Journal of Economic Dynamics & Control 33 (2009) 649–665

Contents lists available at ScienceDirect

Journal of Economic Dynamics & Control

journal homepage: www.elsevier.com/locate/jedc

# Solving heterogeneous-agent models by projection and perturbation

Michael Reiter *

*Department of Economics and Finance, Institute for Advanced Studies (IHS), Stumpergasse 56, Room 305, A-1060 Vienna, Austria*

A R T I C L E I N F O

*Article history:*
Received 29 September 2006
Accepted 21 August 2008
Available online 18 September 2008

*JEL classification:*
C63
C68
E21

*Keywords:*
Heterogeneous agents
Projection methods
Perturbation methods
Invariant distribution

A B S T R A C T

The paper proposes a numerical solution method for general equilibrium models with a continuum of heterogeneous agents that combines elements of projection and of perturbation methods. The basic idea is to solve first for the stationary solution of the model, without aggregate shocks but with fully specified idiosyncratic shocks. Afterwards one computes a first-order perturbation of the solution in the aggregate shocks. This approach allows to include a high-dimensional representation of the cross-sectional distribution in the state vector. The method is applied to a model of household saving with uninsurable income risk and liquidity constraints. Techniques are discussed to reduce the dimension of the state space such that higher order perturbations are feasible.



## 1. Introduction

Stochastic general equilibrium models with incomplete markets and a continuum of heterogeneous agents pose some hard computational problems. The main problem is that the state vector includes the whole cross-sectional distribution of wealth (or whatever the individual state variable of the agent is). This is an infinite-dimensional object.

The most widely used method to solve this class of models is the one pioneered in Krusell and Smith (1998). Alternative methods that have been around for some years are Den Haan (1997) and Reiter (2002). All these approaches have in common that they try to represent the cross-sectional distribution of wealth by a very small number of statistics in order to reduce the dimensionality of the state space. For the models studied in those papers, this approach was shown to work very well. In fact, Krusell and Smith (1998) find that in order to predict future values of aggregate capital with reasonable precision, everything one has to know about today's distribution of capital is the mean. Higher moments of the distribution matter very little. In Krusell and Smith (2006) they explain this approximate aggregation result in detail.

However, this can hardly serve as a general approach to the solution of heterogeneous-agent models. For example, in models where firms follow (S,s)-policies (price setting, inventory holdings, etc.), we expect that the cross-sectional distribution of the relevant variables enters the solution in an essential way, and the solution algorithm may have to keep track of a medium- or high-dimensional representation of this distribution. Some very recent approaches promise to live up to this challenge. Algan et al. (2008) develop a method that uses a flexible parameterization of the cross-sectional distribution and makes heavy use of projection methods. It increases computational efficiency over earlier approaches in a number of ways and will allow to handle a higher number of state variables. Kim et al. (2007) and Preston and Roca (2007) apply perturbation methods. They perturb the solution around the deterministic steady state, with no aggregate or

* Tel.: +43 1 59991 154.
*E-mail address:* michael.reiter@ihs.ac.at

0165-1889/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jedc.2008.08.010

idiosyncratic shocks and a degenerate cross-sectional distribution. The disadvantage of this approach is that it is a local approximation method, and the handling of inequality constraints, for example liquidity constraints, is problematic. The main advantage is that the method will allow to handle higher moments and/or several individual state variables.

In this paper, I propose an alternative procedure, which combines features of projection methods and of perturbation methods.[1] The idea is to compute a solution that is fully nonlinear in the idiosyncratic shocks, but only linear in the aggregate shocks. In a first step, I use ideas of projection methods to find a finite parameterization of the model, including the individual choice functions and the cross-sectional distribution of wealth. We put all the parameters characterizing the economy at time $t$ into the vector $\Theta_t$. We then compute the stationary solution $\Theta^*$ of the finite model without aggregate shocks. In a second step, I use the idea of perturbation to compute an approximation of the dynamics of $\Theta_t$ around the steady state $\Theta^*$. This approximation is linear in the aggregate shocks and in $\Theta_t$ itself. With this approach we can include a very detailed representation of the cross-sectional distribution in the state vector. In the numerical examples of Section 5, the distribution is characterized by up to 1000 state variables.

I apply the method to a model of household saving with uninsurable income risk and liquidity constraints. To make the exposition transparent, I keep the model as simple as possible,[2] but the method has much wider applicability. It can be used for a wide range of models where agents have one individual state variable (wealth, productivity, product price, etc.). Generalizations to more than one continuous individual state variable are straightforward, although this requires to approximate a multidimensional cross-sectional distribution, such that accuracy will probably be reduced. The approach should prove particularly useful in models where the distribution matters, for example when individual policy functions are discontinuous (menu cost models of pricing, or labor market participation, Haefke and Reiter (2006)). Of course, the method is suitable only if linearity in aggregate variables is good enough for the purpose at hand, as it is in many business cycle applications. It is not sufficient, for example, for problems of asset pricing. But even then, the linear solution is a useful first step. With it we can study in detail the dynamics of the cross-sectional distribution that results from small shocks, and we can find out whether some form of state aggregation is possible, and what suitable state variables are. In a further step, one can then solve for the nonlinear dynamics in a reduced state space.

The plan of the paper is as follows. Section 2 presents a simple model that will serve as a test case for the solution method. The method is presented in detail in Section 3. Section 4 explains how parameters and functional forms were chosen. Numerical results are given in Section 5. Section 6 discusses whether a low-dimensional VAR in the moments of the capital distribution is suitable to describe the dynamics of the model. Section 7 concludes.

## 2. The model

There is a continuum of infinitely lived households of unit mass. Households are ex ante identical, and differ ex post through the realization of their individual labor productivity. They supply their labor inelastically. Production takes place in competitive firms with constant-returns-to-scale technology. A government is introduced into this model to the sole purpose of creating some random redistribution of wealth. This helps to identify the effect of the wealth distribution on the dynamics of aggregate capital.

### 2.1. Production

Output is produced by perfectly competitive firms, using the Cobb–Douglas gross production function

$$Y_t = \mathscr{Y}(\bar{K}_{t-1}, L_t, Z_t) = Z_t A \bar{K}_{t-1}^{\alpha} L_t^{1-\alpha}, \quad 0<\alpha<1 \tag{1}$$

where $A$ is a constant. Notice that we follow the convention to choose the time index of a variable as the period in which this variable is determined. Production at the beginning of period $t$ uses $\bar{K}_{t-1}$, the aggregate capital stock determined at the end of period $t-1$. Since labor supply is exogenous, and individual productivity shocks cancel due to the law of large numbers, aggregate labor input is constant and normalized to $L_t = 1$. Capital accumulates as

$$\bar{K}_t = (1-\delta)\bar{K}_{t-1} + Y_t - C_t \tag{2}$$

where $\delta$ is the depreciation rate and $C_t$ is aggregate consumption. The aggregate productivity parameter $Z_t$ follows the AR(1) process

$$\log Z_{t+1} = \rho_z \log Z_t + \varepsilon_{Z,t+1} \tag{3}$$

[1] For a different type of model, a similar idea has been used in Campbell (1998, computational appendix). I am grateful to Tony Smith for pointing this out to me. It solves a macroeconomic model with a continuum of plants, stated as a social planner's problem. To my knowledge, this is the first paper which exploits the idea of reducing an infinite-dimensional problem to a finite-dimensional one by some approximation and then linearize the approximate model around the steady state. The reason why the computational contribution of that paper has not drawn more attention (apart from a very interesting application in Michelacci and Lopez Salido (2007)) may be that people were not aware of its wide applicability, or that they were afraid of the lengthy linearizations that appear in the computational appendices of Campbell (1998) and Michelacci and Lopez Salido (2007). In fact, as will become clear below, the linearization can be left to the computer, so that it puts no burden on the researcher.

[2] The Matlab programs to solve this model are available at http://elaine.ihs.ac.at/~mreiter/hetagent.tar.gz.

where $\varepsilon_Z$ is an i.i.d. shock with expectation 0 and standard deviation $\sigma_Z$. The before tax interest rate $\bar{r}_t$ and the wage $w_t$ are determined competitively:

$$\bar{r}_t = \mathscr{Y}_K(\bar{K}_{t-1}, L_t, Z_t) - \delta \tag{4}$$

$$w_t = \mathscr{Y}_L(\bar{K}_{t-1}, L_t, Z_t) \tag{5}$$

### 2.2. The government

As I said earlier, the only purpose of the government is to create some random redistribution of wealth. The government does this through a tax $\tau_t$ levied at the beginning of $t$ on the capital stock accumulated at the end of period $t-1$. The after tax interest rate $r_t$ is then related to the before tax rate by

$$r_t = \bar{r}_t - \tau_t \tag{6}$$

The revenues are redistributed lump sum, providing a transfer $T_t$ to every household. The government has to balance its budget every period:

$$T_t = \tau_t \bar{K}_{t-1} \tag{7}$$

The tax rate follows an AR(1) process around its steady state value $\tau^*$:

$$\tau_{t+1} - \tau^* = \rho_\tau(\tau_t - \tau^*) + \varepsilon_{\tau,t+1} \tag{8}$$

where $\varepsilon_\tau$ is an i.i.d. shock with expectation 0 and standard deviation $\sigma_\tau$.

### 2.3. The household

There is a continuum of households, indexed by $i$. Households differ ex post by their labor productivity $\xi_{t,i}$. Labor productivity is assumed to be i.i.d., and is normalized to have unit mean:

$$\mathrm{E}\xi_{t,i} = \mathrm{E}_{t-1}\xi_{t,i} = 1 \tag{9}$$

For convenience of exposition I assume that $\xi_{t,i}$ has a continuous distribution with density function $f(\xi)$. The corresponding distribution function is denoted by $F(\xi)$. The density and distribution function of labor income depend on the wage and are denoted by

$$f_t^y(y) = \frac{1}{w_t} f\left(\frac{y}{w_t}\right) \tag{10a}$$

$$F_t^y(y) = F\left(\frac{y}{w_t}\right) \tag{10b}$$

The household supplies inelastically one unit of labor. Its labor earnings are therefore given by

$$\bar{y}_{t,i} = w_t \xi_{t,i} \tag{11}$$

With the government lump sum transfer $T_t$, the net non-capital income is given by

$$y_{t,i} = \bar{y}_{t,i} + T_t \tag{12}$$

Household $i$ enters period $t$ with asset holdings $k_{t-1,i}$ left at the end of the last period. It receives the net interest rate $r_t$ on its assets, such that the available resources after income of period $t$ ('cash on hand') are given by

$$x_{t,i} = (1 + r_t)k_{t-1,i} + y_{t,i} \tag{13}$$

Cash on hand is split between consumption and asset holdings:

$$k_{t,i} = x_{t,i} - c_{t,i} \tag{14}$$

We impose the liquidity constraint

$$k_{t,i} \geqslant \underline{k} \tag{15}$$

The solution of the household problem is given by a consumption function $C(x;\Omega_t)$, or equivalently a savings function

$$K(x;\Omega_t) \equiv x - C(x;\Omega_t) \tag{16}$$

where $\Omega_t$ denotes the vector of state variables, which will be specified later. The first order equation of the household problem is the Euler equation

$$U'(C(x;\Omega_t)) \geqslant \beta \mathrm{E}_t[R(\Omega_{t+1})U'(C(R(\Omega_{t+1})(x - c_{t,i}) + y_{t+1,i};\Omega_{t+1}))] \quad \text{and} \quad C(x;\Omega_t) = x - \underline{k} \tag{17a}$$

or

$$U'(C(x;\Omega_t)) = \beta \mathrm{E}_t[R(\Omega_{t+1})U'(C(R(\Omega_{t+1})(x-c_{t,i})+y_{t+1,i};\Omega_{t+1}))] \quad \text{and} \quad C(x;\Omega_t) < x - \underline{k} \tag{17b}$$

where $R(\Omega_{t+1}) \equiv 1 + r(\Omega_{t+1})$. The expectation in (17) is over the distribution of $\xi_{t+1,i}$ and of $\Omega_{t+1}$. Since the household problem is concave, Eq. (17) together with the constraint (15) are both necessary and sufficient for a solution of the household problem. Notice that the solution depends on the aggregate state $\Omega_t$ only through the expected future dynamics of the factor prices.

### 2.4. Dynamics of the cross-sectional distribution

The cross-sectional distribution of cash on hand at the beginning of period $t$, after the realization of labor income, is continuous, because we have assumed that idiosyncratic productivity $\xi_{t,i}$ has a continuous distribution. Denote the density of cash-on-hand by $\phi_t(\xi)$. The distribution of assets at the end of period $t$ has a mass point of liquidity constrained households which hold the minimum level of assets $\underline{k}$. For $k > \underline{k}$, the distribution is continuous and we denote the density function by $\psi_t(k)$. The cross-sectional distribution function $\Psi_t(k)$ is defined as the fraction of households $i$ such that $k_{t,i} \leqslant k$. End-of-period aggregate capital is given by

$$\bar{K}_t = \int k \, \mathrm{d}\Psi_t(k) \tag{18}$$

The mass point satisfies

$$\Psi_t(\underline{k}) = \int_{-\infty}^{\chi_t} \phi_t(x) \, \mathrm{d}x \tag{19a}$$

Here $\chi_t$ is the value of cash-on-hand where the liquidity constraint starts binding. The cross-sectional density of capital $\psi_t(i)$ satisfies

$$\psi_t(K(x;\Omega_t))K_x(x;\Omega_t) = \phi_t(x), \quad x > \chi_t \tag{19b}$$

where $K_x(x;\Omega_t)$ denotes the partial derivative of the savings function. For the following formula, we use the abbreviation:

$$a_t(k) \equiv (1+r_t)k + T_t \tag{20}$$

for the level of assets after interest and lump sum subsidy, but before labor income. Then the density of cash-on-hand is related to last period's distribution of capital by

$$\phi_t(x) = \Psi_{t-1}(\underline{k}) f_t^y(x - a_t(\underline{k})) + \int_{\underline{k}}^{\infty} f_t^y(x - a_t(k))\psi_{t-1}(k) \, \mathrm{d}k \tag{21}$$

Eqs. (19) and (21) can be combined into a dynamic equation in the distribution of $k$.[3] Using

$$\begin{aligned} \int_{-\infty}^{\chi_t} \int_{\underline{k}}^{\infty} f_t^y(x - a_t(k))\psi_{t-1}(k) \, \mathrm{d}k \, \mathrm{d}x &= \int_{\underline{k}}^{\infty} \psi_{t-1}(k) \int_{-\infty}^{\chi_t} f_t^y(x - a_t(k)) \, \mathrm{d}x \, \mathrm{d}k \\ &= \int_{\underline{k}}^{\infty} \psi_{t-1}(k) F_t^y(\chi_t - a_t(k)) \, \mathrm{d}k \end{aligned} \tag{22}$$

we get

$$\Psi_t(\underline{k}) = \Psi_{t-1}(\underline{k}) F_t^y(\chi_t - a_t(\underline{k})) + \int_{\underline{k}}^{\infty} \psi_{t-1}(k) F_t^y(\chi_t - a_t(k)) \, \mathrm{d}k \tag{23a}$$

and

$$\psi_t(K(x;\Omega_t))K_x(x;\Omega_t) = \Psi_{t-1}(\underline{k}) f_t^y(x - a_t(\underline{k})) + \int_{\underline{k}}^{\infty} f_t^y(x - a_t(k))\psi_{t-1}(k) \, \mathrm{d}k, \quad \forall x > \chi_{t-1} \tag{23b}$$

### 2.5. Equilibrium of the model

We call the set of Eqs. (3), (7), (8), (17) and (23) the *theoretical model*. The 'natural' state variables of this model at time $t$ are the exogenous driving forces $Z_t$ and $\tau_t$ and the cross-sectional distribution of capital holdings inherited from the last period, $\Psi_{t-1}(k)$. We therefore choose $\Omega_t \equiv (Z_t, \tau_t, \Psi_{t-1}(k))$. A stationary equilibrium in $\Omega_t$ consists of a consumption function

[3] We could also write the dynamic equation in the distribution of $x$, but it is better to write it in $k$, because the distribution of $k$ at the end of period $t-1$ is the appropriate state variable in period $t$. The distribution of $x$ in $t$ is already affected by the aggregate shocks of period $t$.

$C(x;\Omega_t)$, a stochastic process of cross-sectional distribution functions $\Psi_t(k)$ and a process of lump sum transfers $T_t$ such that:

(1) the consumption function satisfies the Euler equation (17),
(2) the cross-sectional distribution satisfies the dynamic equations (23) and
(3) the transfer satisfies the government budget constraint (7).

To my knowledge, it has not been proven that a stationary equilibrium in $\Omega_t$ exists; Miao (2006) proves a recursive equilibrium in a larger state space including expected payoffs to households. In the following section, we will introduce a discrete approximation to this economy, and the computations can be seen as a 'constructive proof' that the approximate equation system has a stable solution for sufficiently small aggregate shocks.

## 3. Method

The solution method proposed here has three steps:

(1) Providing a finite representation of the economy. This includes representing the savings function $K(x;\Omega_t) = x - C(x;\Omega_t)$ by a vector $\mathbf{s}_t$; we will choose $\mathbf{s}_t$ as the value of $K(x;\Omega_t)$ at the knot points of a spline approximation. It also includes the vector $\mathbf{p}_t$ which represents the cross-sectional distribution of wealth $\Psi_t(k)$; we will choose for $\mathbf{p}_t$ the probability mass of households within specified intervals of capital holdings (histogram values). The set of equations for this finite number of variables is called the 'discrete model (DM)'.
(2) Computing the steady state of the economy, that is, the stationary economy when the aggregate shocks are identically zero. Concretely, aggregate productivity $Z_t$ and the tax on capital $\tau_t$ are constant. The idiosyncratic productivity shocks $\xi_{t,i}$ are the same as in the full model.
(3) Computing a first-order perturbation of the steady state solution in all the variables of the model. This gives a precise solution of the DM for small aggregate shocks $\varepsilon_Z$, $\varepsilon_\tau$.

### 3.1. Finite approximation of the model equations

#### 3.1.1. Savings, consumption and the Euler equation

The consumption function $C(x;\Omega_t)$ is very well behaved. At least for the steady state, we know (Carroll, 2004) that

(1) $C(x;\Omega_t)$ is concave in $x$.
(2) There is an $\chi_t$ such that $C(x;\Omega_t) = x - \underline{k}$ for $x \leqslant \chi_t$ and $C(x;\Omega_t) < x - \underline{k}$ for $x > \chi_t$.
(3) The limit $\lim_{x\to\infty} \partial C(x;\Omega_t)/\partial x$ exists.

This motivates the following smooth approximation of the savings function. I represent $K(x;\Omega_t)$ by $n_p + 1$ numbers, collected into the vector $\mathbf{s}_t$: the critical value $\chi_t$, and the values $k_{t,i}$, $i = 1,\ldots,n_p$ at $n_p$ knot points $x_{t,i}$ with $x_{t,i} > \chi_t$. The knot points are chosen as $x_{t,0} = \chi_t$ and $x_{t,i} = \chi_t + X_i$, $i = 1,\ldots,n_p$ with some fixed set of grid points $X_i$.[4] Off the knot points $x_{t,i}$, the function $K(x;\Omega_t)$ is then approximated by

$$\hat{K}(x;\mathbf{s}_t) = \begin{cases} \underline{k} & \text{for } x \leqslant \chi_t \\ CSI(x) & \text{for } \chi_t < x \leqslant x_{t,n_p} \\ k_{t,n_p} + CSI'(x_{t,n_p})(x - x_{t,n_p}) & \text{for } x > x_{t,n_p} \end{cases} \tag{24}$$

where $CSI(x)$ stands for 'cubic spline interpolation', the natural cubic spline that interpolates the points $(\chi_t, \underline{k})$, $(x_{t,i}, k_{t,i})$, $i = 1,\ldots,n_p$. [5] Beyond the last knot point $x_{t,n_p}$, we approximate the savings function by a straight line, with the slope given by the derivative of the spline at $k_{t,n_p}$.

We approximate the expectation over the individual productivity shock $\xi$ on the right-hand side of the Euler equation (17) by a finite sum, with quadrature points $\hat{\xi}_i$ and weights $\omega_i^\xi$, $i = 1,\ldots,n_\xi$. The details of the quadrature depend on the distribution of $\xi$, and will be discussed in Section 4.2.

Since we approximate the savings function by a spline with $n_p + 1$ degrees of freedom, we apply a collocation method and require the Euler equation (17) to hold with equality at the knot points $x_{t,i}$:

$$U'(\hat{C}(x_{t-1,i};\mathbf{s}_{t-1})) = \beta \sum_{j=1}^{n_\xi} \omega_j^\xi [R(\mathbf{p}_{t-1}, Z_t) U'(\hat{C}(\hat{X}_{ij};\mathbf{s}_t))] + \eta_{i,t}^c, \quad i = 0,\ldots,n_p \tag{25a}$$

[4] For the $X_i$ I take a grid with many points close to 0, because the consumption function has high curvature close to the kink point $\chi_t$.

[5] Because of the strong nonlinearity of the consumption function around $\chi_t$, my experience is that a spline approximation works better than Chebyshev polynomials.

where $\hat{C}(x;\mathbf{s}_t)$ is short for $x-\hat{K}(x;\mathbf{s}_t)$ and

$$\hat{X}_{ij} \equiv R(\mathbf{p}_{t-1},Z_t)(x_{t-1,i}-\hat{C}(x_{t-1,i};\mathbf{s}_{t-1})) + W(\mathbf{p}_{t-1},Z_t)\hat{\xi}_j + T_t \tag{25b}$$

We use the fact that gross interest rate $R(\mathbf{p}_{t-1},Z_t)$ and wage rate $W(\mathbf{p}_{t-1},Z_t)$ are a function of the end of last period's capital and today's productivity. Eq. (25a) uses the notation of Sims (2001): the $\eta^c_{i,t}$ are the expectation errors that result from the aggregate shocks (idiosyncratic shocks are handled by summing over the quadrature points). They are determined endogenously in the solution of the system (cf. Section 3.3).

### 3.1.2. Wealth distribution

There are several ways to parameterize a distribution function. One approach is to approximate the density function by a linear combination of known, smooth basis functions (polynomials, Gaussian densities, etc.) and force the dynamic equations (23) to hold on a set of collocation points. This is the approach taken in Bohacek and Kejak (2005). The downside of this approach is that we have to take care to enforce the non-negativity of the density function, which will involve the solution of a set of nonlinear equations. Here I follow a different path, which allows to compute the stationary distribution by a system of linear equations (cf. (34)). I first truncate the distribution of capital at a maximum level $\overline{k}$. Then I split the support into $n_d$ small intervals by a grid of points $\underline{k}=\kappa_0,\kappa_1,\ldots,\kappa_{n_d}=\overline{k}$ (I will choose $n_d=1000$ below). The truncation point $\overline{k}$ must be chosen such that in the steady state distribution, a very small fraction of households is close to $\overline{k}$. Then I characterize the distribution by the mass at the lower bound, $p^0_t \equiv \Psi_t(\underline{k})$, and the mass on the intervals $p^i_t \equiv \Psi_t(\kappa_i)-\Psi_t(\kappa_{i-1})$, $i=1,\ldots,n_d$, which are stacked into the vector $\mathbf{p}_t \equiv (p^0_t,p^1_t,\ldots,p^{n_d}_t)$. These probabilities satisfy

$$\begin{aligned} p^i_t &= \int_{\kappa_{i-1}}^{\kappa_i}\psi_t(k)\,\mathrm{d}k = \int_{x_{t,i-1}}^{x_{t,i}}\psi_t(K(x;\Omega_t))K_x(x;\Omega_t)\,\mathrm{d}x \\ &= \int_{x_{t,i-1}}^{x_{t,i}}\left[\Psi_{t-1}(\underline{k})f^y_t(x-a_t(\underline{k})) + \int_{\underline{k}}^{\infty} f^y_t(x-a_t(k))\psi_{t-1}(k)\,\mathrm{d}k\right]\mathrm{d}x \\ &= \Psi_{t-1}(\underline{k})[F^y_t(x_{t,i}-a_t(\underline{k}))-F^y_t(x_{t,i-1}-a_t(\underline{k}))] \\ &\quad + \int_{\underline{k}}^{\infty}\psi_{t-1}(k)[F^y_t(x_{t,i}-a_t(k))-F^y_t(x_{t,i-1}-a_t(k))]\,\mathrm{d}k \end{aligned} \tag{26}$$

where the $x_{t,i}$ are implicitly defined by

$$K(x_{t,i};\Omega_t)=\kappa_i \tag{27}$$

Formula (26) contains expressions of the form $\int_{\underline{k}}^{\infty}\psi_{t-1}(k)F^y_t(x-a_t(k))\,\mathrm{d}k$, which have to be evaluated for many different $x$. To compute this integral, I will assume that the probability mass is approximately uniformly distributed within each interval (constant density function):

$$\psi_t(k) \approx \frac{p^i_t}{\kappa_i-\kappa_{i-1}}, \quad k\in(\kappa_{i-1},\kappa_i), \ \forall i \tag{28}$$

Then we get, using (20), that

$$\begin{aligned} \int_{\underline{k}}^{\infty}\psi_{t-1}(k)F^y_t(x-a_t(k))\,\mathrm{d}k &= \sum_{i=1}^{n_d}\int_{\kappa_{i-1}}^{\kappa_i}\psi_{t-1}(k)F\left(\frac{x-a_t(k)}{w_t}\right)\mathrm{d}k \\ &\approx \sum_i \frac{p^i_{t-1}}{\kappa_i-\kappa_{i-1}}\int_{\kappa_{i-1}}^{\kappa_i}F\left(\frac{x-a_t(k)}{w_t}\right)\mathrm{d}k \\ &= \sum_i \frac{p^i_{t-1}}{\kappa_i-\kappa_{i-1}}\int_{(x-a_t(\kappa_i))/w_t}^{(x-a_t(\kappa_{i-1}))/w_t}F(\xi)\frac{w_t}{R_t}\,\mathrm{d}\xi \\ &= \sum_i \frac{p^i_{t-1}}{\kappa_i-\kappa_{i-1}}\frac{w_t}{R_t}\left[F^{(1)}\left(\frac{x-a_t(\kappa_{i-1})}{w_t}\right)-F^{(1)}\left(\frac{x-a_t(\kappa_i)}{w_t}\right)\right] \end{aligned} \tag{29}$$

Here, $F^{(1)}(x)$ is defined as the integral of the distribution function $F(x)$ with respect to $x$. To evaluate it quickly, we do the following preparatory steps at the beginning of the model solution. First, we compute $F^{(1)}(x)$ at a dense grid of $x$ by numerical integration. Then we use those values to compute a spline approximation of $F^{(1)}(x)$ which preserves the convexity of this function. I use Schumaker's quadratic splines (Judd, 1998, p. 231), exploiting that the derivative of $F^{(1)}(x)$ is $F(x)$. Afterwards we evaluate $F^{(1)}(x)$ always by spline interpolation.

These derivations are summarized in the following linear dynamic equation for $\mathbf{p}_t$:

$$\mathbf{p}_t = \Pi(\hat{\Omega}_t)\mathbf{p}_{t-1} \tag{30a}$$

where the aggregate state vector in the DM is given by $\hat{\Omega}_t \equiv (\mathbf{p}_{t-1},Z_t,\tau_t)$. The elements of the transition matrix $\Pi(\hat{\Omega}_t)$ are (cf. (23a), (26) and (29))

$$\Pi_{0,0}(\hat{\Omega}_t) = F^y_t(\chi_t - a_t(\underline{k})) \tag{30b}$$

$$\Pi_{i,0}(\hat{\Omega}_t) = F_t^y(x_{t,i} - a_t(\underline{k})) - F_t^y(x_{t,i-1} - a_t(\underline{k})), \quad i = 1, \ldots, n_d \tag{30c}$$

$$\Pi_{0,j}(\hat{\Omega}_t) = \frac{1}{\kappa_j - \kappa_{j-1}} \frac{w_t}{R_t} \left[ F^{(1)}\left( \frac{\chi_t - a_t(\kappa_{j-1})}{w_t} \right) - F^{(1)}\left( \frac{\chi_t - a_t(\kappa_j)}{w_t} \right) \right], \quad j = 1, \ldots, n_d \tag{30d}$$

$$\Pi_{i,j}(\hat{\Omega}_t) = \frac{1}{\kappa_j - \kappa_{j-1}} \frac{w_t}{R_t} \left[ F^{(1)}\left( \frac{x_{t,i} - a_t(\kappa_{j-1})}{w_t} \right) - F^{(1)}\left( \frac{x_{t,i} - a_t(\kappa_j)}{w_t} \right) - F^{(1)}\left( \frac{x_{t,i-1} - a_t(\kappa_{j-1})}{w_t} \right) + F^{(1)}\left( \frac{x_{t,i-1} - a_t(\kappa_j)}{w_t} \right) \right], \quad i,j = 1, \ldots, n_d \tag{30e}$$

Notice that $\Pi(\hat{\Omega}_t)$ depends on $\hat{K}(x; \mathbf{s}_t)$ through (27) where we use $\hat{K}(x; \mathbf{s}_t)$ for $K(x; \Omega_t)$.

#### 3.1.3. *The Discrete Model*

With the finite approximations defined above, we can now reduce the model to a finite set of equations in each period $t$. The *DM* consists of the Eqs. (3), (8), (25), (30a) and the government budget constraint

$$T_t = \tau_t \bar{K}(\mathbf{p}_t) \tag{31}$$

where $\bar{K}(\mathbf{p}_t)$ is aggregate capital, the first moment of the cross-sectional distribution of capital. It is computed using the following formula for the $j$-th moment of the capital stock

$$\int k^j \, \mathrm{d}\Psi_t(k) = \underline{k}^j \Psi_t(\underline{k}) + \int_{\underline{k}}^{\infty} k^j \psi_t(k) \, \mathrm{d}k \approx \underline{k}^j p_t^0 + \sum_i p_t^i \frac{\kappa_i^{j+1} - \kappa_{i-1}^{j+1}}{(j+1)(\kappa_i - \kappa_{i-1})} \tag{32}$$

Eq. (32) uses again the approximation (28).

These equations define, for each period $t$, a system of $n_p + n_d + 4$ equations in just as many variables: $\mathbf{s}_t$, $\mathbf{p}_t$ (note that $p_t^0$ is redundant since probabilities sum to unity), $T_t$, $Z_t$ and $\tau_t$. Of those, $\mathbf{p}_t$ are endogenous state variables, $Z_t$ and $\tau_t$ are exogenous state variables, while $\mathbf{s}_t$ and $T_t$ are control or 'jump' variables.

### 3.2. *Solving for the steady state*

Given the finite approximation of the model described above, we are now ready to solve for the stationary state of the model, characterized by constant values of the exogenous aggregate variables $Z_t = 1$ and $\tau_t = \tau^*$. Solving for the steady state then involves a one-dimensional fixed-point problem in the aggregate capital stock $K^*$. This problem is described by the following steps:

(1) Guess the steady state capital stock $K^*$. This determines the net interest rate $r^* = \mathcal{Y}_K(K^*, 1, 1) - \tau^* - \delta$, the wage $w^* = \mathcal{Y}_L(K^*, 1, 1)$ and transfers $T^* = \tau^* K^*$. Then compute the parameters of the savings and consumption function, $\mathbf{s}^*$, so as to satisfy the household optimality condition (25).
(2) Given $r^*$, $T^*$ and $\mathbf{s}^*$, compute the parameters of the cross-sectional distribution $\mathbf{p}^*$ so as to satisfy (30).
(3) Check whether the guess $K^*$ is consistent with the cross-sectional distribution of capital computed in the last step.

These steps are now described in more detail.

#### 3.2.1. *Finding the parameters of the consumption function*

In steady state, the Euler equation (25) reduces to

$$U'(\hat{C}(x_i^*; \mathbf{s}^*)) = \beta \sum_{j=1}^{n_\xi} \omega_j^\xi [(1 + r^*) U'(\hat{C}((1 + r^*) k^*(x_i^*) + w^* \hat{\xi}_j + T^*; \mathbf{s}^*))], \quad i = 0, \ldots, n_p \tag{33}$$

Given $r^*$, $w^*$ and $T^*$, this is a system of $n_p + 1$ nonlinear equations in the $n_p + 1$ components of $\mathbf{s}^*$. This can be solved without problems by a quasi-Newton algorithm (cf. the routine 'broydn.m' in the accompanying Matlab code).

#### 3.2.2. *Computing the invariant distribution*

Given the steady state savings function $\hat{K}(x; \mathbf{s}^*)$, we obtain from (30) the matrix $\Pi^*$ that describes the stationary transition dynamics of the wealth distribution. The finite parameterization of the steady state distribution $\mathbf{p}^*$ is then a non-zero solution of the linear equation system

$$\mathbf{p}^* = \Pi^* \mathbf{p}^* \tag{34}$$

From the theory of Markov chains, and the properties of the consumption function, one can show that there is a unique $\mathbf{p}^*$ such that $\sum_i \mathbf{p}_i^* = 1$.[6] It can be found as the eigenvector of $\Pi^*$ belonging to the unit eigenvalue. Since $\Pi^*$ is moderately sparse, this can be done quickly with the matlab 'eigs' command.[7]

#### 3.2.3. Solving for the aggregate capital stock

Starting from a guess of the aggregate capital stock $K^*$, we have found $\mathbf{p}^*$, which implies a value for $K^*$ by (18). Therefore we are left with a one-dimensional root-finding problem in $K^*$. This can be easily solved by bisection or by Brent's method (cf. Press et al. (1986, Section 9.3), also implemented in the Matlab command 'fzero').

### 3.3. Perturbation of the steady state

In Section 3.1.3 we saw that Eqs. (3), (8), (25), (30a) and (31) define, for each period, a system of $n_p + n_d + 4$ equations in just as many variables. Collect all those variables in the vector $\Theta_t$. Collect the expectation errors in the vector $\eta_t$, and the exogenous shocks in $\varepsilon_t \equiv (\varepsilon_{Z,t}, \varepsilon_{\tau,t})$. The system of equations can then be written in the compact form

$$H(\Theta_{t-1}, \Theta_t, \eta_t, \varepsilon_t) = 0 \tag{35}$$

Denote by $\Theta^*$ the steady state values of $\Theta_t$. Obviously, we have

$$H(\Theta^*, \Theta^*, 0, 0) = 0 \tag{36}$$

The linearized stochastic model is given by

$$H_1(\Theta^*, \Theta^*, 0, 0)(\Theta_{t-1} - \Theta^*) + H_2(\Theta^*, \Theta^*, 0, 0)(\Theta_t - \Theta^*) + H_3\eta_t + H_4\varepsilon_t = 0 \tag{37}$$

where $H_i$, $i = 1, 2, 3, 4$ denotes the partial derivative of $H$ with respect to its $i$-th argument. The analytical computation of the partial derivatives would be tedious. Automatic differentiation would be possible, but is not available in the Matlab environment that I use. I therefore choose the simplest approach, computing the derivatives by forward differencing. This is clearly not computationally efficient compared to automatic differentiation, but it turns out to be good enough.

The system of Eqs. (37) is in the form that can be solved by the package of Sims (2001). The number of state variables is $n_d + 2$, the number of jump variables is $n_p + 2$. In the numerical examples below, I use $n_d = 1000$ and $n_p = 100$. The policy function is then characterized by a $102 \times 1002$ matrix, the state transition function by a $1002 \times 1002$ matrix. Although dynamics is linear in all variables, this allows for a very rich dynamic structure. On a laptop with Intel Centrino 2 GHz processor and Matlab 7.1, it takes about 25 min to build the matrices in (37) by forward differencing, and then about 6 min to solve the model by Sims' method.

### 3.4. State space reduction

With our non-parametric approach, to solve precisely for the steady state distribution we need a fine grid, that is a big $n_d$. In the model with aggregate shocks, the $n_d$ parameters of the cross-sectional distribution, contained in $\mathbf{p}_t$, are then all state variables. For very high $n_d$, say 10,000, one can still solve for the steady state $\mathbf{p}^*$, but solving for the dynamics of the distribution is currently not feasible on a PC. Even if feasible, one might get a sufficiently precise solution by some form of state space reduction. In the following I propose a method to do this.

We can approximate the perturbations in $n_d$-dimensional space by a lower-dimensional smooth parameterization, for example by the linear combination of some smooth basis functions:

$$\mathbf{p}_t \approx \mathbf{p}^* + B\mathbf{b}_t \tag{38}$$

Here $B$ is a $n_d \times n_B$-matrix of known basis functions, and $\mathbf{b}_t$ is the vector of time-varying coefficients. In (38), we separate the parameterization of the invariant distribution $\mathbf{p}^*$ from the parameterization of the deviations, $B\mathbf{b}_t$. The vector of state variables characterizing the deviations of the distribution from the steady state is now $\mathbf{b}_t$ with dimension $n_B$.

The matrix $B$ represents a linear operation that transforms an $n_B$-dimensional parameter vector into an $n_d$-dimensional distribution. If $n_B < n_d$, we cannot expect (30a) to hold for all $n_d$ points of $\mathbf{p}_t$. To define exactly $n_B$ conditions on the distribution, we define an inverse operation to $B$, which assigns to any $n_d$-dimensional distribution an $n_B$-dimensional parameter vector. Denote the matrix of this operation by $P_B$. Then we can replace (30a) by the weaker condition

$$\mathbf{b}_{t+1} = P_B[\Pi(\hat{\Omega}_t)(\mathbf{p}^* + B\mathbf{b}_t) - \mathbf{p}^*] \tag{39}$$

[6] Even the richest household ($k = \overline{k}$) is driven down to $k = \underline{k}$ in a finite number of period if it always draws the worst possible income shock. Starting from $\underline{k}$, the household reaches $\overline{k}$ in a few periods, if it always draws the highest income shocks. Since the income distribution has a connected support, and the consumption function is continuous, the household reaches all intermediate capital values with positive probability in the same number of periods. Then $\Pi^*$ satisfies all the conditions of Stokey and Lucas (1989, Theorem 11.4).

[7] In the numerical experiments reported below, I have done this for matrices up to $n_d = 5000$. In cases where $n_d$ is much bigger, but sparse, one could use even more sophisticated procedures to find the invariant distribution (Virnik, 2007).

Obviously, we will require that $P_B B \mathbf{b}_t = \mathbf{b}_t$. This means that $P_B$ gives the coefficients of a projection into *span*($B$). One could choose the orthogonal projection, $P_B = (B'B)^{-1}B$. This is not optimal in general: the numerical experiments show that it is important to guarantee that this operation preserves the mean of the distribution. More generally, we may want to impose that the projection preserves a set of linear conditions:

$$M'\mathbf{p} = M'BP_B\mathbf{p}, \quad \forall \mathbf{p} \in \Re^{n_d+1} \tag{40}$$

for some given matrix $M$. This can be achieved by including $M$ into $B$, or by choosing

$$P_B = \begin{pmatrix} N'B'B \\ M'B \end{pmatrix}^{-1} \begin{pmatrix} N'B' \\ M' \end{pmatrix}, \quad N \equiv null(M'B) \tag{41}$$

I have done some experimentation with this, and it seemed best to use (41) and to impose only the preservation of the first moment, not of higher moments.

To summarize, the *reduced model* consists of the Eqs. (3), (8), (25), (31) and (39) in the variables $\mathbf{s}_t$, $\mathbf{b}_t$, $T_t$, $Z_t$ and $\tau_t$. For the choice of basis functions, cf. Section 5.3.2.

## 4. Parameter values and functional forms

I solve the model at annual frequency. For the model parameters I use standard values, $\beta = 0.95$, $\alpha = \frac{1}{3}$, $\delta = 0.1$. For the utility function I use CRRA

$$U(c) = \frac{c^{1-\gamma} - 1}{1 - \gamma} \tag{42}$$

with risk aversion parameter $\gamma \in \{1, 5\}$. The parameter $A$ in the production function is set to $A = (\beta^{-1} - 1 + \delta)/\alpha$, such that the steady state capital stock of the representative agent model equals 1.

For the technology shock I choose $\rho_z = 0.8$ and $\sigma_z = 0.014$, which are standard values, in annual terms. I choose the tax shock as uncorrelated, $\rho_\tau = 0$, to create unpredictable short-run redistributions. The variability of the tax shock is set, rather arbitrarily, to $\sigma_\tau = 0.02$. Taxes fluctuate around zero, so $\tau^* = 0$. Note that the variance of the shocks has no other role here than to scale the simulations, since our approximate solution is linear in the aggregate shocks.

### *4.1. The distribution of individual productivity*

I look at two specifications of the labor income distribution. The first one is the log-normal distribution, which has a thin tail. Following Carroll (2001), I choose standard deviation and mean of $\log(\xi_{t,i})$ as $\sigma = 0.2$ and $\mu = -0.5\sigma^2$, which gives an expected value of productivity of 1, as required by (9).

The second distribution is the Pareto distribution, which has a thick tail. The distribution function is given by

$$F(\xi) = 1 - C\xi^{-m}, \quad m > 1 \tag{43}$$

The distribution has support $\xi \in (C^{1/m}, \infty)$. Choosing the constant $C = ((m-1)/m)^m$ gives an expected value of 1. I choose $m = 2$, which is approximately what Saez (2001) finds for the US income distribution. For $m \leqslant 2$, the Pareto distribution has infinite variance. In the computations, however, I truncate the distribution at the point $\bar{\xi}$ where $F(\bar{\xi}) = 1 - 10^{-5}$, so that existence of moments is not an issue.

Comparing results for these two distributions with very different characteristics allows us to study the role of the wealth distribution in the dynamics of aggregate capital.

### *4.2. Quadrature formulas*

In the case of a log-normal distribution, I use Gauss–Legendre quadrature for the $\hat{\xi}_i$ and $\omega_i^\xi$ in Eq. (25a). This is implemented in Matlab by the routine 'qnwnorm' of the Compecon toolkit that comes with Miranda and Fackler (2002).

In the case of a Pareto distribution, I use the following procedure: I choose a set of cumulative probabilities $P_j$, $j = 0, \ldots, n_\xi$ with $P_0 = 0$, $P_j > P_{j-1}$ and $P_{n_\xi} = 1 - 10^{-5}$ (recall that the distribution is truncated at $1 - 10^{-5}$). Then I set

$$\omega_j^\xi \equiv P_j - P_{j-1} \tag{44a}$$

$$\hat{\xi}_j \equiv \frac{1}{P_j - P_{j-1}} \int_{q(P_{j-1})}^{q(P_j)} \xi f(\xi)\,\mathrm{d}\xi \tag{44b}$$

where $q(p)$ is the $p$-th percentile of the distribution of $\xi$. The quadrature points $\hat{\xi}_j$ are the expectations of $\xi$, conditional on $\xi \in (q(P_{j-1}), q(P_j))$. I choose the $P_j$ such that the $\omega_j^\xi$ are $10^{-5}$, $10^{-4}$, $10^{-3}$ and $10^{-2}$, respectively, for the highest intervals. The rest of the probability mass is divided into equal intervals. This makes sure that the top income earners are well represented in the discrete approximation. The numerical experiments reported in Section 5 use $n_\xi = 41$. Increasing this parameter has no significant effect on the results.

Note that this quadrature rule is only used for the Euler equations. The dynamics of the wealth distribution, developed in Sections 2.4 and 3.1.2, uses the distribution function $F(\xi)$.

## 5. Numerical examples

The numerical examples presented here illustrate how the method works with the simple model of Section 2, and which level of accuracy we can expect. After showing some of the dynamic properties of the model, I will investigate the accuracy of the computed solution, proceeding in two steps. First, I estimate the approximation error that comes from replacing the theoretical model of Section 2.5 by the DM of Section 3.1.3. Then I estimate the approximation error that arises from the state space reduction (Section 3.4). The error analysis should be understood as applying to the case of (infinitesimally) small aggregate shocks, even if the shocks are scaled to realistic size for convenience. I do not investigate the error arising from big aggregate shocks.

### *5.1. The dynamics of the economic model*

Table 1 shows some summary statistics (mean, standard deviation, skewness and kurtosis) of the cross-sectional distribution of income and wealth. The lognormal distribution has a standard deviation of about 0.2. It is skewed to the right, and has thin tails (the kurtosis is only slightly above that of the normal distribution, which equals 3). In contrast, the truncated Pareto distribution has much higher variance, is strongly skewed to the right and has fat tails (high kurtosis). Notice that, without truncation, these moments would be infinite. The wealth accumulation process levels out some of these differences: with the lognormal, the wealth distribution has higher skewness and kurtosis than the income distribution, the opposite happens with the Pareto distribution. The last column in the table shows the fraction of households that are close to the liquidity constraint ($k \leqslant \underline{k}+0.1$). We see first that this fraction is smaller when consumers are more risk averse ($\gamma = 5$), since these consumers try harder to avoid the region of low asset holdings. More interestingly, this fraction is much higher when the income distribution is very unequal and has fat tails (Pareto). Households with very high transitory income shocks save approximately the fraction $1 - r$ of this income. With fat tails, there are more of those people every period. This reduces the interest rate and drives more households down to the liquidity constraint.

Table 2 documents the effect of the different shocks on the dynamics of the capital distribution. As one would expect, the tax shocks create more variation in the cross-sectional variance and less variation in the mean of capital than the technology shocks. Interestingly, in the case of a distribution with fat tails, the variability of the cross-sectional variance is not as high as in the thin-tail distribution, but it generates a higher variability of mean capital. This is because there are more households in the region where the consumption function has strong curvature (cf. last paragraph), and therefore the wealth distribution matters more for the dynamics of the capital stock.

**Table 1**
Statistics of the income and stationary wealth distribution

| $F(\xi)$ | $\gamma$ | Std $\xi$ | Skew $\xi$ | Kurt $\xi$ | Mean $k$ | Std $k$ | Skew $k$ | Kurt $k$ | $p$ (bottom) |
|---|---|---|---|---|---|---|---|---|---|
| Logn | 1 | 0.202 | 0.61 | 3.68 | 1.004 | 0.759 | 1.41 | 5.63 | 0.03 |
| Pareto | 1 | 1.373 | 27.90 | 1675.32 | 1.111 | 1.926 | 6.36 | 72.53 | 0.13 |
| Logn | 5 | 0.202 | 0.61 | 3.68 | 1.025 | 0.704 | 1.26 | 5.06 | 0.02 |
| Pareto | 5 | 1.373 | 27.90 | 1675.32 | 1.534 | 2.242 | 6.22 | 71.32 | 0.05 |

$\xi$: individual labor productivity, $k$: individual capital holdings, Std: standard deviation, Skew: skewness, Kurt: kurtosis, $p$(bottom): probability mass in the range $(0, 0.1)$.

**Table 2**
Time series variability of capital

| $F(\xi)$ | $\gamma$ | Shocks $Z$ | | Shocks $\tau$ | | Shocks $Z, \tau$ | |
|---|---|---|---|---|---|---|---|
| | | $\bar{K}$ | Var $k$ | $\bar{K}$ | Var $k$ | $\bar{K}$ | Var $k$ |
| Logn | 1 | 3.11 | 5.81 | 0.13 | 31.43 | 3.12 | 31.96 |
| Pareto | 1 | 2.88 | 3.56 | 0.89 | 10.97 | 3.02 | 11.53 |
| Logn | 5 | 5.53 | 13.18 | 0.42 | 28.87 | 5.54 | 31.74 |
| Pareto | 5 | 4.04 | 4.17 | 1.33 | 12.54 | 4.26 | 13.22 |

$\bar{K}$: 100 times standard deviation of log aggregate capital, Var $k$: 100 times standard deviation of log of cross-sectional variance of $k$.

**Table 3**
Approximation error from discretization

| $F(\xi)$ | $\gamma$ | Euler resid. | | Approx. error distribution $k$ | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Max | Ave | $p$(top) | $\Psi_t(k)$ | $\bar{K}$ | Std $k$ | Skew $k$ | Kurt $k$ |
| Logn | 1 | −4.17 | −7.00 | −12.84 | −3.33 | −4.90 | −2.69 | −2.13 | −1.99 |
| Pareto | 1 | −3.25 | −7.01 | −14.72 | −4.26 | −4.26 | −3.52 | −4.33 | −2.70 |
| Logn | 5 | −3.91 | −7.01 | −12.70 | −3.40 | −4.24 | −2.74 | −2.19 | −2.10 |
| Pareto | 5 | −2.89 | −6.56 | −12.77 | −4.46 | −3.80 | −3.38 | −3.82 | −2.79 |

All numbers are given as decimal log. For the exact definition of the Euler residual, cf. the text. $p$(top) is the probability mass in the range $(0.9\bar{k}, \bar{k})$. Other statistics: difference between distribution with 1000 and with 5000 intervals. $\Psi_t(k)$: maximum difference of the distribution function. $\bar{K}$-Kurt($k$): relative difference in mean, standard deviation, skewness and kurtosis.

### 5.2. The approximation error from discretization

Table 3 shows some statistics to estimate the error that we make by the discretizations of Section 3.1. The first two columns document the residual of the household Euler equation (25a). It is computed as $(1/C_t)[C_t - U'^{-1}(\beta(1+r^*)\mathrm{E}_t(U'(C_{t+1}))]$, such that it expresses the residual as the relative error in consumption. The average absolute residual is computed as the unweighted average on the range between $\chi^*$ and $10\bar{K} + y^{\max}$, 10 times average capital plus the maximum realization of income. We see that the average residual is indeed very low, but the maximum residual is about 3 orders of magnitude higher. The maximum occurs in the region near the critical point where the liquidity constraint starts binding. Here the consumption function has high curvature and is more difficult to approximate by a spline, although I have allocated half of all the knot points of the spline in this region. One can further reduce the residual by increasing the number of parameters, $n_p$, which obviously comes at a higher computational cost.

The discretization of the cross-sectional distribution of wealth developed in Section 3.1.2 implies two types of errors. The first one comes from truncating the distribution at $\bar{k}$. The column '$p$(top)' in Table 3 shows that this error can be neglected: only about $10^{-13}$ of all households are in the range between $0.9\bar{k}$ and $\bar{k}$. The second and more serious error arises insofar as the assumption in Eq. (28), which says that the density is constant within intervals, is not satisfied. To get an estimate of the order of magnitude of this error, I compute the steady state distribution both for $n_d = 1000$ and 5000 and measure the difference. The table gives the maximum absolute distance between the two distribution functions, and the relative difference between the mean, standard deviation, skewness and kurtosis. We see that the mean is hit quite precisely, but higher moments become more and more difficult to pin down. Note that the difference in the invariant distributions is a tough criterion, since even minimal differences in the transition matrix can add up to sizable differences in the distribution over an infinite horizon. From this perspective, accuracy seems quite good.

### 5.3. The approximation error from state space reduction

What distinguishes the present method from other ones is that it uses a high number of state variables. In the examples of this section, I use about 1000 state variables. It is then natural to ask whether this is in fact necessary, or whether an approximate model using a much smaller state vector would already give sufficient accuracy. To address this question, I consider the linearized *DM* of Eq. (37), with $n_d = 1000$, as the true model. Then I compare the exact solution to the outcome of the reduced model of Section 3.4. Notice that, in the exact model, the law of motion that households use for making their decision is fully consistent with the dynamics of the distribution that results from the individual policy function (unlike methods where households base their decision on a small number of states).

#### 5.3.1. Measuring the accuracy of approximate models

I will use two different accuracy measures. The first measure, (52), addresses the main concern in Den Haan (2007), who suggests to use long-term forecasts to assess accuracy. In fact, the statistic (52) is the unconditional variance of an infinite horizon forecast. Den Haan (2007) also stresses the importance to report maximum, not average errors. This is crucial in a nonlinear context, but not so much here because we deal with linear processes. Assuming normally distributed shocks, the forecast errors will also be normal with mean zero, so the variance gives us complete information about the shock. Nevertheless, the second statistic, (53), is designed to provide information about the worst-case error. It is the maximum error of a forecast up to 100 periods, under an extreme, highly unlikely realization of the path of shocks.

*Accuracy measure* 1: *Unconditional mean squared error*: The solution of the exact model (37) can be written as

$$\Theta_t = A\Theta_{t-1} + B\varepsilon_t \tag{45}$$

where the matrices $A$ and $B$ are output of the solution algorithm. Any approximate model (denoted by a hat) that we consider has the same form

$$\hat{\Theta}_t = \hat{A}\hat{\Theta}_{t-1} + \hat{B}\varepsilon_t \tag{46}$$

While the state variables of the approximate model are not the same as of the exact model, both models are driven by the same shocks $\varepsilon$. We can then combine them into

$$\Theta_t^* = A^*\Theta_{t-1}^* + B^*\varepsilon_t \tag{47}$$

$$\Theta^* \equiv \begin{bmatrix} \Theta_t \\ \hat{\Theta}_t \end{bmatrix}, \quad A^* \equiv \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix}, \quad B^* \equiv \begin{bmatrix} B \\ \hat{B} \end{bmatrix} \tag{48}$$

The unconditional variance–covariance matrix $\Sigma_{\Theta^*}$ of $\Theta^*$ is obtained by solving the discrete Lyapunov equation

$$\Sigma_{\Theta^*} = A^*\Sigma_{\Theta^*}A^{*\mathrm{T}} + B^*\Sigma_\varepsilon B^{*\mathrm{T}} \tag{49}$$

The $i$-th non-central moment $m_{i,t}$ is a linear function of the state variables, both in the exact and the approximate models. So we can write the moments of the exact and any approximate model as $m_{i,t} = H_i\Theta_t$ and $\hat{m}_{i,t} = \hat{H}_i\hat{\Theta}_t$, respectively, with known row vectors $H_i$ and $\hat{H}_i$. From (45) and (46) it is clear that the unconditional expected error is zero:

$$\mathrm{E}(m_{i,t} - \hat{m}_{i,t}) = 0 \tag{50}$$

The unconditional mean squared error is then given by

$$\mathrm{E}[(m_{i,t} - \hat{m}_{i,t})^2] = [H_i \quad -\hat{H}_i]\Sigma_{\Theta^*} \begin{bmatrix} H_i^{\mathrm{T}} \\ -\hat{H}_i^{\mathrm{T}} \end{bmatrix} \tag{51}$$

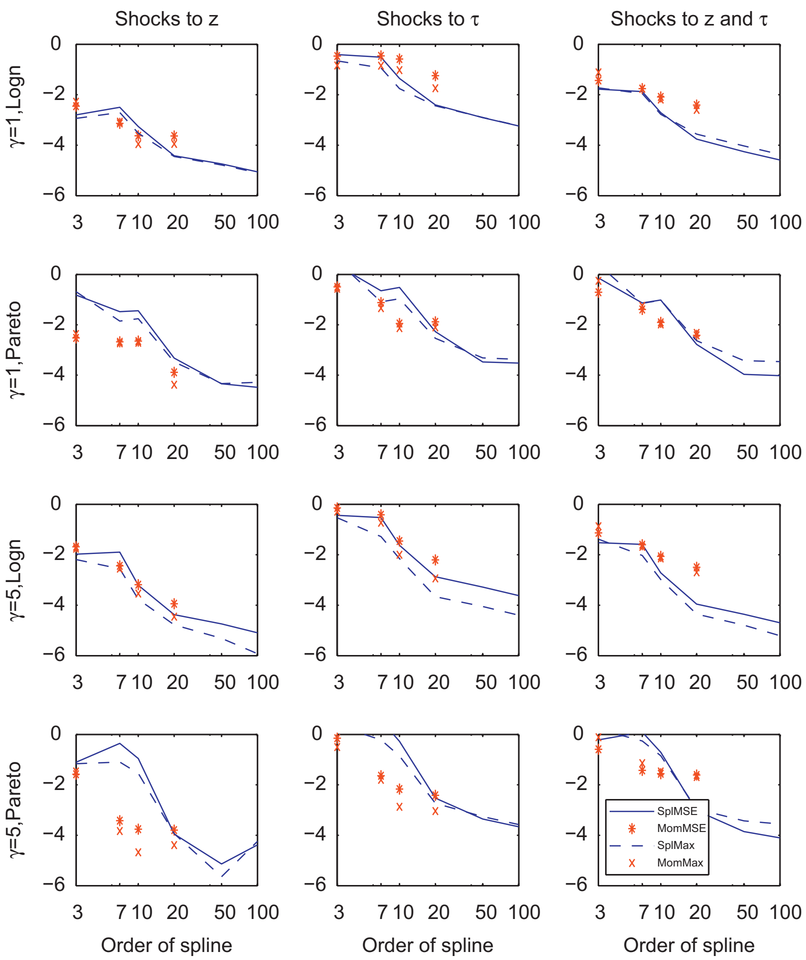


**Fig. 1.** Accuracy measures, log 10 of Eqs. (52) and (53).

Below we will report the root mean squared error, normalized by the standard deviation of $m_{i,t}$:

$$\sqrt{\frac{[H_i - \hat{H}_i]\Sigma_{\Theta^*}\begin{bmatrix} H_i^{\mathrm{T}} \\ -\hat{H}_i^{\mathrm{T}} \end{bmatrix}}{H_i \Sigma_\Theta H_i^{\mathrm{T}}}} \tag{52}$$

This statistic can be interpreted as the average of an infinite-horizon forecast error under a typical realization of the driving shocks.

*Accuracy measure* 2: *Maximum absolute error with extreme shock realization*: In the second experiment, we start from the deterministic steady state and simulate both the exact model (45) and the approximate model (46) for 100 periods, such that the same shock hits during 100 model periods. More precisely, I assume $\varepsilon_{Z,t} = \sigma_z$ and $\varepsilon_{\tau,t} = \sigma_\tau$ for $t = 1, \ldots, 100$ (in the case where $Z$ is the only shock, $\sigma_\tau = 0$ and vice versa). Then I report the maximum relative error

$$\frac{\max_{t=1}^{100} |m_{i,t} - \hat{m}_{i,t}|}{\max_{t=1}^{100} |m_{i,t}|} \tag{53}$$

*5.3.2. The accuracy of spline approximations of the distribution*

We now analyze the model with a reduced state space, presented in Section 3.4. For the basis functions (the columns of $B$), it is natural to try some standard smooth functions. Here I am going to use a spline basis. More precisely, since capital extends over a large state space, I use splines not in the original $k$, but in the nonlinear transformation $g = \log(k+1)$. To bring the model in form (46), let the vector $\hat{\Theta}_t$ comprise the variables mentioned at the end of Section 3.4.

Fig. 1 displays the approximation error of aggregate capital $\bar{K}$. The solid line gives the root mean squared error (52) of the reduced model. The dashed line gives the maximum error from an untypical realization of shocks, (53), for the same model. The stars and crosses in the graph refer to the VAR model explained in Section 6. The first surprising result is that the two lines are so close to each other. It means that, as long as we consider very small shocks, a long untypical realization of shocks creates similar aggregation errors as a an average realization. This might well be different for big shocks, where a long series of positive or negative shocks would drive us further away from steady state, such that nonlinear effects may become more important.

The picture shows some more things. First, the errors get reduced by 2 or 3 orders of magnitude if we use 100 state variables for the distribution rather than just 3 or 10. Second, the Pareto distribution requires a higher-dimensional approximation to achieve the same level of accuracy. Third, accuracy is much lower in the case of tax shocks. While these shocks cause only small fluctuations in $\bar{K}$, as can be seen from Table 2, they create complicated dynamics that is difficult to capture in a low-dimensional representation. The impulse response analysis in Section 6.3 will shed more light on this issue. I do not want to speculate what level of accuracy is 'sufficient'. However, in the case of tax shocks, a low-order approximation gives forecast errors of the same order of magnitude as the variation in capital itself, which is clearly unsatisfactory.

**Table 4**
Average errors of first four moments, from model with spline bases

| Shock | $n=3$ | | | | $n=100$ | | | |
|---|---|---|---|---|---|---|---|---|
| | 1 | 2 | 3 | 4 | 1 | 2 | 3 | 4 |
| *Lognormal income distribution*, $\gamma=1$ | | | | | | | | |
| $Z$ | −2.80 | 1.22 | 2.26 | 3.12 | −5.06 | −3.14 | −2.73 | −2.47 |
| $\tau$ | −0.41 | 1.37 | 2.16 | 2.94 | −3.24 | −2.76 | −2.59 | −2.43 |
| $Z,\tau$ | −1.78 | 1.36 | 2.17 | 2.95 | −4.59 | −2.79 | −2.59 | −2.44 |
| *Pareto income distribution*, $\gamma=1$ | | | | | | | | |
| $Z$ | −1.98 | 1.19 | 2.24 | 3.13 | −5.09 | −3.36 | −3.12 | −2.59 |
| $\tau$ | −0.44 | 1.35 | 2.18 | 3.00 | −3.62 | −3.00 | −2.87 | −2.88 |
| $Z,\tau$ | −1.53 | 1.29 | 2.20 | 3.04 | −4.70 | −3.10 | −2.91 | −2.77 |
| *Lognormal income distribution*, $\gamma=5$ | | | | | | | | |
| $Z$ | −0.82 | 2.73 | 4.21 | 4.63 | −4.48 | −2.20 | −0.86 | −0.48 |
| $\tau$ | 0.37 | 2.83 | 4.07 | 4.73 | −3.52 | −2.00 | −0.90 | −0.29 |
| $Z,\tau$ | −0.15 | 2.81 | 4.08 | 4.72 | −4.02 | −2.02 | −0.90 | −0.32 |
| *Pareto income distribution*, $\gamma=5$ | | | | | | | | |
| $Z$ | −1.10 | 2.55 | 3.94 | 4.17 | −4.38 | −2.29 | −1.04 | −0.86 |
| $\tau$ | 0.28 | 2.64 | 3.89 | 4.95 | −3.66 | −2.67 | −1.72 | −0.92 |
| $Z,\tau$ | −0.22 | 2.62 | 3.90 | 4.51 | −4.10 | −2.50 | −1.40 | −0.87 |

Results refer to (52). Numbers are given as decimal logs.

Table 4 presents approximation errors not only for $\bar{K}$, but for the first four non-central moments of the cross-sectional distribution of capital. The first column of both tables indicates the driving shock: '$Z$' means that the technology shock was the only driving force, '$\tau$' means that the tax shock was the only driving force, '$Z$,$\tau$' means that both shocks are present. The remaining columns provide the statistic (52) for different spline orders and different dimension of the VAR. The table shows that movements of higher order moments are very difficult to get right, particularly in the case of the Paretian distribution. This is in line with the findings of Section 5.2. With splines of order $n = 3$, the errors are larger than the fluctuations of the moments itself. Even with $n = 100$, prediction errors of the third and fourth moment are large. We learn that fluctuations of aggregate capital can be predicted very well despite of large errors in higher moments. This reflects again the basic finding that higher moments do not have a big impact in this model.

## 6. Forecasting future aggregate capital by moments of the distribution

For a model that is very similar to ours, Krusell and Smith (1998) find that future values of the aggregate capital stock, and therefore of factor prices, can be very well predicted using only the following information: today's aggregate level of capital, and current and future realizations of the exogenous driving shocks. We will now explore whether this is also true in the present model.

### 6.1. A VAR model in moments

We define the state vector $\hat{\Theta}_t \equiv [(m_t - m^*)^{\mathrm{T}} \log(z_t)\tau_t]^{\mathrm{T}}$, consisting of the exogenous driving processes and the vector of the first $n_m$ non-central moments $m_t$. In these variables, we define again an approximate model of the form (46). This time, the matrices $\hat{A}$ and $\hat{B}$ are not the outcome of a model solution, as they were in the model with spline approximations, but come from fitting a VAR model to the data series generated by the true model (45). Since we are dealing with linear models, we need not resort to simulations of the model, but can determine $\hat{A}$ and $\hat{B}$ by the asymptotic formulas given in Appendix A. In this way, our results are not contaminated by any sampling error.

Given $\hat{A}$ and $\hat{B}$, we can again compute the forecast errors described in Section 5.3.1. Notice that these errors are not the ones that would result from an application of the Krusell/Smith method. However, they are probably a good indication of the accuracy that this method can achieve in the present model. At least for the case of only one moment, I could confirm that a suitable version of the Krusell/Smith algorithm has approximation errors of the same order of magnitude as the VAR model with 1 moment.[8]

### 6.2. Numerical results

The stars and crosses in Fig. 1 display the root mean squared error (52) and the maximum error (53) for the VAR model of dimensions 1–4 (this is unrelated to the numbers on the $x$-axis). The RMSE for the first four non-central moments of the cross-sectional distribution of capital is given in Table 5, which is analogous to Table 4.

If only technology shocks are present, the results confirm what earlier papers have reported: a forecast based only on the first moment (average capital) plus the exogenous shock appears to be very precise. If we express the error in terms of $R^2$ (forecast error variance relative to variance of aggregate capital), as is often done in the literature, it looks quite impressive. In the log-normal case with $\gamma = 1$, for example, the infinite horizon error reported shown in the graph would have $R^2 = 0.999988$. For the one-step ahead forecast error, the $R^2$ would be 0.9999997.

Things change once we allow for tax shocks that redistribute capital. In this case, forecasts based on one moment have an error of between 30 and 70 percent of the fluctuation of capital, clearly an unsatisfactory outcome. Four moments are necessary to bring the forecast error down to about 1 percent. I do not claim that the specification with the big tax shocks is realistic; the purpose of the exercise is only to create a laboratory in which one can test a method that uses a high-dimensional representation of the distribution. It suggests that the success of the one-moment forecast in the model with technology shock does not mean that the distribution would not matter, but rather that the distribution does not change very much. Once we have sufficient time-series variation in the distribution of capital, the one-moment model breaks down. Admittedly, the tax-shock model is quite extreme in this respect. Table 2 tells us that the cross-sectional variance of capital fluctuates 10–300 times more than mean capital.

In any case, the good news is that the inclusion of moments up to order 4 increases accuracy significantly, usually by about two orders of magnitude. Table 5 shows that not only the first moment, but the higher moments themselves can be predicted quite accurately. This means that a Krusell/Smith-type algorithm which uses higher moments has the potential of being very accurate in this model. This should be explored in future work; the existing literature usually stops at one or two moments.

[8] Working with linear approximations to the model solutions, one can define a linearized version of the Krusell/Smith algorithm, in which the estimation uses the asymptotic formulas of Appendix A, and thus avoid any sampling error. Explaining this algorithm and its properties is beyond the scope of the present paper, but is the subject of ongoing work.

**Table 5**
Average errors of first four moments, from VAR in moments

| Shock | $d=1$ | $d=2$ | | $d=3$ | | | $d=4$ | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | 1 | 1 | 2 | 1 | 2 | 3 | 1 | 2 | 3 | 4 |
| *Lognormal income distribution*, $\gamma=1$ | | | | | | | | | | |
| $Z$ | −2.45 | −3.14 | −1.31 | −3.63 | −2.35 | −2.42 | −3.63 | −2.99 | −3.05 | −3.21 |
| $\tau$ | −0.47 | −0.46 | −0.70 | −0.58 | −1.30 | −1.59 | −1.25 | −2.12 | −2.40 | −2.64 |
| $Z,\tau$ | −1.43 | −1.76 | −0.67 | −2.08 | −1.46 | −1.72 | −2.41 | −2.08 | −2.34 | −2.55 |
| *Pareto income distribution*, $\gamma=1$ | | | | | | | | | | |
| $Z$ | −1.68 | −2.44 | −1.33 | −3.19 | −2.35 | −2.45 | −3.95 | −3.28 | −3.37 | −3.52 |
| $\tau$ | −0.14 | −0.41 | −0.64 | −1.45 | −1.82 | −2.11 | −2.20 | −2.84 | −3.11 | −3.34 |
| $Z,\tau$ | −1.13 | −1.61 | −0.83 | −2.06 | −1.64 | −1.86 | −2.50 | −2.30 | −2.51 | −2.63 |
| *Lognormal income distribution*, $\gamma=5$ | | | | | | | | | | |
| $Z$ | −2.52 | −2.67 | −2.08 | −2.65 | −2.00 | −2.53 | −3.89 | −3.06 | −3.22 | −2.97 |
| $\tau$ | −0.49 | −1.12 | −1.49 | −1.95 | −2.03 | −2.36 | −1.89 | −2.13 | −2.38 | −1.83 |
| $Z,\tau$ | −0.71 | −1.39 | −1.34 | −1.91 | −1.88 | −1.95 | −2.42 | −2.15 | −2.11 | −1.62 |
| *Pareto income distribution*, $\gamma=5$ | | | | | | | | | | |
| $Z$ | −1.58 | −3.42 | −2.56 | −3.76 | −2.91 | −2.98 | −3.79 | −3.29 | −3.43 | −3.90 |
| $\tau$ | −0.15 | −1.64 | −1.89 | −2.17 | −2.63 | −2.43 | −2.42 | −2.82 | −2.81 | −2.01 |
| $Z,\tau$ | −0.59 | −1.43 | −1.38 | −1.56 | −1.64 | −2.17 | −1.61 | −1.62 | −2.11 | −1.98 |

Results refer to (52). Numbers are given as decimal logs.

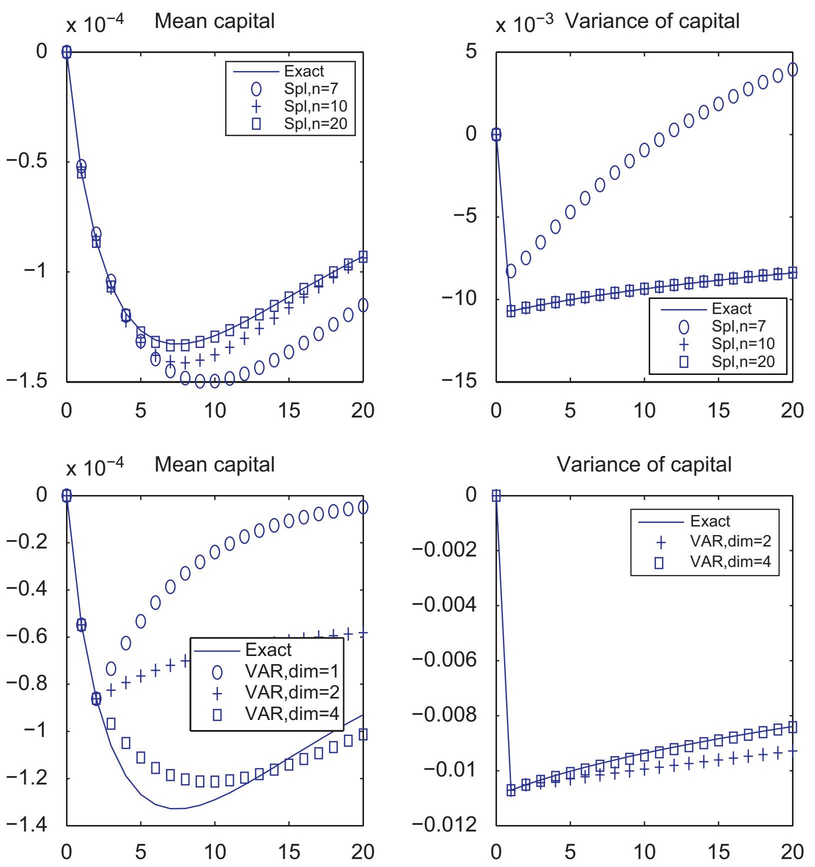


**Fig. 2.** Impulse response of aggregate capital to $\tau$-shock, lognormal distribution, $\gamma=1$.

### 6.3. Response to a tax shock

Fig. 2 displays impulse responses to a shock $\varepsilon_\tau$ in the capital tax. The two panels on the left-hand side show the response of mean capital, the panels on the right-hand side show the response of the cross-sectional variance of capital. The two panels in the upper part give the response of the exact model and the reduced model with spline bases. The panels in the lower part show the exact model and the VAR model in moments.

The tax shock brings a sudden reduction in the cross-sectional variance of capital. Since the consumption function is a concave function (cf. Section 3.1.1), this causes an increase in aggregate consumption: wealth was redistributed from richer to poorer households, who have a higher propensity to consume. This causes a reduction in the future capital stock. Notice that this effect really comes from redistribution, not from a shift in the consumption function: the expected future decrease in capital raises the real interest rate, which reduces consumption for a given level of assets. The shift in the consumption function therefore has a stabilizing effect and dampens the decrease of capital.

The model with spline approximations captures this, although the precision is low for $n<20$. The response of the VAR model yields some interesting insights. In the first two periods after the shock, the impulse responses are exact up to machine precision. This is because the VAR captures precisely the effect of the present shock, and indirectly also of the past shock, through the element $\tau$ in the state vector. From the third period onwards, however, the VAR model 'forgets' the shock (remember that $\rho_\tau = 0$). The VAR with one moment then predicts a slow return to the steady state, because, by construction, information on the distribution does not matter. The picture illustrates nicely that a solution based on only one moment does not even qualitatively capture the effect of changes in the distribution. As we already know from Section 6.2, adding higher moments greatly increases accuracy.

## 7. Conclusions

This paper has presented a new technique to solve models with a continuum of heterogeneous agents. The method allows to keep track of the dynamics of the cross-sectional distribution in a precise way. This was achieved by linearizing the model solution in the aggregate shocks, while maintaining the nonlinearity of the solution in the individual shocks.

Several things remain to be done. First, to provide a theoretical analysis of the perturbation approach along the lines of Jin and Judd (2002) or Judd (2004) for this class of models. Second, to carry the analysis on to higher order perturbations. Third, to identify and to study plausible and relevant models where the high-dimensional representation of the distribution function is essential to obtain a sufficient accuracy of the model.

## Acknowledgments

I am grateful to Ken Judd for very helpful discussions on the topics of this paper, and to James Costain, an anonymous referee and the editor, Wouter den Haan, for very useful comments on the first version of the paper. Support from Spanish ministry Grants SEJ2004-03619 and SEC2002-01601 is gratefully acknowledged.

## Appendix A. Asymptotic estimation of aggregate VAR

Assume the true model is

$$\Theta_t = A\Theta_{t-1} + B\varepsilon_t \tag{A.1}$$

where the $\varepsilon$ is a vector of i.i.d. shocks. We think of $\Theta$ as being a high-dimensional state vector, and want to estimate a VAR in a lower-dimensional vector of statistics $y$, which is related to $\Theta$ by

$$y_t = H\Theta_t \tag{A.2}$$

with known matrix $H$. The estimated model is

$$y_t = \hat{A}y_{t-1} + \hat{B}\varepsilon_t + u_t \tag{A.3}$$

The error term $u_t$ is supposed to capture the aggregation error. The normal equations for this model, which are equivalent to the ML estimator are

$$\frac{1}{T}\sum_{i=1}^{T} y_t[y'_{t-1} \;\; \varepsilon'_t] = \frac{1}{T}\sum_{i=1}^{T}[\hat{A} \;\; \hat{B}]\begin{bmatrix} y_{t-1} \\ \varepsilon_t \end{bmatrix}[y'_{t-1} \;\; \varepsilon'_t] \tag{A.4}$$

Using (A.2) and (A.1) we can write (A.4) as

$$\frac{1}{T}\sum_{i=1}^{T} H(A\Theta_{t-1} + B\varepsilon_t)[\Theta'_{t-1}H' \;\; \varepsilon'_t] = \frac{1}{T}\sum_{i=1}^{T}[\hat{A} \;\; \hat{B}]\begin{bmatrix} H\Theta_{t-1} \\ \varepsilon_t \end{bmatrix}[\Theta'_{t-1}H' \;\; \varepsilon'_t] \tag{A.5}$$

In the limit $T \to \infty$, the means converge to their unconditional expectations, and we get

$$[HA\mathrm{E}[\Theta\Theta']H' \; HB\mathrm{E}[\varepsilon\varepsilon']] = [\hat{A} \; \hat{B}] \begin{bmatrix} H\mathrm{E}[\Theta\Theta']'H' & 0 \\ 0 & \mathrm{E}[\varepsilon\varepsilon'] \end{bmatrix} \tag{A.6}$$

$\mathrm{E}[\Theta\Theta']$ is found by solving a discrete Lyapunov equation as in (49).

We apply this to the estimation problem of Section 6 by setting

$$y_t = \begin{bmatrix} m_t - m^* \\ \log z_t \\ \tau_t - \tau^* \end{bmatrix}$$

## References


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

Bohacek, R., Kejak, M., 2005. Projection methods for economies with heterogeneous agents. CERGE-EI Working Paper 258.

Campbell, J., 1998. Entry, exit, embodied technology, and business cycles. Review of Economic Dynamics 1 (2), 371–408.

Carroll, C.D., 2001. A theory of the consumption function, with and without liquidity constraints. Journal of Economic Perspectives 15 (3), 23–45.

Carroll, C.D., 2004. Theoretical foundations of buffer stock saving. NBER Working Paper No. 10867.

Den Haan, W.J., 1997. Solving dynamic models with aggregate shocks and heterogenous agents. Macroeconomic Dynamics 1, 355–386.

Den Haan, W.J., 2007. Assessing the accuracy of the aggregate law of motion in models with heterogeneous agents. Manuscript.

Haefke, C., Reiter, M., 2006. Endogenous labor market participation and the business cycle. IZA DP 2029.

Jin, H.-H., Judd, K.L., 2002. Perturbation Methods for General Dynamic Stochastic Models. Stanford University.

Judd, K.L., 1998. Numerical Methods in Economics. MIT Press, Cambridge, London.

Judd, K.L., 2004. Existence, uniqueness, and computational theory for time consistent equilibria: a hyperbolic discounting example. Manuscript.

Kim, J., Kim, S., Kollmann, R., 2007. Perturbation analysis of a heterogeneous agent economy with incomplete asset markets. Manuscript.

Krusell, P., Smith, A., 2006. Quantitative macroeconomic models with heterogeneous agents. In: Blundell, R., Newey, W.K., Persson, T. (Eds.), Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress. Cambridge University Press, New York, NY.

Krusell, P., Smith, A.A., 1998. Income and wealth heterogeneity in the macroeconomy. Journal of Political Economy 106 (5), 867–896.

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

Michelacci, C., Lopez Salido, D., 2007. Technology shocks and job flows. Review of Economic Studies 74, 1195–1227.

Miranda, M.J., Fackler, P.L., 2002. Applied Computational Economics and Finance. MIT Press, Cambridge, MA.

Press, W., Flannery, B., Teukolsky, S., Vetterling, W., 1986. Numerical Recipes. Cambridge University Press, Cambridge.

Preston, B., Roca, M., 2007. Incomplete markets, heterogeneity and macroeconomic dynamics. NBER Working Paper 13260.

Reiter, M., 2002. Recursive solution of heterogenous agent models. Manuscript, Universitat Pompeu Fabra.

Saez, E., 2001. Using elasticities to derive optimal income tax rates. Review of Economic Studies 68, 205–229.

Sims, C.A., 2001. Solving linear rational expectations models. Computational Economics 20 (1–2), 1–20.

Stokey, N.L., Lucas, R.E., 1989. Recursive Methods in Economic Dynamics. Harvard University Press, Cambridge, MA.

Virnik, E., 2007. An algebraic multigrid preconditioner for a class of singular M-matrices. SIAM Journal on Scientific Computing 29 (5), 1982–1991.