Solving the Incomplete Markets Model With Aggregate Uncertainty by Backward Induction*

Michael Reiter
Institute for Advanced Studies, Vienna

this version: September 17, 2008

Abstract

The paper describes a method to solve models with a continuum of agents, incomplete markets and aggregate uncertainty. I use backward induction on a finite grid of points in the aggregate state space. The aggregate state includes a small number of statistics (moments) of the cross-sectional distribution of capital. For any given set of moments, agents use a specific cross-sectional distribution, called "proxy distribution", to compute the equilibrium. Information from the steady state distribution as well as from simulations can be used to chose a suitable proxy distribution.

*I am grateful to Wouter den Haan, Ken Judd, an anonymous referee and seminar participants at the 2003 CFS-JEDC conference, at SITE 2004, and at the Paris School of Economics in 2007 for helpful comments on earlier drafts of this paper. Support from Spanish ministry grants SEJ2004-03619 and SEC2002-01601 is gratefully acknowledged.

1 Introduction

The paper describes a method to solve the model with incomplete markets and aggregate uncertainty as specified in Den Haan, Judd, and Juillard (2008). The method is applicable to a large class of models with a continuum of agents and incomplete markets.

Some well known solution methods (Krusell and Smith (1998), Den Haan (1997)) iterate between solving the household problem, conditional on an assumed aggregate law of motion, and updating the aggregate law of motion, based on the policy function of the household. Iteration continues until consistency is achieved between the solution of the individual household problem and the aggregate law of motion. The idea of the method I present here is to solve the method in one run of backward iterations, on a discrete grid of points in the aggregate state space. Consistency between individual and aggregate solution is enforced in each step of the backward iteration. This can be done separately point for point of the grid of aggregate states.

Following Krusell and Smith, I work with the approximation that household decisions are based not on the complete state of the model, which is infinite-dimensional because it contains the whole cross-sectional distribution, but rather on an n_m-vector of statistics (or "moments") m, which characterize this distribution. This gives rise to a major complication: in order to check whether at the point in the state space characterized by m, individual behavior is consistent with the aggregate law of motion, it is not sufficient to know the statistics m, but we need to know the whole cross-sectional distribution. We therefore need a selection mechanism that tells us which distribution to use when statistics are equal to m. I call this a "Distribution Selection Function" (DSF), a mapping from aggregate states to cross-sectional distributions:

$$DSF: \qquad (m, a) \rightarrow F(k, \varepsilon; m, a) \tag{1}$$

Here, a is the exogenous state variable (aggregate productivity). The DSF provides a family of distributions, parameterized by the statistics m. I call the distribution $F(k, \varepsilon; m, a)$ selected by (1) the "proxy distribution"; it stands in for the true cross-sectional distribution, which the household does not know or is not supposed to utilize.

The advantages of this method are:

- It is fast, because it does not need the repeated simulation and solution of the model. This becomes more important if more than 1 or 2 statistics are used to describe the distribution.
- It depends very little (in simple cases not at all) on simulations of the model, thereby avoiding problems of sampling error that come with simulations.
- It does not rely on a parameterization of the aggregate transition law. At any point m in the state space we solve for next period's statistics m' (conditional on the aggregate shock).

The disadvantage is that the method relies on a DSF. In Krusell and Smith (1998), the distributions are generated in the course of the solution.

2 The Method

The outline of the algorithm is the following:

1. Choose a discrete representation of the economy. This includes
   - a finite parameterization of the cross-sectional distribution function (not to be confounded with the vector of statistics m);
   - a finite parameterization of the household value function (alternatively one could approximate the consumption function).
2. Solve for the stationary state of the economy without aggregate shocks.
3. Choose a DSF which uses the information from the solution without aggregate shocks.
4. Choose a set of statistics m which describe the cross-sectional distribution and which are taken as arguments of the household value function in the model with aggregate shocks.
5. Solve the model with aggregate shocks by backward induction. Use the stationary distribution as a tool for the DSF.
6. Optional: simulate the model, and use the distributions arising in the information to obtain a more accurate DSF.
7. Optional: iterate between 5. and 6. to further improve the DSF, until no increase in accuracy is found.

2.1 Discretization

2.1.1 Representing distributions

In many applications it may be appropriate to use a smooth representation of the cross-sectional density (for example, an exponential of polynomials as in Algan, Allais, and Den Haan (2008)). Here I use a non-smooth approximation, namely a histogram. The reason is that the theoretical model implies a stationary distribution with a large number of discrete point masses, because the idiosyncratic shock has only two realizations. Then the households at the liquidity constraint will enter the range of positive capital as a discrete point mass, if they receive a positive employment shock. In the calibration that we investigate below, the fraction of households at the liquidity constraint is so small that the distribution appears to be almost continuous, but other calibrations imply a distribution with many spikes. My parameterization accounts for this possibility.

A discrete approximation D of a distribution is then given by

1. a grid $\kappa_D^{\varepsilon,j}$ for $j = 0, \ldots, n_d$ and each employment status $\varepsilon \in \{u, e\}$

2. a set of probabilities $p_D^{\varepsilon,j}$, $j = 1, \ldots, n_d$, $\varepsilon \in \{u, e\}$ with $\sum_{j=1}^{n_d} p_D^{u,j} = u$ and $\sum_{j=1}^{n_d} p_D^{e,j} = 1 - u$, which denote the mass of households with employment status ε and capital $k \in \left(\kappa_D^{\varepsilon,j-1}, \kappa_D^{\varepsilon,j}\right)$.[1]

Underlying this discretization is the assumption that, for each ε, the cross-sectional density is constant within any interval $k \in \left(\kappa_D^{\varepsilon,j-1}, \kappa_D^{\varepsilon,j}\right)$. The expectation of any function $g(k, \varepsilon)$ over this distribution is therefore given by

$$\mathrm{E}_D\left[g(k,\varepsilon)\right] \equiv \sum_{\varepsilon\in\{e,u\}} \sum_{j=1}^{n_d} \frac{p_D^{\varepsilon,j}}{\kappa_D^{\varepsilon,j} - \kappa_D^{\varepsilon,j-1}} \int_{\kappa_D^{\varepsilon,j-1}}^{\kappa_D^{\varepsilon,j}} g(k,\varepsilon)\, dk \tag{2}$$

The point mass of households at the borrowing constraint is handled by a very small interval at zero: $\kappa_D^{\varepsilon,0} = 0$ and $\kappa_D^{\varepsilon,1} = 10^{-6}$. Choosing this interval not too small is helpful in making the mapping from the time-t-distribution to the $(t+1)$-distribution smooth. The fraction of households with zero assets is therefore represented by $p_D^{\varepsilon,1}$.

2.1.2 Approximating the value function

In the model with aggregate uncertainty, the household is assumed to base its decisions on the following set of state variables: the individual employment state ε and assets k, the aggregate productivity a and the vector of statistics m. They are the arguments of the household value function.

In each step of the backward iteration, the value function $V(k, \varepsilon; m, a)$ and its derivative $V_k(k, \varepsilon; m, a)$ are computed at a grid of the aggregate and individual state variables. Denote by $\mathcal{K} = \{\kappa_1, \ldots, \kappa_{n_i}\}$ the set of grid points of individual capital, and by $\mathcal{M} = \{\mathfrak{m}_1, \ldots, \mathfrak{m}_{n_a}\}$ the set of grid points of aggregate statistics (moments). With two moments, for example, each $\mathfrak{m}_\nu$ is a vector with 2 elements. I use the following interpolation scheme in the aggregate variables m off the grid $\mathcal{M}$.

$$V(\kappa_j, \varepsilon, m, a) = \sum_{\nu=1}^{n_a} \omega_\nu(m) V(\kappa_j, \varepsilon, \mathfrak{m}_\nu, a), \tag{3a}$$

$$V_k(\kappa_j, \varepsilon, m, a) = \sum_{\nu=1}^{n_a} \omega_\nu(m) V_k(\kappa_j, \varepsilon, \mathfrak{m}_\nu, a), \qquad j = 1, \ldots, n_i, \qquad \varepsilon \in \{u, e\} \tag{3b}$$

where the weights $\omega_\nu(m)$ satisfy $\sum_{\nu=1}^{n_a} \omega_\nu(m) = 1$ for any m. In a multilinear interpolation, the number of nonzero $\omega_\nu(m)$ is $2^{dim(m)}$, and we have $\omega_\nu(m) \geq 0$. I also use a piecewise cubic interpolation, which usually gives more accurate approximations, but has $4^{dim(m)}$ nonzero weights and implies some negative weights.

Notice that the value function is concave in k. Using an interpolation with $\omega_\nu(m) \geq 0$ guarantees that the interpolated values $V(k, \varepsilon, m, a)$, $V_k(k, \varepsilon, m, a)$ are compatible with a concave function. Choosing then a shape-preserving interpolation in capital (cf. Section 2.5) such as Schumaker's quadratic splines (Judd 1998, p.231), the interpolant will

[1] Notice that the $p_t^{\varepsilon,j}$ in Den Haan (2008) are normalized differently, adding up to unity for each ε.

be concave. It turns out to be safer to start with this interpolation, and use the resulting value function as starting point for a solution with more general interpolation weights. In the latter case, I check in each step whether the interpolant is concave.

2.2 The steady state without aggregate shocks

For the steady state calculation, we can assume that the distribution at the beginning and the end of the period are defined on the same grid, for both employment states: $\kappa_D^{\varepsilon,j} = \kappa_j$. A distribution D is then characterized by the probabilities only, which are stacked into the column vector

$$\vec{p}(D) = \begin{pmatrix} p_D^{u,1} & \dots & p_D^{u,n_d} & p_D^{e,1} & \dots & p_D^{e,n_d} \end{pmatrix}^T \tag{4}$$

Solving for the stationary state implies the following steps.

1. Guess the aggregate end-of-period capital stock K. Since effective labor input is given exogenously, knowing K we get the gross interest rate R and the wage rate w. Knowing factor prices, we can compute the household decision rules $k'(k, \varepsilon)$ by dynamic programming.

2. Compute the transition matrix $\tilde{T}^\varepsilon$ from beginning-of-period capital to end-of-period capital. The element $\tilde{T}^\varepsilon_{l,j}$ can be interpreted as describing the probability of being in interval (κ_{l-1}, κ_l) at the end of the period, conditional on being in the interval (κ_{j-1}, κ_j) at the beginning of the period. In the notation of Den Haan (2008), this is given by

$$\tilde{T}^\varepsilon_{l,j} = \frac{\partial(F^{\varepsilon,\kappa_l} - F^{\varepsilon,\kappa_{l-1}})}{\partial p^{\varepsilon,j}} \tag{5}$$

   The transition from end-of-this-period's distribution of capital D to next period's D' is then given by

$$\vec{p}(D') = T\vec{p}(D) \tag{6}$$

$$T \equiv \begin{bmatrix} \tilde{T}^u & 0 \\ 0 & \tilde{T}^e \end{bmatrix} \left(\begin{pmatrix} \pi(u,u) & \pi(e,u) \\ \pi(u,e) & \pi(e,e) \end{pmatrix} \otimes I_{n_d} \right) \tag{7}$$

   Here, $\pi(u, e)$ is the individual probability of moving from unemployment to employment, etc. I_{n_d} is the n_d-identity matrix, and $\otimes$ denotes Kronecker product.

3. Find the vector of probabilities $\vec{p}(D^*)$ that satisfies $\vec{p}(D^*) = T\vec{p}(D^*)$. In our model this is unique; we can find it as the eigenvector of T belonging to the eigenvalue of 1. Since T is sparse, this problem can be solved very quickly (for example with the Matlab command 'eigs'). The vector $\vec{p}(D^*)$ gives the steady state distribution D^*.

4. See whether the guess of K was correct by checking the condition

$$K = \mathrm{E}_{D^*}[k] \tag{8}$$

Use a one-dimensional nonlinear solver to find the K such that (8) is met. I use Brent's method (cf. Press, Flannery, Teukolsky, and Vetterling (1986, Section 9.3)) as implemented in the Matlab command 'fzero'.

2.3 The Distribution Selection Function

The process of picking a proxy distribution $\mathcal{D}(m, a)$ goes in two steps. First, we determine a "reference distribution" $\mathcal{R}(m, a)$, which is an informed guess of what the distribution with (m, a) should approximately look like, but which may not yet satisfy the moments m precisely. In the next step, we choose the proxy distribution that is closest to the reference distribution under the constraint that it has moments m.

2.3.1 Reference distribution from stationary distribution

Since we have to solve the model before we can simulate it, we first have to compute a solution that does not use any information from simulation. For that we can use the ergodic distribution that arises in the steady state without aggregate shocks. Denote this by D^*. Then the reference distribution is given by

$$p^{u,j}_{\mathcal{R}(m,a)} = p^{u,j}_{D^*} \frac{u(a)}{u^*} \qquad j = 1, \ldots, n_d \tag{9a}$$

$$p^{e,j}_{\mathcal{R}(m,a)} = p^{e,j}_{D^*} \frac{1 - u(a)}{1 - u^*} \qquad j = 1, \ldots, n_d \tag{9b}$$

$$\kappa^{\varepsilon,j}_{\mathcal{R}(m,a)} = \phi \kappa^{\varepsilon,j}_{D^*}, \qquad j = 1, \ldots, n_d, \quad \varepsilon \in \{e, u\} \tag{9c}$$

Distributions are scaled by the respective unemployment rate $u(a)$ relative to steady state unemployment u^*. The constant ϕ rescales the grid points used to approximate the distribution and is chosen such that $\mathrm{E}_{\mathcal{R}(m,a)}[k] = m_1$ (the first moment m_1 is aggregate capital, cf. Section 2.4).

2.3.2 Reference distribution from simulated distributions

If we have already simulated the model with an earlier solution, we can study the distributions that arise in the simulation and obtain a reference distribution that is closer to the typically realized distributions. We assume that all the distributions in the simulation were defined on the same grid of κ_j; then the reference distribution shares the same grid, and we base the reference distribution on averages of distributions D_t obtained in a simulation:

$$p^{\varepsilon,j}_{\mathcal{R}(m,a)} = \frac{1}{T} \sum_{t=1}^{T} p^{\varepsilon,j}_{D_t}, \qquad j = 1, \ldots, n_d, \quad \varepsilon \in \{e, u\} \tag{10a}$$

$$\kappa^{\varepsilon,j}_{\mathcal{R}(m,a)} = \phi \kappa^{\varepsilon,j}_{D^*}, \qquad j = 1, \ldots, n_d, \quad \varepsilon \in \{e, u\} \tag{10b}$$

where again the constant ϕ is chosen such that $\mathrm{E}_{\mathcal{R}(m,a)}[k] = m_1$. In the experiments below, the distributions D_t where obtained from a simulation over 7000 periods, where the 5000 first periods were dropped, so that $T = 2000$.

It should be stressed that the only use of simulation that is made in solving the model is in finding the D_t used in (10).

2.3.3 From the reference distribution to the proxy distribution

The DSF picks the distribution that is closest to the reference distribution (in a mean-square sense) and exactly satisfies the moment conditions. Its probabilities $p^{j,\varepsilon}_{\mathcal{D}(m,a)}$ for $j = 1, \ldots, n_d$ and $\varepsilon \in \{e, u\}$ are the solution to the constrained least squares problem

$$\min_{p^{e,1},\ldots,p^{e,n_d},p^{u,1},\ldots,p^{u,n_d}} \sum_{\varepsilon\in\{e,u\}} \sum_{j=1}^{n_d} \left[p^{\varepsilon,j} - p^{\varepsilon,j}_{\mathcal{R}(m,a)}\right]^2 \tag{11a}$$

subject to

$$\sum_{j=1}^{n_d} p^{u,j} = u(a), \qquad \sum_{j=1}^{n_d} p^{e,j} = 1 - u(a) \tag{11b}$$

$$\sum_{\varepsilon\in\{e,u\}} \sum_{j=1}^{n_d} \frac{p^{\varepsilon,j}}{\kappa^{\varepsilon,j}_{\mathcal{R}(m,a)} - \kappa^{\varepsilon,j-1}_{\mathcal{R}(m,a)}} \int_{\kappa^{\varepsilon,j-1}_{\mathcal{R}(m,a)}}^{\kappa^{\varepsilon,j}_{\mathcal{R}(m,a)}} M_l(k,\varepsilon)\, dk = m_l, \qquad l = 1, \ldots, n_m \tag{11c}$$

In other words, we select the function that minimizes the mean squared difference to the reference function, under the condition that the moments of the proxy function are right. It would be straightforward to replace the criterion in (11a) by a weighted mean square, but (11a) is what was used in the examples below. Reiter (2002) gives a simple algorithm that finds an approximate solution to (11) very fast.

Notice that in the case where only the mean is used as a state variable, the proxy distribution is equal to the reference distribution, since the constant ϕ is chosen such that the moment condition is satisfied.

2.4 Moments

We assume that the household bases its decision at time t not on the whole distribution, or its finite representation D_t, but only on a set of n_m statistics of the form

$$m_{l,t} = \mathrm{E}_{D_t}[M_l(k,\varepsilon)] \qquad l = 1, \ldots, n_m \tag{12}$$

where the $M_l(k,\varepsilon)$ are known integrable functions. Obviously, one can also use nonlinear one-to-one transformations of the m_l in (12) as state variables.

For the first moment I always use mean capital $m_{1,t} = K_t$, therefore $M_1(k,\varepsilon) = k$. Together with exogenous labor supply, this implies that the factor prices R_t and w_t are a function of m_{t-1} and a_t.

For the solutions with two moments, I use $M_2(k,\varepsilon) = C^*(k,\varepsilon)$, which means that the second "moment" is the aggregate amount of consumption obtained by applying the steady state consumption function to the current distribution. The idea is that this statistic is the best indicator for aggregate household consumption, and should therefore be helpful in predicting future aggregate capital. Other statistics such as the unconditional variance of capital have turned out not to be very successful in this respect.

2.5 Solving the model with aggregate shocks by backward induction

I solve the model by backward iteration on the value function. Using the Euler equation and iterating on the consumption function would be an alternative; it leads to very similar results, but I found the value function approach to give somewhat smaller Euler residuals for low levels of assets.

Knowing the household value function in the stationary state, $V^*(k,\varepsilon)$, we can use it to initialize the value function in the model with aggregate shocks, setting $V'(k,\varepsilon;m,a) = V^*(k,\varepsilon)$ for all m and a.

Now assume we have computed an estimate of the household value function $V'(k',\varepsilon';m',a')$. We interpret this as the value function of the "next period"; the prime denotes next-period variables. We apply a backward iteration step to obtain a new estimate, the value function of "this period". Notice that here all variables and the value function are taken at the end of the period.

Substep 1: Compute the equilibrium. For any point $(\mathfrak{m}_\nu, a)$ in the grid of aggregate states $\mathcal{M}$, and for any realization of next period's aggregate shock a', we have to find the equilibrium values of m'. This is done in the following way:

1. Make a guess m'.

2. From Equs. (3) we obtain interpolated values $V(\kappa_j,\varepsilon',m',a')$ and $V_k(\kappa_j,\varepsilon',m',a')$ for all $\kappa_j \in \mathcal{K}$. Then we use the endogenous grid point method of Carroll (2006) to compute the value function as a function of cash-on-hand, after the employment shock has realized and income has accrued. This is done as follows. From the first order condition $u_c(c(x,\varepsilon')) = {V'}_k(x-c(x,\varepsilon'),\varepsilon',m',a')$ we know that any point (κ_j,ε') is reached by choosing the consumption level $c^{j,\varepsilon'} \equiv {u_c}^{-1}\left({V'}_k(\kappa_j,\varepsilon',m',a'_j)\right)$. This implies that we reach κ_j from the cash-on-hand level $x_{j,\varepsilon'} \equiv \kappa_j + c^{j,\varepsilon'}$. This gives us the household value $\tilde{V}$ as a function of cash-on-hand:

$$\tilde{V}(x_{j,\varepsilon'},\varepsilon',m',a') = u\left[{u_c}^{-1}\left({V'}_k(\kappa_j,\varepsilon',m',a'_j)\right)\right] + V'(\kappa_j,\varepsilon',m',a'_j) \tag{13a}$$

$$\tilde{V}_x(x_{j,\varepsilon'},\varepsilon',m',a') = {V'}_k(\kappa_j,\varepsilon',m',a'_j) \tag{13b}$$

   At levels of x off the endogenous grid $(x_1,\ldots,x_{n_i})$, the values $\tilde{V}(x,\varepsilon',m',a')$ and $\tilde{V}_x(x,\varepsilon',m',a')$ are computed using interpolation with Schumaker splines.[2] Then

[2]The first grid point $x_{1,u}$ gives the level of cash-on-hand at which the liquidity constraint starts binding

consumption as a function of cash-on-hand is given by

$$c(x, \varepsilon', m', a') = {u_c}^{-1}\left(\tilde{V}_x(x, \varepsilon', m', a')\right) \tag{14}$$

3. Start from the proxy distribution $\mathcal{D}(m, a)$ and use the consumption function (14) to obtain the distribution of capital at the end of the next period, D', applying the procedure explained in Den Haan (2008).

4. Iterate on the guess m' until it is consistent with the statistics of the implied distribution:

$$m'_l(a') = \mathrm{E}_{D'}\left[M_l(k, \varepsilon)\right], \qquad l = 1, \ldots, n_m \tag{15}$$

This is a n_m-dimensional fixed-point problem, which I solve by a quasi-Newton method.

Substep 2: Update the value function. Having found the equilibrium values of $m'(\mathfrak{m}_\nu, a, a')$ for each a', we update the value function for each $\kappa_j \in \mathcal{K}$ by

$$V(\kappa_j, \varepsilon; \mathfrak{m}_\nu, a) = \beta \sum_{\varepsilon'=1}^{2} \sum_{a'=1}^{2} \pi(a, \varepsilon; a', \varepsilon') \tilde{V}\left(\tilde{x}_{j,\varepsilon'}, \varepsilon', m'(\mathfrak{m}_\nu, a, a'), a'\right) \tag{16a}$$

$$V_k(\kappa_j, \varepsilon; \mathfrak{m}_\nu, a) = \beta \sum_{\varepsilon'=1}^{2} \sum_{a'=1}^{2} \pi(a, \varepsilon; a', \varepsilon') R(m, a') \tilde{V}_x\left(\tilde{x}_{j,\varepsilon'}, \varepsilon', m'(\mathfrak{m}_\nu, a, a'), a'\right) \tag{16b}$$

where $\tilde{x}_{j,\varepsilon'}$ is short for

$$\tilde{x}_{j,\varepsilon'} \equiv R(\mathfrak{m}_\nu, a')\kappa_j + w(\mathfrak{m}_\nu, a')\left[(1-\tau)\bar{l} I_{\{\varepsilon'=e\}} + \mu I_{\{\varepsilon'=u\}}\right]$$

Notice that the $\tilde{x}_{j,\varepsilon'}$ are obtained by going forward from κ_j, while the points $x_{j,\varepsilon'}$ on the endogenous grid in Substep 1 are obtained by going backward from κ_j. The value function at $\tilde{x}_{j,\varepsilon'}$ is again computed by Schumaker interpolation. Factor prices and income in period $t+1$ are determined by the period-t moments m, which include the capital stock, and period-$(t+1)$ level of aggregate productivity, which determines effective labor input.

The backward iteration step is repeated until convergence is achieved. The criterion I use is that the maximum relative change in V_k is less than 10^{-6} in an iteration.

The backward iteration steps are critical for the computational effort. The following tricks are helpful to speed things up:

---

for the unemployed (it never binds for the employed). To compute $\tilde{V}$ for values of $x < x_{1,u}$, notice that $c = x$ and then we get

$$\tilde{V}(x, u, m, a) = \tilde{V}(x_{1,u}, u, m, a) + u(x) - u(x_{1,u})$$
$$\tilde{V}_x(x, u, m, a) = u_c(x)$$

1. Most importantly, the use of the endogenous grid points method explained above.

2. Acceleration steps. We need not solve for the equilibrium in every iteration. In most steps, we can just keep the $m'(\mathfrak{m}_\nu, a, a')$ of the last iteration and only update the value function (Substep 2). In the computations reported below, I was using 7 acceleration steps for each full step with equilibrium finding, after 20 initial full steps.

3. The quasi-Newton method is started by computing the Jacobian. We need not do this in each iteration, but reuse the Jacobian computed in the last iteration at the same point $(\mathfrak{m}_\nu, a, a')$. The Jacobian is only recomputed if a Newton step fails.

2.6 Simulating the model

Without using "Aggregate Law of Motion" (No-ALM)

At time t, the economy is described by the productivity a_t and the (discretized) cross-sectional distribution D_t. Conditional on the realization of a_{t+1}, we solve for next period's endogenous state m_{t+1}. This is done exactly as in Substep 1 of Section 2.5, with the exception that we use the distribution D_t obtained in the simulation rather than the proxy distribution $\mathcal{D}(m, a)$ in point 3. For V' we take of course the value function obtained after the backward iterations have converged. With the consumption function obtained from (14), we then compute D_{t+1} as explained in point 3 of Substep 1.

Using "Aggregate Law of Motion" (With-ALM)

For solving the model, we do not rely on a parameterization of the aggregate law of motion (ALM). Consistency between individual and aggregate behavior is checked using the DSF. In this sense, the DSF takes the role that the ALM has, for example, in the method of Krusell and Smith (1998). The analogue to simulating the aggregate variables with the ALM is therefore to simulate the model as above, but replacing in every period the actual distribution D_t by the proxy distribution with the same statistics, $\mathcal{D}\left(.,.;\mathrm{E}_{D_t}\left[M(k,\varepsilon)\right], a_t\right)$.

Computing Euler residuals

In the model without aggregate shocks, Euler residuals are computed as in Den Haan (2008). Since that paper does not present Euler residuals for the case of aggregate uncertainty, I report them in Table 1. They are computed as follows. In the course of the long simulation, I compute residuals at periods $t = 1$, $t = 101$, $t = 201$ etc. Assume that during the simulation we have reached the point (D, a). In the course of solving for D, we have already obtained $\tilde{V}(x_{j,\varepsilon}, \varepsilon, D, a)$ for the endogenous grid points $x_{j,\varepsilon}$, as explained in Substep 1 of Section 2.5. We obtain the consumption function from (14), and thereby get next period's cash-on-hand $x'(x, a')$ for any a'. Then we simulate the model one period forward (No-ALM) as explained above, separately for each possible value a', and thereby obtain $D'(a')$. In the process we have computed $\tilde{V}(x_{j,\varepsilon'}, \varepsilon, D'(a'), a')$ for all possible realizations

of a' and ε'. From the $\tilde{V}(.)$ we obtain consumption functions $c(.,.;D'(a'),a')$ as in (14). The consumption choice implied by the conditional expectation is then given by

$$\bar{c}(x,\varepsilon;D,a) \equiv \min\left\{ {u_c}^{-1}\left[\beta\sum_{\varepsilon'=1}^{2}\sum_{a'=1}^{2}\pi(a,\varepsilon;a',\varepsilon')R(D,a')u_c\left(c(x'(x,a'),\varepsilon';D'(a'),a')\right)\right], x\right\}$$

and the Euler residual is reported as $(c-\bar{c})/\bar{c}$. The residuals are computed at all the midpoints of the cash-on-hand grid. Of the absolute Euler residuals, I report the maximum (over k and ε, averaged over t), and the mean (over the first 1000 points of x, as required by Den Haan (2008), and over ε and t).

3 Numerical tests

3.1 Model parameters

For the approximation of the value function, I use $n_i = 500$ grid points in capital, mostly located close to the bound where the constraint starts binding, to get very small Euler residuals in steady state. For the aggregate grid, I use $n_a = 8$ points for each moment, that makes 64 grid points in the two-moment case. The distribution is represented as a histogram of $n_d = 1500$ points. The first 1000 points go from 0 to 100 in steps of 0.1, the last 500 points to from 100 to 392.6 (ten times the steady state capital of the complete markets model) in steps of equal size.

3.2 Results

Table 1 presents accuracy checks for 4 different solutions of the model. "Mom1" and "Mom2" refer to solutions using 1 or 2 statistics, respectively. "Dstst" and "Dsimul" indicate that the reference distribution was obtained from the steady state or from simulated distributions, respectively. "Long Sim." refers to the simulation with 10000 periods, using the values of the aggregate shock given by Den Haan (2008). "Short Sim. 1" is the simulation where the aggregate shock equals $a = a^b$ for 100 periods and $a = a^g$ for 100 periods. "Short Sim. 2" is a simulation where the aggregate shock equals $a = a^g$ for 200 periods and $a = a^b$ for 200 periods. Two accuracy measures are given, namely absolute Euler residuals and the absolute percentage difference in the capital stock between the With-ALM and the No-ALM simulation. Both the max and the mean over the simulations are provided. It turns out that Euler residuals and With-ALM/No-ALM differences are highly correlated, both across simulations and across different variants of the solution. This suggests that those measures are indeed good indicators of the quality of the solution.

To interpret the size of the Euler residuals, note that the mean and max residual in the steady state are 3.07e-4 and 1.33e-6, respectively. The increase in the mean residual when there are aggregate shocks mainly reflects the mistake that households make in predicting the future interest rate. In terms of With-ALM/No-ALM differences, the

maximum difference is around 0.3 percent, the mean absolute difference around 0.1 percent. With respect to both accuracy measures, the two-moment solution is consistently the better one, although the differences in accuracy between solutions are not big. Learning from the simulation seems difficult; having the reference distribution based on simulated distributions does not improve accuracy by much. I have also tried some more complicated ways to learn from the simulations, but the results were not consistently better.

Finally, the method is reasonably fast. Programs are written in Matlab. On a PC with a Intel Core 2 Duo S775 E6750 2×2.67 GHz processor, solving for the steady state takes about a minute, the same is true for the 1-moment solution. The 2-moment solution takes about 10 minutes. Reducing the grid points in the value function approximation (from now $n_i = 500$ to 100, for example) would correspondingly reduce computing time; the loss of accuracy affects in particular the maximum Euler residual.

	Long Sim.		Short Sim. 1		Short Sim. 2	
	max	mean	max	mean	max	mean
Euler Residuals, average over simulations						
Mom1 Dstst	5.54e-5	2.28e-5	6.09e-5	5.21e-5	5.86e-5	4.98e-5
Mom1 Dsimul	5.51e-5	2.09e-5	6.69e-5	5.95e-5	6.05e-5	5.08e-5
Mom2 Dstst	5.25e-5	1.65e-5	3.80e-5	2.28e-5	3.96e-5	3.20e-5
Mom2 Dsimul	5.24e-5	1.50e-5	3.98e-5	2.59e-5	3.86e-5	3.06e-5
Relative Difference K, With-ALM vs. No-ALM						
Mom1 Dstst	3.63e-3	1.31e-3	3.25e-3	2.02e-3	4.05e-3	2.00e-3
Mom1 Dsimul	3.44e-3	1.11e-3	3.53e-3	2.06e-3	3.73e-3	1.81e-3
Mom2 Dstst	2.79e-3	1.05e-3	2.19e-3	7.06e-4	3.82e-3	1.56e-3
Mom2 Dsimul	2.42e-3	8.51e-4	1.73e-3	6.99e-4	3.42e-3	1.34e-3

Euler residuals are reported as in Den Haan (2008).
Relative differences are reported as $\left|K^{No-ALM}/K^{With-ALM}-1\right|$.

Table 1: Accuracy checks

References

Algan, Y., O. Allais, and W. J. Den Haan (2008). Solving heterogeneous-agent models with parameterized cross-sectional distributions. *Journal of Economic Dynamics and Control 32*(3), 875–908.

Carroll, C. D. (2006). The method of endogenous gridpoints for solving dynamic stochastic optimization problems. *Economics Letters 91*(3), 312–320.

Den Haan, W. J. (1997). Solving dynamic models with aggregate shocks and heterogenous agents. *Macroeconomic Dynamics 1*, 355–86.

Den Haan, W. J. (2008). Comparison of solutions to the incomplete markets model with aggregate uncertainty. *Journal of Economic Dynamics and Control*, this issue.

Den Haan, W. J., K. Judd, and M. Juillard (2008). Computational suite of models with heterogeneous agnets: Model specifications. *Journal of Economic Dynamics and Control*, this issue.

Judd, K. L. (1998). *Numerical Methods in Economics.* Cambridge and London: MIT Press.

Krusell, P. and A. A. Smith (1998). Income and wealth heterogeneity in the macroeconomy. *Journal of Political Economy 106*(5), 867–96.

Press, W., B. Flannery, S. Teukolsky, and W. Vetterling (1986). *Numerical Recipes.* Cambridge University Press.

Reiter, M. (2002). Recursive computation of heterogeneous agent models. Manuscript.