needs to specify how this limited set of moments evolve over time. Thus, instead of calculating  $\overrightarrow{\Upsilon}^e(\cdot)$  and  $\overrightarrow{\Upsilon}^u(\cdot)$ , we now calculate  $[\overrightarrow{m}_{-1}^{u,\hat{c}}, \overleftarrow{m}^{u,1}, \dots, \overleftarrow{m}^{u,N_M}] = \overrightarrow{\Gamma}_n^u(s; \psi_n^{\Gamma^u})$  and  $[\overleftarrow{m}^{e,c}, \overleftarrow{m}^{e,1}, \dots, \overleftarrow{m}^{e,N_M}] = \overrightarrow{\Gamma}_n^e(s; \psi_n^{\Gamma^e})$ , where  $\overrightarrow{\Gamma}_n^w(s)$  is an  $n^{\text{th}}$ -order polynomial with coefficients  $\psi_n^{\Gamma^w}$ . To simplify the notation we will typically write  $\overrightarrow{\Gamma}^w(s)$ , but one should keep in mind that this is an approximating function with coefficients that are determined by the algorithm. In the implementation of the algorithm, we set—as in Krusell and Smith (1998)— $N_M$  equal to 1.

## 4.2 Overview

An important part of this algorithm is to avoid Monte Carlo integration by approximating the densities  $\overleftarrow{f}^e$  and  $\overleftarrow{f}^u$  with a flexible functional forms. To determine this functional form, we use the moments that are used as state variables,  $m$ , as well as some additional information that we will refer to as "reference moments". The reference moments are higher-order moments that are helpful in getting the shape of the distribution right.

The algorithm uses the following iterative procedure to solve the model.

- Given transition laws,  $\overrightarrow{\Gamma}^e(s)$  and  $\overrightarrow{\Gamma}^u(s)$ , solve for  $c(\varepsilon, k, s)$  and  $k'(\varepsilon, k, s)$ . This is discussed in Section 4.3.
- Use the solutions for the individual policy functions,  $c(\varepsilon, k, s)$  and  $k'(\varepsilon, k, s)$  to obtain information about the "reference moments". This is discussed in Section 4.4.
- Given solutions for the individual policy functions,  $c(\varepsilon, k, s)$  and  $k'(\varepsilon, k, s)$ , solve for  $\overrightarrow{\Gamma}^e(s)$  and  $\overrightarrow{\Gamma}^u(s)$ . This is discussed in Section 4.6. This requires setting up a grid of the aggregate state variables  $[a, a_{-1}, m]$  and a procedure to link the values of the moments  $m$  and the reference moments with an explicit cross-sectional density. This procedure is discussed in Section 4.5.
- Iterate until the transitions laws used to solve for the individual policy functions are close to the transition laws implied by the individual policy functions.

## 4.3 Procedure to solve for individual policy functions

The procedure to solve for individual policy functions relies on standard projection methods, except that we modify the standard procedure to deal with the inequality constraint on capital. In this section, we describe how to solve for the individual policy rules taking the aggregate policy rules  $\overrightarrow{\Gamma}^e(s)$  and  $\overrightarrow{\Gamma}^u(s)$  as given. The first-order conditions of the agent are given by<sup>11</sup>

$$\begin{aligned} c(\varepsilon, k, s)^{-\gamma} &= \mathbf{E} [\beta c(\varepsilon', k', s')^{-\gamma} (1 + r(s'))] \text{ for } k' > 0, \\ c(\varepsilon, k, s)^{-\gamma} &\geq \mathbf{E} [\beta c(\varepsilon', k', s')^{-\gamma} (1 + r(s'))] \text{ for } k' = 0, \text{ and} \end{aligned} \quad (6)$$

$$c + k' = r(s)k + w(s)\bar{l}\varepsilon + (1 - \delta)k, \quad (7)$$

---

<sup>11</sup>We have suppressed the  $i$  superscript for notational convenience.

In this system  $w(s)$  and  $r(s)$  only depend on  $a$  and the aggregate capital stock and can be solved directly from Equations 3 and 4. The conditional expectation in Equation 6 is a function of the individual and aggregate state variables. To solve the individual problem we approximate this conditional expectation with a flexible functional form. That is,

$$\text{E} [\beta c(\varepsilon', k', s')^{-\gamma} (1 + r(s'))] \approx \Psi_n(k, \varepsilon, s; \psi_n^E), \quad (8)$$

where  $\Psi_n(\cdot)$  is an  $n^{\text{th}}$ -order polynomial and  $\psi_n^E$  its coefficients. Let  $\bar{k}(\varepsilon, s)$  be the capital stock such that

$$k' = 0 \quad \text{if } k \leq \bar{k}(\varepsilon, s). \quad (9)$$

Then  $\partial k' / \partial k = 0$  for  $k < \bar{k}(\varepsilon, s)$ . This implies that  $\text{E}[\beta c(\varepsilon', k', s')^{-\gamma} (1 + r(s'))]$  as a function of  $k$  is flat for  $k < \bar{k}$  and nondifferentiable at  $k = \bar{k}$ . When  $k < \bar{k}(\varepsilon, s)$  one does not need the approximation  $\Psi_n(\cdot)$ , since  $k' = 0$  and one can solve consumption from the budget constraint. To calculate the approximation for the conditional expectation, we only use grid points at which  $k' > \bar{k}(\varepsilon, s)$ . This means that the grid is no longer fixed within the algorithm and we loose some of the optimality properties of using Chebyshev grid points, but we found that with this procedure we can obtain more accurate solutions.

Besides this modification, our procedure to solve for the individual policy rules is a standard application of projection methods as discussed in Judd (1992). In particular, we use the following procedure.

- Construct a grid for the values of individual and aggregate state variables.
- Use  $\bar{\psi}$  as the initial value for  $\psi_n^E$ . Given the value  $\bar{\psi}$ , it is straightforward to solve for  $c(\varepsilon, k, s)$  and  $k'(\varepsilon, k, s)$  from the first-order condition and the budget constraint.
- At each grid point calculate  $k'$ ,
- For all possible realizations of  $a'$  and  $\varepsilon'$  calculate  $\beta c(\varepsilon', k', s')^{-\gamma} (1 + r(s'))$ . This requires calculating  $m'$  but this is easy since  $\vec{\Gamma}^e$  and  $\vec{\Gamma}^u$  are given.<sup>12</sup> Next, calculate  $\text{E}[\beta c(\varepsilon', k', s')^{-\gamma} (1 + r(s'))]$  by weighting the possibly outcomes with the probabilities.
- Perform a projection to obtain a new estimate for  $\psi_n^E, \hat{\psi}$ .
- Use a weighted average of  $\hat{\psi}$  and  $\bar{\psi}$  as a new initial value for  $\psi_n^E$ .

---

<sup>12</sup>For details see A.1.

- Iterate until the coefficients have converged.<sup>13</sup>

#### 4.4 Procedure to generate reference moments

The reference moments are used to ensure that the functional form of the cross-sectional distribution is appropriate without using too many moments explicitly as state variables. Note that an extra state variable increases the dimension of the grid and the set of arguments of each function, whereas an extra reference moment does not. Given the complexity of the system one has to rely on simulations to obtain information about the shape of the distribution. Thus, we also use a simulation procedure to obtain reference moments, but we propose a new simulation procedure that reduces the amount of sampling variation. This new simulation procedure is discussed in Section 5.

The simulation generates a time series with for each period a set of observations  $\overleftarrow{m}_t^{w,j}$  for  $w \in \{e, u\}$  and  $j \in \{N_M + 1, \dots, N_{\overline{M}}\}$ . The simplest way to proceed would be to use the sample averages as the reference moments, but we let the reference moments depend on  $a$ .<sup>14</sup>

#### 4.5 Procedure to find cross-sectional distribution

At each grid point, we know the values of  $a$  and  $a_{-1}$  as well as the values of  $\overleftarrow{m}^{w,j}$  for  $w \in \{e, u\}$  and  $j \in \{c, 1, \dots, N_M\}$ . We also have a set of higher-order reference moments  $\overleftarrow{m}^{w,j}$ ,  $w \in \{e, u\}$  and  $j \in \{N_M + 1, \dots, N_{\overline{M}}\}$ . Let  $P(k; \rho^w)$  be the exponential of a polynomial of order an  $N_{\overline{M}}$  with coefficients  $\rho^w$ . One way to solve for

---

<sup>13</sup>As pointed out by Reiter (2006), one doesn't need the law of motion for  $m'$  to solve the individual problem. Using the parameterized cross-sectional distribution, one could in principle use quadrature methods to directly calculate the values of  $m'$  when needed. By doing this one could in each iteration not only update the individual policy rules but also the law of motion for the aggregate state variables. This is likely to speed up the algorithm *if* it is on course towards the fixed point, but the simultaneous updating might make the algorithm less stable.

<sup>14</sup>Without complicating the algorithm, one could let the higher-order moments depend on all the elements of the aggregate state variables,  $s$ , that is

$$\overleftarrow{m}^{w,j} = \Phi_n^w(s), \tag{10}$$

where  $\Phi_n^w$  is an  $n^{\text{th}}$ -order flexible functional form. One can obtain the coefficients of  $\Phi_n^w$  by a simple regression and at each aggregate grid point it is trivial to use  $\Phi_n^w$  to determine the set of reference moments. We have not done so, because—as pointed out by Young (2005)—higher-order moments do not exhibit a clear relationship *on average* to lower-order moments.

$\rho^w$  is to solve the following system of  $N_{\overline{M}} + 1$  equations and unknowns.

$$\begin{aligned}
& \int_0^{\infty} \left[ k - \overleftarrow{m^{w,1}} \right] P(k; \rho^w) dk = 0 \\
& \int_0^{\infty} \left[ (k - \overleftarrow{m^{w,1}})^2 - \overleftarrow{m^{w,2}} \right] P(k; \rho^w) dk = 0 \\
& \qquad \qquad \qquad \dots \\
& \int_0^{\infty} \left[ (k - \overleftarrow{m^{w,1}})^{N_{\overline{M}}} - \overleftarrow{m^{w,N_{\overline{M}}}} \right] P(k; \rho^w) dk = 0 \\
& \int_0^{\infty} P(k; \rho^w) dk = 1
\end{aligned} \tag{11}$$

This is a non-linear system of equations and without additional restrictions on  $P(k, \rho^w)$ , there are no algorithms with reliable convergence properties. It is feasible to solve this system, but it required in our experience a lot of work in finding good initial conditions. By adopting a particular class of approximating polynomials one can reduce this problem to a convex optimization problem for which convergence is no longer problematic.

Our alternative procedure parameterizes the density with a regular polynomial, but the basis functions are constructed in a specific way. In particular, the polynomial of order  $N_{\overline{M}}$  is written as:

$$\rho_0^w \exp \left( \begin{array}{c} P(k, \rho^w) = \\ \rho_1^w \left[ k - \overleftarrow{m^{w,1}} \right] + \\ \rho_2^w \left[ \left( k - \overleftarrow{m^{w,1}} \right)^2 - \overleftarrow{m^{w,2}} \right] + \dots + \\ \rho_{N_{\overline{M}}}^w \left[ \left( k - \overleftarrow{m^{w,1}} \right)^{N_{\overline{M}}} - \overleftarrow{m^{w,N_{\overline{M}}}} \right] \end{array} \right). \tag{12}$$

When the density is constructed in this particular way, the coefficients, except for  $\rho_0^w$ , can be found with the following minimization routine:

$$\min_{\rho_1^w, \rho_2^w, \dots, \rho_{N_{\overline{M}}}^w} \int_0^{\infty} P(k, \rho^w) dk. \tag{13}$$

The first-order conditions correspond exactly to the first  $N_{\overline{M}}$  equations in (11).  $\rho_0^w$  does not appear in these equations, but  $\rho_0^w$  is determined by the condition that the density integrates to one.

The Hessian (times  $\rho_0^w$ ) is given by

$$\int_0^\infty X \left( \overleftarrow{m}^{w,1}, \dots, \overleftarrow{m}^{w,N_M} \right) X \left( \overleftarrow{m}^{w,1}, \dots, \overleftarrow{m}^{w,N_M} \right)' P(k, \rho^w) dk, \quad (14)$$

where  $X$  is an  $(N_M \times 1)$  vector and the  $i^{\text{th}}$  element is given by

$$\begin{aligned} & \left( k - \overleftarrow{m}^{w,1} \right) \quad \text{for } i = 1 \\ & \left( k - \overleftarrow{m}^{w,1} \right)^i - \overleftarrow{m}^{w,i} \quad \text{for } i > 1 \end{aligned} \quad (15)$$

The Hessian is positive semi-definite since  $X$  does not depend on  $\rho^w$ .<sup>15</sup> Consequently, this is a convex optimization problem. Our minimization routine (BFGS) does not exploit the fact that this is a convex optimization problem. Even without exploiting the convexity in the minimization routine we found that finding a solution for  $\rho^w$  was an enormous improvement. Most importantly, it avoided the need for good initial conditions. When using exponentials to parameterize the density and using (11) to find the coefficients, the algorithm often got stuck and had to be restarted with better initial conditions. Even when good initial conditions were given then the convex problem made it possible to find a solution much faster (even though the minimization routine didn't exploit the convexity).<sup>16</sup>

## 4.6 Procedure to solve for aggregate laws of motion

The procedure is characterized by the following steps.

- We construct a grid with values for  $a$ ,  $a_{-1}$ ,  $\overrightarrow{m}_{-1}^{u,c}$ , and  $\overleftarrow{m}^{w,j}$  for  $w \in \{e, u\}$  and  $j \in \{1, \dots, N_M\}$ . Here  $\overrightarrow{m}_{-1}^{u,c}$  is the fraction of unemployed agents that chose  $k' = 0$  last period and  $\overleftarrow{m}^{w,j}$  is the  $j^{\text{th}}$  moment of the distribution of strictly positive capital holdings for agents with employment status  $w$ . Given values for  $a$ ,  $a_{-1}$ , and  $\overrightarrow{m}_{-1}^{u,c}$ , we can calculate  $\overleftarrow{m}^{e,c}$  and  $\overleftarrow{m}^{u,c}$ . The grid values for this period's and last period's aggregate state are the two possible realizations and

<sup>15</sup>Note that evaluated at the solution for  $\rho^w$ , it is a covariance matrix.

<sup>16</sup>In particular, we tried finding the coefficients using standard exponentials, i.e.,  $\rho_0 \exp(\rho_1 k + \rho_2 k^2 + \dots)$ , and using our preferred functional forms. The best way to find the coefficients of the exponential—in our experience—is to solve (??) using a Newton procedure. As starting values we used in both procedures the solution to a slightly perturbed set of values for the target moments. With the Newton procedure it took on average 57 seconds to find the solution, whereas with our preferred procedure it took on average 0.24 seconds, more than 200 times as fast.

we use Chebyshev nodes to locate the grid points for the other state variables. These are the "x-values".

- Using quadrature methods, we calculate end-of-period moments,  $\overline{m^{w,j}}$  for  $j \in \{c, 1, \dots, N_M\}$ , at each grid point. These are the "y-values". The parameterization of the cross-sectional distribution discussed in the last section makes it possible to use Simpson quadrature to calculate end-of-period moments.
- Using the y-values and the x-values, we perform a projection step to find the coefficients of the approximating function,  $\overrightarrow{\Gamma}_n^e(s; \psi_n^{\Gamma^e})$  and  $\overrightarrow{\Gamma}_n^u(s; \psi_n^{\Gamma^u})$ .

## 5 Simulating a panel with a non-random cross-section

An important contribution of this paper is to develop a simulation procedure that eliminates the amount of cross-sectional sampling variation in the simulation of a panel. Simulations play a relatively minor role in the algorithm proposed here and are only used to get information on the shape of the cross-sectional distribution. But this simulation procedure can also be used in algorithms that rely on simulations to determine the law of motion of the aggregate state variables, such as, the algorithm used in Krusell and Smith (1998). This section discusses the new procedure. Accuracy tests for this procedure are proposed in the next section.

The idea of the simulation procedure proposed is to stay close to the idea that there is a continuum of agents in the model. This implies that—conditional on the realization of the aggregate shock—there is no cross-sectional sampling variation with our simulation procedure, just as there is none in the true solution. The standard procedure in the literature uses a large but finite number of agents,  $N_N$ . Since sampling variation only vanishes at rate  $\sqrt{N_N}$  one may need a high value of  $N_N$  to get accurate moments and indeed quite high values are used in the literature (between 10,000 and 100,000). Having an accurate procedure to simulate a cross-sectional distribution is also useful if one is interested in a particular region of the distribution. If the simulated series do not visit these regions frequently then one would not accurately describe what happens there even if the value of  $N_N$  is high enough to accurately describe the moments used as state variables.

The procedure works as follows

- Use a random number generator to draw a time series for the aggregate productivity shock,  $\{a_t\}_{t=0}^{N_T}$ . Although the procedure eliminates the cross-sectional

sampling variation, there is still sampling variation due to the stochastic nature of the aggregate productivity shock.

- In period 1, the procedure starts with the following information. First, the fractions of employed and unemployed agents with zero capital stock at the beginning of the period,  $\overleftarrow{m}_1^{e,c}$ , and  $\overleftarrow{m}_1^{u,c}$ . Second,  $N_{\widetilde{M}}$  centralized moments of the distribution of strictly positive beginning-of-period capital holdings for the unemployed and the employed,  $\overleftarrow{m}_1^{w,j}$  for  $w \in \{e, u\}$  and  $j \in \{1, \dots, N_{\widetilde{M}}\}$ . Since this procedure is relatively cheap and not part of a complex fixed-point calculation one can set  $N_{\widetilde{M}}$  fairly high. We set  $N_{\widetilde{M}} = N_{\overline{M}} = 6$ , but these parameters do not have to be equal to each other.
- The moments  $\overleftarrow{m}_1^{w,j}$  for  $w \in \{e, u\}$  and  $j \in \{1, \dots, N_{\widetilde{M}}\}$  determine the densities of positive capital holdings for the employed and unemployed,  $P(k; \rho_1^e)$  and  $P(k; \rho_1^u)$ . That is, using the procedure discussed in Section 4.5, we find the coefficients of the densities in period 1,  $\rho_1^e$  and  $\rho_1^u$ , so that the moments of  $P(k; \rho_1^e)$  and  $P(k; \rho_1^u)$  correspond to the specified moments.
- Use  $P(k; \rho_1^e)$ ,  $P(k; \rho_1^u)$ ,  $\overleftarrow{m}_1^{e,c}$ , and  $\overleftarrow{m}_1^{u,c}$ , i.e., the distribution of beginning-of-period capital holdings together with the individual policy rules to calculate the end-of-period moments,  $\overrightarrow{m}_1^{w,j}$  for  $w \in \{e, u\}$  and  $j \in \{c, 1, \dots, N_{\widetilde{M}}\}$ , and  $\overrightarrow{m}_1^{u,c}$ . We use Simpson quadrature to do the integration.
- Use the values of the productivity shocks in period 1 and 2, i.e.,  $a_1$  and  $a_2$ , together with the end-of-period moments for period 1 to calculate beginning-of-period moments for period 2,  $\overleftarrow{m}_2^{w,j}$  and  $\overleftarrow{m}_2^{w,c}$  for  $w \in \{e, u\}$  and  $j \in \{c, 1, \dots, N_{\widetilde{M}}\}$ . Recall that this simply takes care of the effects of changes in the employment status on the cross-sectional distribution. Details are given in Appendix A.1.
- Use the procedure discussed in 4.5 to find the values for  $\rho_2^e$  and  $\rho_2^u$ .
- Repeat the procedure for the next period until  $t = N_T$ .

To ensure that the sample used to obtain information about the cross-sectional distribution has reached (or is at least close to) its ergodic distribution one should disregard an initial set of observations. For the particular model we study in this paper, we found that if the initial distribution is not close to the ergodic set, then one has to disregard a large number of initial observations, since it can take quite a

while before the economy has reached the ergodic distribution. After some experimentation, one has a good idea about a reasonable initial distribution and then this is less of a problem.

For some policy functions, it may be the case that some higher-order moments of the cross-sectional distribution do not exist or that higher-order moments are on an explosive path. In our numerical procedure we integrate over a finite range of capital holdings so this problem cannot occur. To make sure that the numerical procedure doesn't hide diverging properties of the true model it is important to check whether the results are robust to changing the upper bound of capital stock.

## 6 Results and accuracy

In this section, we discuss the accuracy of the aggregate policy function and the parameterized cross-sectional distribution. We also discuss the accuracy of our simulation approach. Tests to check the accuracy of the individual policy function are standard and these are discussed in A.3. Parameter settings of the numerical procedure, such as the order of the polynomial and the number of grid points, are given in A.2.

### 6.1 Parameter values

In the benchmark specification, parameter values are taken from Krusell and Smith (1998) and are reported in Tables 1 and 2. The discount rate, coefficient of relative risk aversion, share of capital in GDP, and the depreciation rate take on standard values. Unemployed people are assumed to earn a fixed fraction of 15% of the wage of the employed.<sup>17</sup> The value of  $\Delta^a$  is equal to 0.01 so that productivity in a boom,  $1+\Delta^a$ , is two percent above the value of productivity in a recession,  $1-\Delta^a$ . Business cycles are symmetric and the expected duration of staying in the same regime is eight quarters. The unemployment rate in a boom,  $u^g$ , is equal to 4% and the unemployment rate in a recession,  $u^b$ , is equal to 10%. The time endowment,  $\bar{l}$ , is chosen to normalize total labor supply in the recession to one. The average unemployment duration is 2.5 quarters conditional on staying in a recession and equal to 1.5 quarters conditional on staying in a boom. These features correspond with the transition probabilities reported in Table 2.

---

<sup>17</sup>This is the only change relative to Krusell and Smith (1998) who set  $\mu = 0$ . This has little effect on the properties of the model but avoids the possibility of agents having zero consumption.

Table 1: Benchmark calibration

Parameters	$\beta$	$\gamma$	$\alpha$	$\delta$	$\bar{l}$	$\mu$	$\Delta^a$
Values	0.99	1	0.36	0.025	1/0.9	0.15	0.01

Table 2: Transition probabilities

$s, \varepsilon / s', \varepsilon'$	$1-\Delta^a, 0$	$1-\Delta^a, 1$	$1+\Delta^a, 0$	$1+\Delta^a, 1$
$1-\Delta^a, 0$	0.525	0.35	0.03125	0.09375
$1-\Delta^a, 1$	0.038889	0.836111	0.002083	0.122917
$1+\Delta^a, 0$	0.09375	0.03125	0.291667	0.583333
$1+\Delta^a, 1$	0.009155	0.115885	0.024306	0.850694

## 6.2 Aggregate policy function

In this section, we address the accuracy of the aggregate policy function. In Section 6.2.1, we establish the accuracy of the functional form taking the parameterization of the cross-sectional distribution as given. In Section 6.2.2, we establish whether more moments are needed as state variables. In Section 6.2.3, we describe a more demanding accuracy test by taking a multi-period perspective.

### 6.2.1 Accuracy of functional form of aggregate policy function

The aggregate policy functions,  $\vec{\Gamma}_n^e(a, a_{-1}, m; \psi_n^{\Gamma^e})$  and  $\vec{\Gamma}_n^u(a, a_{-1}, m; \psi_n^{\Gamma^u})$  capture the law of motion of the end-of-period values of the three moments that are used as state variables. The approximation uses a tensor product polynomial with at most first-order terms for  $a$  and  $a_{-1}$ , since  $a$  can take on only two values, and up to second-order terms for the elements of  $m$ .

Accuracy is evaluated using a grid of the aggregate state variables on which the three variables with continuous support can take on a fine range of values. In particular,  $\{\overleftarrow{m}^{e,1}\} = \{35, 35.2, \dots, 42.4\}$ ,  $\{\overleftarrow{m}^{u,1}\} = \{33.5, 33.7, \dots, 41.5\}$ , and  $\{\overrightarrow{m}_{-1}^{u,c}\} = \{0, 0.05\%, \dots, 0.2\%\}$ . At each grid point, we use the values of  $m$  and the reference moments to obtain the corresponding density exactly as they are calculated in the algorithm. Whether this parameterization of the cross-sectional distribution is accurate will be discussed below. Using the parameterized density and the individual policy function, we calculate  $\overrightarrow{m}^{u,c}$ ,  $\overrightarrow{m}^{e,1}$ , and  $\overrightarrow{m}^{u,1}$ . These explicitly calculated values are compared with those generated by the approximations  $\vec{\Gamma}_n^e(s)$  and  $\vec{\Gamma}_n^u(s)$ .

Table 3 reports for each of the three statistics the average and maximum absolute

% error across this fine set of grid points.<sup>18</sup> The errors for the first-order moment of the capital stock of the employed are small. The maximum error is 0.012% and the average error is 0.0059%. The moments for the unemployed are somewhat bigger, but still relatively small. In particular, the maximum error for the first-order moment is equal to 0.92% and the average error is equal to 0.24%. This maximum is attained when the value of  $\overleftarrow{m}^{e,1}$  takes on the highest and  $\overleftarrow{m}^{u,1}$  the lowest grid value, which is an unlikely if not impossible combination to occur. The average and maximum error for  $\overrightarrow{m}^{u,c}$  are 0.2 and 0.84 percentage points (pp), respectively.<sup>19</sup> There are two reasons why these two numbers are not problematic. First, given that both the actual and the approximation predict very low fractions of agents at the constraint, these errors are of no importance and it wouldn't make sense to spend computing time on improving the part of  $\overrightarrow{\Gamma}^u(s)$  that determines  $\overrightarrow{m}^{u,c}$ . Second, in Section 6.2.3, we show using a simulation that the economy doesn't get close to points in the state space where such large errors are observed. In fact, using a simulation of 1,000 periods we find an average error of 0.0076 percentage points and a maximum error of 0.071 percentage points.

Table 3: Accuracy of functional form of aggregate policy function

Moment	approximation error	
	average	maximum
$\overrightarrow{m}^{e,1}$	5.9e-3%	1.2e-2%
$\overrightarrow{m}^{u,1}$	2.4e-1%	9.2e-1%
$\overrightarrow{m}^{u,c}$	2.0e-1pp	8.4e-1pp

### 6.2.2 Number of moments as state variables

The algorithm uses  $\overrightarrow{m}_{-1}^{u,c}$ ,  $\overleftarrow{m}^{e,1}$ , and  $\overleftarrow{m}^{u,1}$  as state variables and in this section we analyze whether additional moments should be used as state variables. That is, *conditional on staying within the class of cross-sectional distributions pinned down by the reference moments* does it make a difference if additional moments are used as state variables. In particular, we check whether changes in the second-order moment

<sup>18</sup>Since  $\overrightarrow{m}^{u,c}$  is a small number and already a percentage, we express the error for  $\overrightarrow{m}^{u,c}$  in terms of percentage points difference an not as a percentage.

<sup>19</sup>This maximum difference for  $\overrightarrow{m}^{u,c}$  is also attained at the unlikely combination of a very high value for  $\overleftarrow{m}^{e,1}$  and very low value for  $\overleftarrow{m}^{u,1}$ . At this grid point, the value from our approximation is equal to 0.152% and the recalculated value is 0.996%.

matter for the key set of moments the agents predict, i.e.,  $\overrightarrow{m_{-1}^{u,c}}$ ,  $\overleftarrow{m^{e,1}}$ , and  $\overleftarrow{m^{u,1}}$ . To do this we calculate at each of the aggregate grid points  $\overrightarrow{m^{u,c}}$ ,  $\overrightarrow{m^{e,1}}$ , and  $\overrightarrow{m^{u,1}}$  in two different ways. First, when  $\overleftarrow{m^{e,2}}$  and  $\overleftarrow{m^{u,2}}$  take on its reference values, i.e., the average observed in the simulated series (conditional on the value of  $a$ ). Second, when  $\overleftarrow{m^{e,2}}$  and  $\overleftarrow{m^{u,2}}$  take on the maximum values observed in the simulation, but the values of  $\overleftarrow{m^{e,j}}$ , and  $\overleftarrow{m^{u,j}}$  for  $j > 2$  are still equal to the reference moments.

Table 4 reports for each of the three statistics the average and maximum absolute % change across the grid points when the variance increases.

Table 4: Effect of increase in variance under reference distribution

Moment	average change	maximum change
$\overrightarrow{m^{e,1}}$	5.2e-3%	1.1e-1%
$\overrightarrow{m^{u,1}}$	1.9e-1%	6.4e-1%
$\overrightarrow{m^{u,c}}$	1.6e-1pp	5.8e-1pp

The effect of the increase in the variance on  $\overrightarrow{m^{e,1}}$  and  $\overrightarrow{m^{u,1}}$  is small, especially considering that the increase in the variance is enormous. Again, the largest changes occur at unlikely grid points and the changes for  $\overrightarrow{m^{u,c}}$  are larger. Of course, it is not surprising that an increase in the variance has an effect on the fraction of agents choosing a zero capital stock, since the increase in the variance increases the fraction of agents close to zero. Given the lack of importance of agents at the constraint, it doesn't make sense to add the second-order moment as a state variable.

### 6.2.3 Multi-period perspective

A word of caution is warranted in drawing conclusions about accuracy from the type of one-period tests performed in the last two sections. The reason is that small errors can accumulate over time if they do not average out. To investigate this issue we compare values for  $\overrightarrow{m^{e,1}}$ ,  $\overrightarrow{m^{u,1}}$ , and  $\overrightarrow{m^{u,c}}$  generated by two different procedures. First, we generate these statistics with our simulation procedure that explicitly integrates over the choices made by the agents in the economy. This procedure does not use our approximations  $\overrightarrow{\Gamma^e}(a, a_{-1}, m)$  and  $\overrightarrow{\Gamma^u}(a, a_{-1}, m)$ . Second, we generate these statistics using only our approximations  $\overrightarrow{\Gamma^e}(a, a_{-1}, m)$  and  $\overrightarrow{\Gamma^u}(a, a_{-1}, m)$ . It is important to point out that this second procedure only uses  $\overrightarrow{\Gamma^e}(a, a_{-1}, m)$  and  $\overrightarrow{\Gamma^u}(a, a_{-1}, m)$ , and involves nothing more than basic algebra. That is, the output of our aggregate law

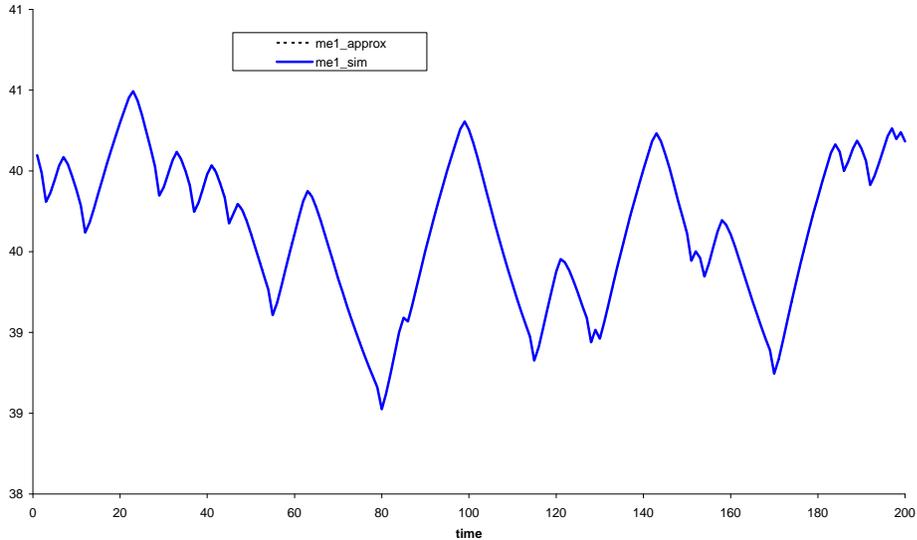


Figure 1:  $\overrightarrow{m^{e,1}}$  generated using the approximation  $\overrightarrow{\Gamma^e}(s)$  or the simulation on a continuum of agents

of motion will be used as the input in the next period.<sup>20</sup> This comparison, thus, is truly a multi-period accuracy test.

The results for  $\overrightarrow{m^{e,1}}$ ,  $\overrightarrow{m^{u,1}}$ , and  $\overrightarrow{m^{u,c}}$  are plotted in Figures 1, 2, and 3 respectively. The graphs make clear that our approximate aggregate laws of motion do a magnificent job of tracking the movements of  $\overrightarrow{m^{e,1}}$  and  $\overrightarrow{m^{u,1}}$ . In fact, one cannot even distinguish the moments generated by  $\overrightarrow{\Gamma^e}(s)$  and  $\overrightarrow{\Gamma^u}(s)$  from the corresponding moments generated by explicit integration over the individual policy rules. Some differences between the two procedures are visible for  $\overrightarrow{m^{u,c}}$ , but our approximate aggregate laws of motion track the changes in  $\overrightarrow{m^{u,c}}$  well. As mentioned above, in a sample of 1,000 observations the average and maximum absolute difference are 0.0076 and 0.071 percentage points, respectively.

### 6.3 The parameterized cross-sectional distribution

Parameterization of the cross-sectional distribution with a flexible functional form serves two objectives in our algorithm. First, it enables the algorithm to calculate

<sup>20</sup>After adjusting, of course, for the employment-status flows.

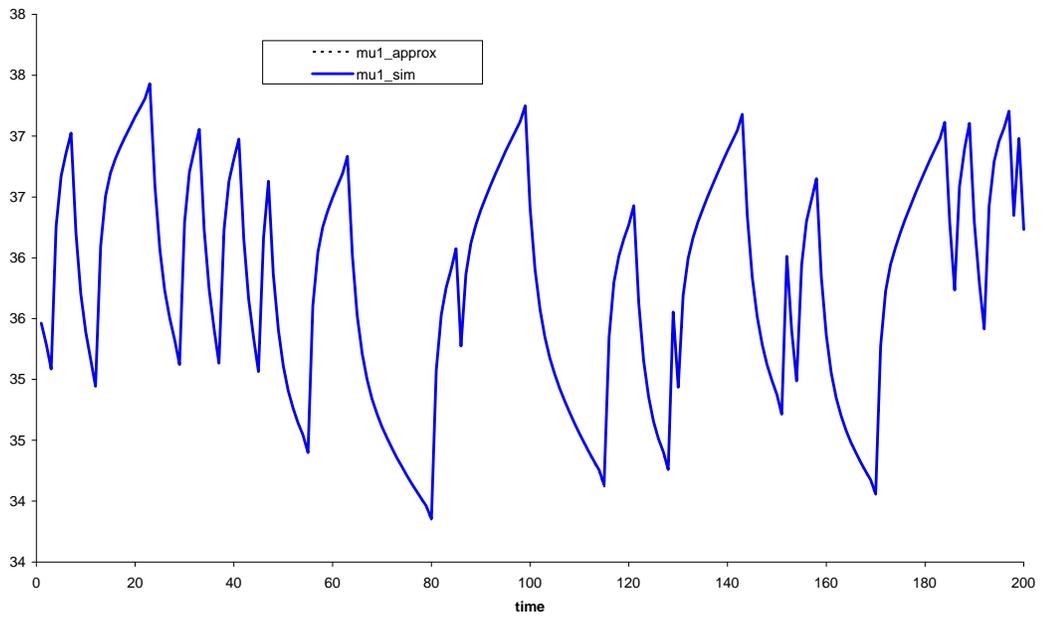


Figure 2:  $\overrightarrow{m^{u,1}}$  generated using the approximation  $\overrightarrow{\Gamma^u}(s)$  or the simulation on a continuum of agents

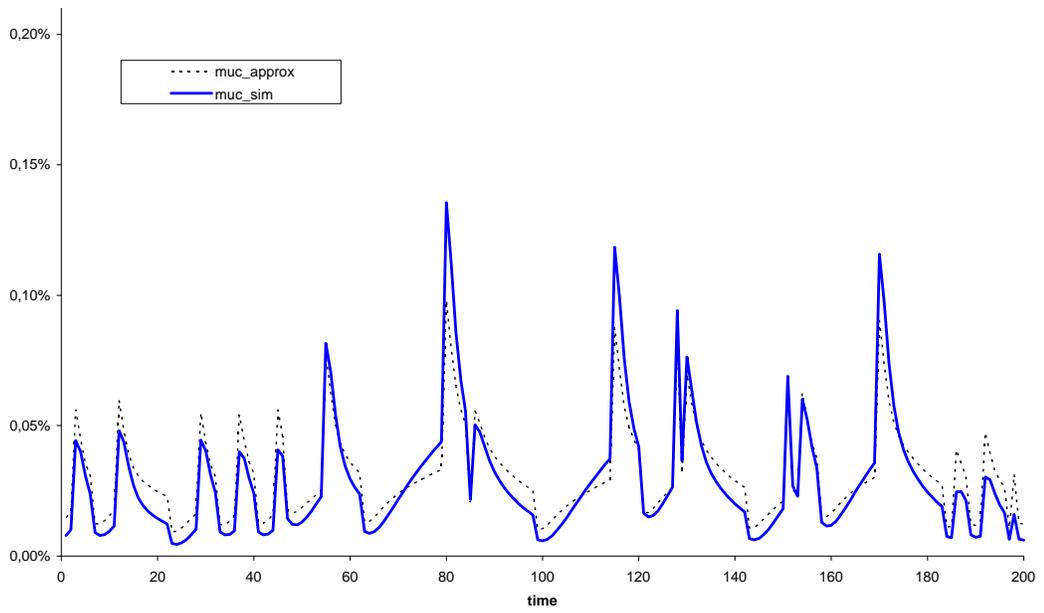


Figure 3:  $\overrightarrow{m}^{u,c}$  generated using the approximation  $\overrightarrow{\Gamma}^{u,c}(s)$  or the simulation on a continuum of agents

the aggregate laws of motion,  $\overrightarrow{\Gamma}^e(s)$  and  $\overrightarrow{\Gamma}^u(s)$ , with standard projection techniques, since with a parameterized density (i) next period's moments can be calculated on a prespecified grid and (ii) next period's moments can be calculated with quadrature instead of the less efficient Monte Carlo techniques. Second, it makes it possible to simulate the economy without cross-sectional variation, which improves the procedure to find the reference moments.

An accurate representation of the cross-sectional distribution may not be necessary for an accurate solution of the model. What is needed for an accurate solution of the model are accurate aggregate laws of motion,  $\overrightarrow{\Gamma}^e(s)$  and  $\overrightarrow{\Gamma}^u(s)$ , since the agent is only interested in predicting future prices and to determine these one does not need the complete distribution, just the set of statistics that determine prices, that is,  $\overrightarrow{m}_{-1}^{u,c}$ ,  $\overleftarrow{m}^{e,1}$ , and  $\overleftarrow{m}^{u,1}$ .

In this section, we take on the more demanding test to check whether the cross-sectional parametrization,  $P(k; \rho^e)$  and  $P(k; \rho^u)$ , are accurate. This is a more demanding test for the following reasons. First, it requires that all  $N_{\overline{M}}$  moments used to pin down the distribution are accurately calculated instead of just the  $N_M$  moments that are used as state variables. More importantly, because the shape of the cross-sectional distribution is endogenous and time-varying, the functional form used must be flexible enough to capture the unknown and changing shapes.

To check the accuracy of our simulation procedure and, thus, the accuracy of our parameterized densities, we do the following. We start in Section 6.3.1 with a comparison between the simulated time path of moments generated by our parameterized densities, with those generated by a standard simulation using  $N_N$  agents. The alternative simulation is, of course, subject to sampling variation, but the advantage of the standard simulation procedure is that there is no functional restriction on the cross-sectional distribution at all.

A second accuracy test consists of checking whether the results settle down if  $N_{\overline{M}}$  increases. This is done in Section 6.3.2. The last accuracy test checks whether moments of order higher than  $N_{\overline{M}}$  are calculated precisely.

### 6.3.1 Comparison between simulation procedures

In this section, we compare the moments generated by our new simulation approach with those generated by the standard Monte Carlo simulation procedure. We report Monte Carlo simulations with 10,000 agents and 100,000 agents.<sup>21</sup> The same

---

<sup>21</sup>In our implementation of the simulation procedure we impose—as in Krusell and Smith (1998)—that  $u_t = u^g(a^b)$  when  $a_t = a^g(a^b)$ . We also ensure that the flows into and out of (un)employment expressed as a fraction of the total population correspond to those one would find with a continuum of agents. Our alternative simulation procedure uses a continuum of agents and

individual policy function is used under the two approaches. We plot time paths of generated moments when we impose the value of  $a$  to alternate deterministically between  $1 - \Delta^a$  and  $1 + \Delta^a$  every 100 periods so that the behavior of the economy during a transition between regimes becomes clear. A set of initial observations is discarded so that effects of the initial distribution are no longer present.<sup>22</sup>

Figure 4 reports the evolution of the end-of-period first moment of the employed  $\overrightarrow{m^{e,1}}$ . Note that the group of the employed is at least 9,000 (90,000) in the Monte Carlo simulations with 10,000 agents (100,000 agents). Although the sampling variation is still visible in the graph with 10,000 agents, it is small relative to the observed changes in the calculated moment. It has virtually disappeared in the simulation with 100,000 agents.

Figure 5 shows the evolution of the end-of-period first-order moment of the unemployed  $\overrightarrow{m^{u,1}}$ . The number of observations in this group is much smaller, amounting to 400 agents (4,000 agents) during recessions in the Monte Carlo simulation with 10,000 (100,000) agents. The sampling variation is substantial in the simulation with 10,000 agents and is still visible in the simulation with 100,000 agents.

The effect of sampling variation is very clear when we consider the fraction of agents at the constraint  $\overrightarrow{m^{u,c}}$ , whose evolution is reported in Figure 6. In this case, even the Monte Carlo simulation with 100,000 agents displays severe sampling variation. The fraction of agents at the constraint is small, and the unemployed do not own much capital. Any inaccuracy due to the high sampling variation is, thus, likely to be inconsequential for the properties of any aggregate series.<sup>23</sup>

A good way to document the higher accuracy of the new simulation procedure is to look at the transition from the bad regime (high unemployment rate) to the good regime (low unemployment rate). The new simulation procedure clearly shows a sharp increase in the fraction of agents at the constraint when the economy enters the good regime. In contrast, for the standard simulation procedure this increase is not always present and there are many other spikes. The true solution of the model should exhibit such a spike. In this economy, employed agents become unemployed every period. When the economy moves from the high-unemployment to the low-unemployment regime, then the flow out of employment into unemployment drops sharply. This means that after the regime change, a smaller fraction of employed

---

automatically imposes the correct stocks and flows.

<sup>22</sup>Details about the simulation procedure are given in Table 9.

<sup>23</sup>In this model with only two discrete possibilities for the employment status one could over-sample the unemployed but this would be problematic to do if the states that occur less frequently have continuous support.

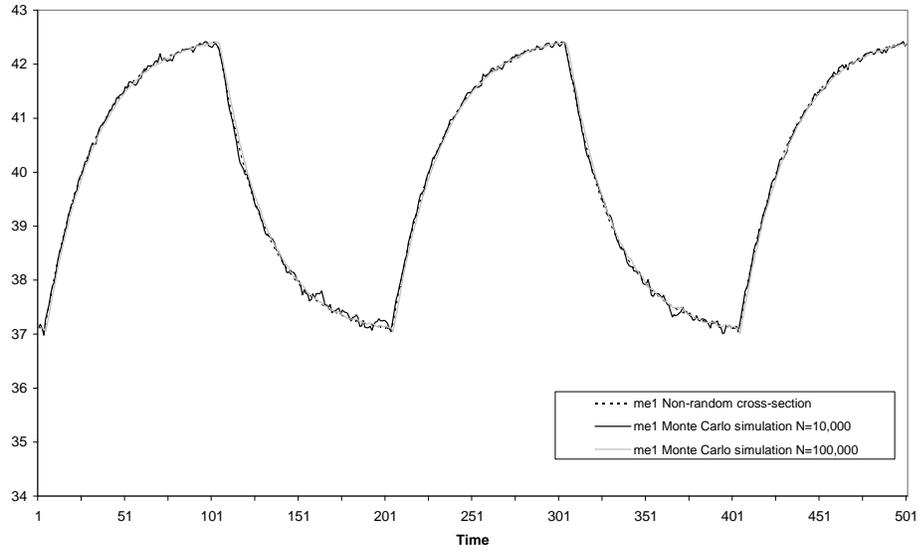


Figure 4:  $\overrightarrow{m^{e,1}}$  generated using a finite and a continuum of agents

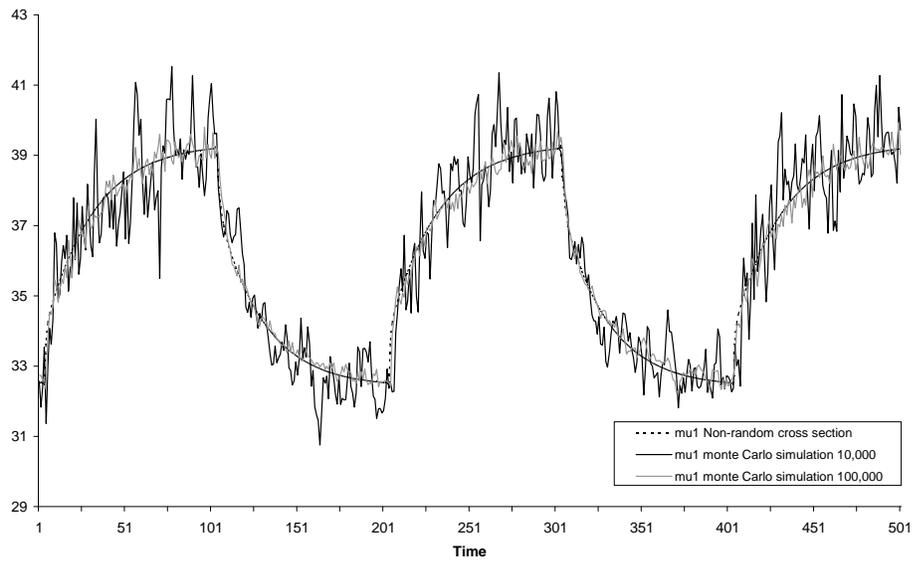


Figure 5:  $\overrightarrow{m^{u,1}}$  generated using a finite and a continuum of agents

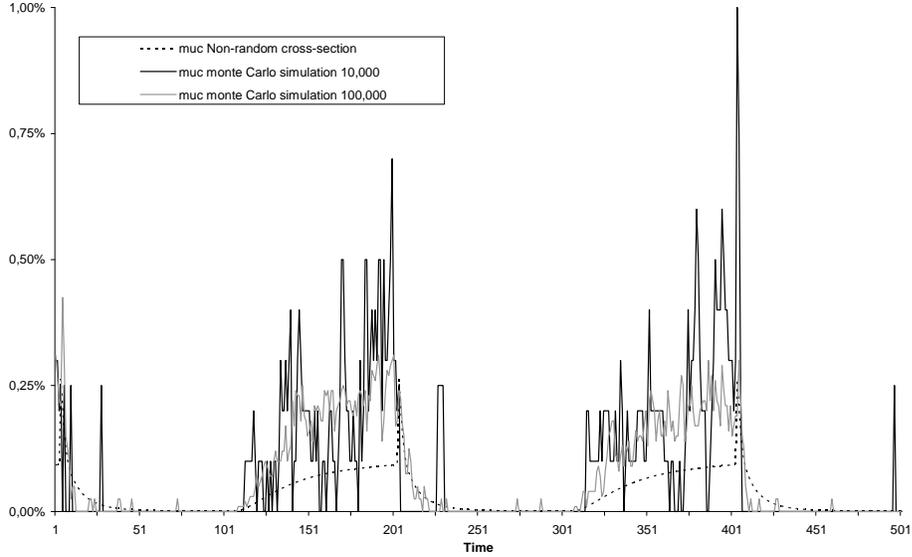


Figure 6:  $\overrightarrow{m^{u,c}}$  generated using a finite and a continuul of agents

becomes unemployed every period, and thus an unemployed agent is much less likely to have been employed the previous period. Consequently, after a change to the low-unemployment regime, a larger fraction of unemployment agents will have a zero capital stock.

To conclude, our new simulation procedure clearly is more accurate than standard Monte Carlo simulation procedures. However, for the current model the simulation procedure is accurate where it matters, namely in tracing the mean capital stock of the employed.

Eventually, not only our proposed simulation method is more accurate than Monte Carlo simulation, but it is also an order of magnitude faster. For instance, Monte Carlo simulations over 3500 periods would take 1 hour and 6 minutes with 10,000 agents and seven hours and 40 minutes with 100,000 agents. In contrast, our simulation procedure takes only 7 minutes for the same number of periods.<sup>24</sup>

### 6.3.2 Increasing $N_{\overline{M}}$

In this paper, we use  $N_{\overline{M}}$  moments to characterize the cross-sectional distribution and  $N_M$  of these are state variables. In Section 6.2.2, we analyzed whether additional

<sup>24</sup>Using a 2.66 GHz Intel Core 2 Duo E6700 processor.

moments are necessary as state variable. Here we address the question whether increasing  $N_{\overline{M}}$  makes a difference.

To check the importance of  $N_{\overline{M}}$ , we simulate an economy using different values of  $N_{\overline{M}}$  to parameterize the cross section. We check when the results settle down. The idea of the test is made clear in Figures ?? and ?? that plot the second and sixth-order moment of the distribution for the unemployed for different values of  $N_{\overline{M}}$ . The figures make clear that  $N_{\overline{M}}$  has to be sufficiently high. When we increase  $N_{\overline{M}}$  from 2 to 4 then the generated moments change enormously. A further increase from 4 to 5 still causes some changes, but when we increase  $N_{\overline{M}}$  from 5 to 6 then the changes are very minor.

Table 5 reports the results for all moments. It corroborates the results from the figures. Using a value of  $N_{\overline{M}}$  equal to 2 is clearly way too low to generate an accurate set of moments. An increase in  $N_{\overline{M}}$  from 5 to 6, however, only causes minor changes in the generated moments. For example, the average change for  $\overleftarrow{m}^{e,2}$  is only 0.25% and the average change for  $\overleftarrow{m}^{u,2}$  is equally small. The average (maximum) change for  $\overleftarrow{m}^{e,6}$  and  $\overleftarrow{m}^{u,6}$  are equal to 2.3% (3.7%) and 2.4% (3.7%), respectively. These are, thus, somewhat, higher, but as made clear in the graph, these errors are low relative to the observed variation in the moments.

Table 5: The effect of increasing  $N_{\overline{M}}$

Moment	% change relative to $N_{\overline{M}} = 6$					
	average			maximum		
	$N_{\overline{M}} = 2$	$N_{\overline{M}} = 4$	$N_{\overline{M}} = 5$	$N_{\overline{M}} = 2$	$N_{\overline{M}} = 4$	$N_{\overline{M}} = 5$
$\overleftarrow{m}^{e,2}$	32.9%	2.5%	0.25%	35.9%	2.8%	0.54%
$\overleftarrow{m}^{e,3}$	90.2%	12.5%	1.5%	95.1%	14.5%	2.4%
$\overleftarrow{m}^{e,4}$	60.2%	6.5%	1.4%	63.6%	7.9%	2.2%
$\overleftarrow{m}^{e,5}$	90.6%	9.7%	2.4%	95.0%	14/-%	3.5%
$\overleftarrow{m}^{e,6}$	78.0%	7.1%	2.3%	80.9%	12.1%	3.7%
$\overleftarrow{m}^{u,2}$	32.2%	2.4%	0.25%	36.0%	2.8%	0.55%
$\overleftarrow{m}^{u,3}$	91.9%	12.6%	1.5%	95.5%	14.6%	2.4%
$\overleftarrow{m}^{u,4}$	60.3%	6.6%	1.4%	63.6%	7.9%	2.2%
$\overleftarrow{m}^{u,5}$	92.0%	9.2%	2.4%	95.4%	13.1%	3.5%
$\overleftarrow{m}^{u,6}$	78.0%	6.7%	2.4%	80.8%	11.2%	3.7%

### 6.3.3 The shape of the distribution

There are two aspects to our approximation of the cross-sectional distribution. First, the class of functions used and second the value of  $N_{\overline{M}}$ . The coefficients are chosen so that the first  $N_{\overline{M}}$  moments are correct. But higher-order moments are implied by the values of the first  $N_{\overline{M}}$  moments *and* the class of functions chosen. For example, when one uses a normal distribution, then one can impose any mean and variance, but skewness and kurtosis are no free parameters.

So for any finite value of  $N_{\overline{M}}$ , the class of approximating polynomials used, thus, imposes certain restrictions on the function form. Here we check those restrictions along a simulated time path by comparing the  $j^{\text{th}}$ -order moments for  $j > N_{\overline{M}}$  implied by the parameterized cross-section with those calculated by integration of the individual's policy function.

In particular, we do the following. Draw a long time series for the aggregate productivity shock,  $a$ . Let  $\rho_1^e$  and  $\rho_1^u$  be the parameters of the cross-sectional distribution in the first period and let  $m_1^{e,c}$  and  $m_1^{u,c}$  be the fraction of employed and unemployed agents with zero capital holdings at the beginning of the period. With this information, we calculate the end-of-period values of the first  $N_{\overline{M}}$  moments.<sup>25</sup>

$$\overrightarrow{m_1^{w,1}} = \int_0^\infty k(\varepsilon^w, \mathbf{k}, s) P(\mathbf{k}; \rho_1^w) dk, \quad w \in \{e, u\}. \quad (16)$$

$$\overrightarrow{m_1^{w,j}} = \int_0^\infty \left[ k(\varepsilon^w, \mathbf{k}, s) - \overrightarrow{m_1^{w,1}} \right]^j P(\mathbf{k}; \rho_1^w) dk, \quad 1 < j \leq N_{\overline{M}}, \quad w \in \{e, u\}. \quad (17)$$

In exactly the same way, we calculate higher-order moments. That is,

$$\overrightarrow{m_1^{w,j}} = \int_0^\infty \left[ k(\varepsilon^w, \mathbf{k}, s) - \overrightarrow{m_1^{w,1}} \right]^j P(\mathbf{k}; \rho_1^w) dk, \quad j > N_{\overline{M}}, \quad w \in \{e, u\}. \quad (18)$$

Now we check whether these higher-order moments ( $j > N_{\overline{M}}$ ) are similar to the moments implied by the parameterized cross-sectional distribution. To do this, we use the first  $N_{\overline{M}}$  end-of-period moments to calculate the coefficients of the corresponding

---

<sup>25</sup>By indicating in bold the variable that we are integrating over, we make clear that we are integrating over the argument of the policy function not the outcome of the policy function.

approximating density,  $\overrightarrow{\rho}_1^e$  and  $\overrightarrow{\rho}_1^u$ . Next, we calculate higher-order moments implied by this parameterized cross-sectional density. That is,

$$\overrightarrow{\mathfrak{S}}_1^j = \int_0^\infty \left[ \mathbf{k} - \overrightarrow{m}_1^{w,1} \right]^j P(\mathbf{k}; \overrightarrow{\rho}_1^w) d\mathbf{k}, \quad j > N_{\overline{M}}, \quad w \in \{e, u\}. \quad (19)$$

If the shape of the cross-sectional distribution is not too restrictive, then the implied higher-order moments correspond to the explicitly calculated higher-order moments.

We perform this exercise using values for  $N_{\overline{M}}$  equal to 2 and 6 and then calculate the average and maximum error observed along the simulation of 2,000 observations with the error term defined as follows:<sup>26</sup>

$$\frac{\left| \overleftarrow{\mathfrak{S}}^{w,n} - \overleftarrow{m}^{w,n} \right|}{\overleftarrow{m}^{w,n}}.$$

Table 6: Implied and actual higher-order moments;  $N_{\overline{M}} = 2$

$N_{\overline{M}} = 2$	Employed		Unemployed	
Error (%)	Average	Max	Average	Max
$\frac{\left  \overleftarrow{\mathfrak{S}}^{w,3} - \overleftarrow{m}^{w,3} \right }{\overleftarrow{m}^{w,3}}$	5.6%	9.1%	7.0%	15.3%
$\frac{\left  \overleftarrow{\mathfrak{S}}^{w,4} - \overleftarrow{m}^{w,4} \right }{\overleftarrow{m}^{w,4}}$	2.8E-1%	3.2E-1%	2.1E-1%	6.4E-1%
$\frac{\left  \overleftarrow{\mathfrak{S}}^{w,5} - \overleftarrow{m}^{w,5} \right }{\overleftarrow{m}^{w,5}}$	4.7%	7.5%	5.6%	11.7%
$\frac{\left  \overleftarrow{\mathfrak{S}}^{w,6} - \overleftarrow{m}^{w,6} \right }{\overleftarrow{m}^{w,6}}$	5.7E-1%	7.5E-1%	9.3E-1%	2.6%
$\frac{\left  \overleftarrow{\mathfrak{S}}^{w,7} - \overleftarrow{m}^{w,7} \right }{\overleftarrow{m}^{w,7}}$	3.7%	6.1%	5.3%	10.3%
$\frac{\left  \overleftarrow{\mathfrak{S}}^{w,8} - \overleftarrow{m}^{w,8} \right }{\overleftarrow{m}^{w,8}}$	6.5E-1%	1.2%	2.3%	5.7%
$\frac{\left  \overleftarrow{\mathfrak{S}}^{w,9} - \overleftarrow{m}^{w,9} \right }{\overleftarrow{m}^{w,9}}$	2.9%	5.0%	5.6%	10.3%
$\frac{\left  \overleftarrow{\mathfrak{S}}^{w,10} - \overleftarrow{m}^{w,10} \right }{\overleftarrow{m}^{w,10}}$	5.6E-1%	1.7%	4.0%	9.2%

<sup>26</sup>Note that the accuracy measure is actually defined for beginning-of-period moments, but this is simply a transformation of end-of-period values taking into consideration the change in the employment status.

Table 7: Implied and actual higher-order moments;  $N_{\overline{M}} = 6$

$N_{\overline{M}} = 6$	Employed		Unemployed	
Error (%)	Average	Max	Average	Max
$\overleftarrow{\mathfrak{S}}^{w,7} - \overleftarrow{m}^{w,7}$	2.8E-2%	7.3E-1%	1.0E-1%	2.2E-1%
$\overleftarrow{\mathfrak{S}}^{w,8} - \overleftarrow{m}^{w,8}$	4.3E-2%	1.0E-1%	1.8E-1%	4.3E-1%
$\overleftarrow{\mathfrak{S}}^{w,9} - \overleftarrow{m}^{w,9}$	9.3E-2%	2.3E-1%	3.8E-1%	8.8E-1%
$\overleftarrow{\mathfrak{S}}^{w,10} - \overleftarrow{m}^{w,10}$	1.3E-1%	3.1E-1%	5.6E-1%	1.3%

Tables 6 and 7 report the errors for  $N_{\overline{M}} = 2$  and  $N_{\overline{M}} = 6$ , respectively. When  $N_{\overline{M}}$  is equal to 2, then observed error terms are large for the odd-numbered moments. That is, the shape of the distribution implied by our class of approximating functions doesn't capture the correct shape of the distribution when a second-order approximation is used. The results are much better when  $N_{\overline{M}}$  is equal to 6. Now we observe much smaller errors. The largest errors are for the 10-th order moment of the capital stock of the unemployed. For this moment, the maximum error is 1.3%, which is not excessively high.

## 7 Concluding comments

In this paper, we have developed a new algorithm to solve models with heterogeneous agents and aggregate uncertainty. We used the algorithm to solve the model in Krusell and Smith (1998) and found our numerical solution to have similar properties to the one obtained with simulation procedures. The ability to obtain similar numerical outcomes with quite different algorithms builds confidence in the results. Accuracy tests, of course, can do the same but they also have its limits, especially in high-dimensional models.<sup>27</sup> The models the profession will consider in the future are likely to become more complex. Since different algorithms have different strengths and weaknesses it is important to have a variety of algorithms to choose from. This paper helps to create a richer portfolio of algorithms to solve these complex models especially because the building blocks used are so different from the popular alternative.

<sup>27</sup>In particular, small errors that accumulate over time to non-trivial magnitudes may be hard to detect.

# A Appendix

## A.1 Details on Transition equations

This appendix describes how the change in employment status that occurs at the beginning of each period affects the moments of the cross-sectional distribution. Although, we use centralized moments as state variables, we actually do not use centralized moments here. It is easier to first do the transformation for non-centralized moments and then calculate the centralized moments.

**From beginning to end-of-period.** Let  $g_{a,w}$  be the mass of agents with employment status  $w$  when the economy is in regime  $a$ . At the beginning of the period we have to following groups of agents:

1. Unemployed with  $k = 0$ , whose mass is equal to  $\overleftarrow{m}^{u,c} g_{a,u}$
2. Unemployed with  $k > 0$ , whose mass is equal to  $(1 - \overleftarrow{m}^{u,c}) g_{a,u}$
3. Employed with  $k = 0$ , whose mass is equal to  $\overleftarrow{m}^{e,c} g_{a,e}$
4. Employed with  $k > 0$ , whose mass is equal to  $(1 - \overleftarrow{m}^{e,c}) g_{a,e}$

Agents in group #1 choose  $k' = 0$ , while agents in group #2 either set  $k' = 0$  or  $k' > 0$ . Let the fraction of agents that set  $k' = 0$  be equal to  $\zeta_{u,k>0}^{k'=0}$ . Thus, the fraction of unemployed agents that set  $k' = 0$ , is equal to

$$\overrightarrow{m}^{u,c} = \overleftarrow{m}^{u,c} + \zeta_{u,k>0}^{k'=0} (1 - \overleftarrow{m}^{u,c}).$$

The  $i^{\text{th}}$  moment of the capital stock chosen by agents in group #2 is equal to

$$\mu_{u,k>0}^{k'\geq 0,i} = \zeta_{u,k>0}^{k'=0} \times 0^i + (1 - \zeta_{u,k>0}^{k'=0}) \times \overrightarrow{m}^{u,i}.$$

Thus,

$$\overrightarrow{m}^{u,i} = \frac{\mu_{u,k>0}^{k'\geq 0,i}}{(1 - \zeta_{u,k>0}^{k'=0})}$$

where  $\mu_{u,k>0}^{k'\geq 0,i} = \int_0^{+\infty} k(0, k, s)^i P(k, \rho^u) dk$ .

$\overrightarrow{m}^{e,c} = 0$ , since employed agents never choose a zero capital stock. To calculate  $\overrightarrow{m}^{e,i}$ , we need the (mean of the) capital stock chosen by those in group #3, i.e.,

$$\mu_{e,k=0}^{k'>0,i} = k(1, 0, s)^i,$$

and the mean of the capital stock chosen by those in group 4, i.e.,

$$\mu_{e,k>0}^{k'>0,i} = \int_0^{+\infty} k(e, k, s)^i P(k, \rho^e) dk.$$

Weighting the two means gives

$$\overrightarrow{m}^{e,i} = \frac{\overleftarrow{m}^{e,c} \mu_{e,k=0}^{k'>0,i} + (1 - \overleftarrow{m}^{e,c}) \mu_{e,k>0}^{k'>0,i}}{\overleftarrow{m}^{e,c} + (1 - \overleftarrow{m}^{e,c})}$$

**From end-of-period to beginning of next period.** At the end of the current period, we have to following groups of agents:

1. Unemployed with  $k' = 0$ , whose mass is equal to  $\overrightarrow{m}^{u,c} g_{a,u}$
2. Unemployed with  $k' > 0$ , whose mass is equal to  $(1 - \overrightarrow{m}^{u,c}) g_{a,u}$
3. Employed with  $k' > 0$ , whose mass is equal to  $g_{a,e}$

$g_{ww',aa'}$  stands the mass of agents with employment status  $w$  that have employment status  $w'$  in the next period, conditional on the values of  $a$  and  $a'$ . For all combinations of  $a$  and  $a'$  we have,

$$g_{uu,aa'} + g_{eu,aa'} + g_{ue,aa'} + g_{ee,aa'} = 1.$$

The number of agents with  $k' = 0$  as a fraction of all agents that are unemployed in the following period is equal to

$$\overleftarrow{m}^{u,c'} = \frac{g_{uu,aa'}}{g_{uu,aa'} + g_{eu,aa'}} \overrightarrow{m}^{u,c}.$$

and the fraction of all employed agents at the constraint is equal to

$$\overleftarrow{m}^{e,c'} = \frac{g_{ue,aa'}}{g_{ue,aa'} + g_{ee,aa'}} \overrightarrow{m}^{u,c}$$

Next period's  $i^{\text{th}}$ -order moments of the distributions with *strictly positive* capital holdings are equal to

$$\overleftarrow{m}^{u,i'} = \frac{g_{uu,aa'}(1 - \overrightarrow{m}^{u,c})\overrightarrow{m}^{u,i} + g_{eu,aa'}\overrightarrow{m}^{e,i}}{g_{uu,aa'}(1 - \overrightarrow{m}^{u,c}) + g_{eu,aa'}}$$

and

$$\overleftarrow{m}^{e,i'} = \frac{g_{ue,aa'}(1 - \overrightarrow{m}^{u,c})\overrightarrow{m}^{u,i} + g_{ee,aa'}\overrightarrow{m}^{e,i}}{g_{ue,aa'}(1 - \overrightarrow{m}^{u,c}) + g_{ee,aa'}}$$

## A.2 Parameters of numerical procedure

Tables 8 and 9 report the parameter settings used in implementing the numerical procedure.

Table 8: Construction of the grid

State variables	$k, \varepsilon, a_{-1}, a, \overrightarrow{m}^{u,c}, \overleftarrow{m}^{e,1}, \overleftarrow{m}^{u,1}$
Number of grid points	$M_k = 50, M_{\overrightarrow{m}^{u,c}} = 5, M_{\overleftarrow{m}^{e,1}} = 5, M_{\overleftarrow{m}^{u,1}} = 5$ $M_\varepsilon = M_a = M_{a_{-1}} = 2$ $k : [0, 99]$
range of values	$M_{\overrightarrow{m}^{u,c}} : [0, 0.002]$ $M_{\overleftarrow{m}^{e,1}} : [35, 42.4]$ $M_{\overleftarrow{m}^{u,1}} : [33.5, 41.5]$
location of points	Chebyshev nodes

Table 9: Simulation procedure

Number of agents	Non-random cross-section: $\infty$ Monte-Carlo: 10,000 or 100,000 agents
Number of periods	3,500 (first 1,000 discarded)
Initial distribution	Stationary distribution in the bad state

### A.3 Accuracy of individual policy function

In this section, we discuss the accuracy of the numerical solution for the individual policy function, taking as given the transition laws for the cross-sectional distribution. We check the accuracy of our policy function using a grid of capital holdings that is much more dense than the one used to obtain the numerical solution and the grid points for the aggregate state variables. In particular, we use  $k=\{0, 0.1, 0.2, \dots, 99\}$ . At each grid point, we first calculate the consumption level implied by the numerical approximation, i.e.,  $c^{app}$ , and next period's capital.

Next, we calculate the conditional expectation, not by using the approximation, but by explicit (numerical) integration of  $\beta c'^{-\gamma}(r' + 1 - \delta)$  over the possible realizations of  $a'$  and  $\varepsilon'$ . The numerical approximation for the policy function is used to evaluate next period's consumption values for the different realizations of the idiosyncratic and aggregate random variable. Let  $c$  be the current-period value of consumption implied by the explicitly calculated conditional expectation on that grid point. Accuracy is measured using  $|c - c^{app}|/c^{app}$ . We calculate the average and the maximum across these approximation errors. Table 10 reports the average across aggregate grid points and the maximum. Even the maximum errors are small for most grid points. The average of the maximum errors for the unemployed is equal to 0.67%. For the employed errors are substantially smaller.

Table 10: Euler equation error

Unemployed		Employed	
average	maximum	average	maximum
0.17%	5.9%	0.15%	1.1%

The maximum percentage error for the unemployed is equal to 5.9%. Such a maximum error is high, but it occurs when consumption takes on a value that is low even for an unemployed agent, namely 0.690 according to our approximation and 0.651 according to the explicitly calculated conditional expectation.<sup>28</sup>

## References

- [1] Christiano, Lawrence J. and Fisher, Jonas D., 2000, Algorithms for solving dynamic models with occasionally binding constraints, *Journal of Economic Dynamics and Control* 24, p. 1179-1232.

<sup>28</sup>Average consumption for the unemployed is 2.77.

- [2] Den Haan, Wouter J. and Albert Marcet, 1990, Solving the stochastic growth model by parameterizing expectations, *Journal of Business and Economic Statistics* 14, p. 31-34.
- [3] Den Haan, Wouter J., 1996, Heterogeneity, aggregate uncertainty and the short-term interest rate, *Journal of Business and Economic Statistics* 14, p. 399-411.
- [4] Den Haan, Wouter J., 1997, Solving dynamic models with aggregate shocks and heterogeneous agents, *Macroeconomic Dynamics* 1, p. 355-386.
- [5] Judd, Kenneth L., 1992, Projection methods for solving aggregate growth models, *Journal of Economic Theory* 58, p. 410-452.
- [6] Judd, Kenneth L., 1998, *Numerical methods in economics*, MIT Press, Cambridge, MA.
- [7] Krueger, Dirk and Felix Kubler, 2004, Computing equilibrium in OLG models with stochastic production, *Journal of Economic Dynamics and Control* 28, p. 61-76.
- [8] Krusell, Per and Smith, Anthony, 1998, Income and wealth heterogeneity in the macroeconomy, *Journal of Political Economy*, p. 868-896.
- [9] Miao, Jianjun, 2006, Competitive equilibria of economies with a continuum of consumers and aggregate shocks, *Journal of Economic Theory* 128, p. 274-298.
- [10] Preston, Bruce, and Roca, Mauro, 2006, Incomplete markets, heterogeneity and macroeconomic dynamics, manuscript Columbia University.
- [11] Reiter, Michael, 2002, Recursive computation of heterogeneous agent models, manuscript, Universitat Pompeu Fabra.
- [12] Reiter, Michael, 2006, Solving heterogeneous agent models by projection and perturbation, manuscript, Universitat Pompeu Fabra.
- [13] Rios-Rull, J. Victor, 1997, Computation of equilibria in heterogeneous agent models, Federal Reserve Bank of Minneapolis Staff Report 231.
- [14] Veracierto, Marcello, 2002, Plant-level irreversible investment and equilibrium business cycles, *American Economic Review* 92, p. 181-197
- [15] Young, Eric, 2005, Approximate aggregation, manuscript, University of Virginia.