Identification and estimation of heterogeneous agent models: A likelihood approach*

Juan Carlos Parra-Alvarez†[a,b], Olaf Posch[b,c] and Mu-Chun Wang[c]

[a]Aarhus University, [b]CREATES, [c]Hamburg University

October 2017

Abstract

In this paper, we study the statistical properties of heterogeneous agent models with incomplete markets. Using a Bewley-Hugget-Aiyagari model we compute the equilibrium density function of wealth and show how it can be used for likelihood inference. We investigate the identifiability of the model parameters based on data representing a large cross-section of individual wealth. We also study the finite sample properties of the maximum likelihood estimator using Monte Carlo experiments. Our results suggest that while the parameters related to the household's preferences can be correctly identified and accurately estimated, the parameters associated with the supply side of the economy cannot be separately identified leading to inferential problems that persist even in large samples. In the presence of partially identification problems, we show that an empirical strategy based on fixing the value of one of the troublesome parameters allows us to pin down the other unidentified parameter without compromising the estimation of the remaining parameters of the model. An empirical illustration of our maximum likelihood framework using the 2013 SCF data for the U.S. confirms the results from our identification experiments.

Keywords: Heterogeneous agent models, Continuous-time, Fokker-Planck equations, Identification, Maximum likelihood.

JEL classification: C10, C13, C63, E21, E24.

---

*We would like to thank Benjamin Moll, Klaus Wälde, George-Marios Angeletos, Kjetil Storesletten and participants to the Fall Midwest Macro Meeting (Rochester, 2015), the 30th Annual Congress of the European Economic Association (Mannheim, 2015), the 9th Nordic Summer Symposium in Macroeconomics (Smögen, 2015) and the 10th NNH-UiO Workshop on Economic Dynamics (Oslo, 2015) for their helpful comments and discussions. The first two authors acknowledge financial support from the Center for Research in Econometric Analysis of Time Series (DNRF78), CREATES, funded by The Danish National Research Foundation.

†Corresponding author: Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, 8210 Aarhus V, Denmark. Email address: `jparra@econ.au.dk`

1 Introduction

Heterogeneous agent models have become an extensively used tool in macroeconomics for the study and evaluation of the welfare implications and desirability of business cycle stabilization policies. They have also been used to address questions related to social security reforms, the precautionary savings behavior of agents, employment mobility and wealth inequality. A comprehensive review on the developments made in the field during the last two decades can be found in Ríos-Rull (1995, 2001) and Heathcote et al. (2009). More recently, they have also started to be used for the study of monetary and fiscal policies, and their distributional implications (see Kaplan et al. (2016), Ozkan et al. (2016) and Wong (2016)).

Currently, the main workhorse in the heterogeneous agent literature is based on the contributions of Bewley (Undated), Huggett (1993) and Aiyagari (1994). Their theories are motivated by the empirical observation that individual earnings, savings, wealth and labor exhibit much larger fluctuations over time than per-capita averages, and accordingly significant individual mobility is hidden within the cross-sectional distributions. These ideas have been formalized with the use of dynamic and stochastic general equilibrium models of a large number of rational consumers that are subject to idiosyncratic income fluctuations against which they cannot fully insure due to market incompleteness.

To date, calibration is the standard methodology used to examine the quantitative properties of these models. Kydland and Prescott (1982) introduced calibration into macroeconomics with subsequent developments made by Prescott (1986), Cooley and Prescott (1995) and Gomme and Rupert (2007). This procedure fixes the value of the model parameters to those encountered in external sources, or to values such that the model generates moments that match certain observed aggregate macroeconomic statistics. Nonetheless, and despite being a very illustrative methodology for the study of a model dynamics, calibration does not allow us to make statements regarding the uncertainty surrounding these values, their statistical significance and on how well the models fit the data.

On the other hand, the use of econometric methods provide some important advantages over the calibration approach by allowing: (i) to impose on the data the restrictions arising from the economic theory associated with a particular model; (ii) to assess the uncertainty surrounding the parameter values which ultimately provides a framework for inference and hypothesis testing, (iii) the use of standard tools for model selection and evaluation. However, the use of econometric techniques seems novel for the type of heterogeneous agents models consider in this paper. Some notable exceptions include the recent contributions by Benhabib et al. (2015), Abbott et al. (2016) and Luo and Mongey (2017) where models without aggregate shocks are estimated using limited information methods, and those of Winberry (2016), Mongey and Williams (2017) and Williams (2017) where models with aggregate shocks are solved using the methodology proposed in Reiter

(2009) and then estimated using full information methods.

One possible explanation of why heterogeneous agent models have just recently started to be statistically estimated is that their solution imposes a computational burden that makes any econometric procedure infeasible. It is a well known fact that the numerical approximation of the density function of the state variables of the model increases considerably the computing time of the model's solution. However, Achdou et al. (2014) and Achdou et al. (2017) have made important advances in the solution of continuous-time heterogeneous agent models that have proven successful in reducing this computational complexities enabling the implementation of standard econometric methods.

The first contribution of this paper is to introduce a simple framework to estimate the structural parameters of heterogeneous agent models by exploiting the information content in the cross-sectional distribution of wealth. Our approach relies on the ability to compute the model's implied stationary probability density function of wealth which allows us to build the likelihood function of the model. Since the density function encompasses all the restrictions imposed by the economic model, the maximum likelihood estimator belongs to the class of full information estimators.

In general, the computation of the probability density function of wealth in heterogeneous agent models is not straightforward as it turns out to be a complicated endogenous and non-linear object that usually has to be numerically approximated. However, Bayer and Wälde (2010a,b, 2011, 2013), Achdou et al. (2014) and Gabaix et al. (2016) have recently suggested the use of Fokker-Planck equations for the derivation and analysis of endogenous distributions in macroeconomics[1]. These partial differential equations describe the entire dynamics of any probability density function in a very general manner without the need to impose any particular functional form. When combined with the standard Hamilton-Jacobi-Bellman equation that describes the optimal behavior of economic agents, they form a system of coupled partial differential equations that can be numerically solved with high degree of accuracy and efficiency on the entire state-space of the model using the finite difference methods described in Candler (1999) and Achdou et al. (2017).

A condition for the maximum likelihood estimator to deliver consistent estimates of the model parameters, and a valid asymptotic inference is identification (see Newey and McFadden (1986)). Roughly speaking, identification refers to the fact that the likelihood function must have a unique maximum at the true parameter vector and at the same time display enough curvature in all of its dimensions. Lack of identification leads to misleading statistical inference that may suggest the existence of some features in the data that are actually absent. Therefore, it is important to verify the identification condition prior to the application of any estimation strategy. The recent contributions of Canova and Sala (2009), Iskrev (2010), Komunjer and Ng (2011) and Ríos-Rull et al. (2012) point out in that direction by providing tools that can be used to study the identifiability

[1]The Fokker-Planck equations are also often called Kolmogorov Forward equations and both terms are equally used in the economic literature.

of parameters in structural macroeconomic models.

The second contribution of this paper is to investigate whether it is possible, and to what extent, to (locally) identify the structural parameters of heterogeneous agent models in a likelihood-based framework using a large cross-sectional sample of individual wealth. Given that the mapping between the deep parameters of the model and the likelihood function is highly nonlinear and not available in closed form, we investigate the identification power of the maximum likelihood estimator in an indirect way by using some of the simulation and graphical diagnostics proposed in Canova and Sala (2009).

To illustrate our approach, Section 2 introduces a continuous-time version of an otherwise standard Bewley-Hugget-Aiyagari model in which a large number of households face idiosyncratic and uninsurable income risk in the form of exogenous shocks to their productivity. In the context of this prototype economy, we then characterize and solve for the stationary competitive equilibrium which equip us with a time-invariant distribution of wealth that can be used for estimation and/or identification analysis. In Section 3 we show how to use this time-invariant density of wealth to compute the model's likelihood function. It also introduces the concept of identification within our maximum likelihood framework, and summarizes the different types of identification issues that could potentially arise in heterogeneous agent models.

Section 4 studies the behavior of the model's implied density function of wealth when the sampling process is known. We are specially interested in investigating the ability to identify the population density function along different dimensions of the parameter space. Since the analysis we conduct is independent of the data, we call it population identification analysis. Section 5 examines the finite sample properties of the maximum likelihood estimator using a Monte Carlo experiment. We pay particular attention to the potential biases and the precision of the estimates in different dimensions of the parameter space, and their implications for some of the model implied steady state macroeconomic aggregates.

A standard practice in macroeconomics when identification problems emerge is to fix the parameters that are believed to be unidentifiable to arbitrary values, and estimate the remaining ones. Section 6 investigates the consequences of following such an strategy. Section 7 provides an empirical illustration of our proposed framework by estimating the parameters of a Bewley-Aiyagari-Hugget model for the U.S. using the wealth data reported in the 2013 Survey of Consumer Finances. Section 8 concludes.

2 A prototypical heterogeneous agent model

For our study we consider a prototypical heterogeneous agent model *á la* Bewley-Hugget-Aiyagari set up in continuous-time following Achdou et al. (2017). In our economy there is no aggregate uncertainty and we assume that all aggregate variables are in their steady state, while at the individual level households face idiosyncratic uninsurable risk and variables change over time in a stochastic way.

2.1 Households

Consider an economy with a continuum of unit mass of infinitely lived households where decisions are made continuously in time. Each household consists of one agent, and we will speak of households and agents interchangeably. Household i, with $i \in (0,1)$, has standard preferences over streams of consumption, c_t, defined by:

$$U_0 \equiv \mathbb{E}_0 \int_0^\infty e^{-\rho t} u(c_t)\mathrm{d}t, \quad u' > 0,\ u'' < 0, \tag{1}$$

where $\rho > 0$ is the subjective discount rate, and where the instantaneous utility function is given by:

$$u(c_t) = \begin{cases} \frac{c_t^{1-\gamma}}{1-\gamma} & \text{for } \gamma \neq 1 \\ \log(c_t) & \text{for } \gamma = 1, \end{cases}$$

where $\gamma > 0$ denotes the coefficient of relative risk aversion. At time $t = 0$, the agent knows his initial wealth and income levels and chooses the optimal path of consumption $\{c_t\}_{t=0}^{\infty}$ subject to:

$$\mathrm{d}a_t = (ra_t + we_t - c_t)\mathrm{d}t, \quad (a_0, e_0) \in [\underline{a}, \infty) \times \mathcal{E}, \tag{2}$$

where a_t denotes the household's financial wealth per unit of time and r the interest rate. Wealth increases if capital income, ra_t, plus labor income, we_t, exceeds consumption, c_t. At every instant of time, households face uninsurable idiosyncratic and exogenous shocks to their endowment of efficiency labor units, e_t, making their labor income stochastic (see Castañeda et al. (2003)); w denotes the wage rate per efficiency unit which is the same across households and determined in general equilibrium together with the interest rate[2]. The fact that there are no private insurance markets for the household specific endowment shock can be explained, for example, by the existence of private information on the employee side, like his real ability, that could give rise to adverse selection and moral hazard problems. This would prevent private firms to provide insurance against income fluctuations. However, the wealth accumulation process in Equation (2) creates a mechanism used by agents to self-insure themselves against labor market shocks and allows for consumption smoothing.

Following Huggett (1993), the endowment of efficiency units can be either high, e_h, or low, e_l. The endowment process follows a continuous-time Markov Chain with state space $\mathcal{E} = \{e_h, e_l\}$ described by:

$$\mathrm{d}e_t = -\Delta_e \mathrm{d}q_{1,t} + \Delta_e \mathrm{d}q_{2,t}, \quad \Delta_e \equiv e_h - e_l \quad \text{and} \quad e_0 \in \mathcal{E}. \tag{3}$$

The Poisson process $q_{1,t}$ counts the frequency with which an agent moves from a high to a low efficiency level, while the Poisson process $q_{2,t}$ counts how often it moves from a low to a high

[2] Alternatively, the efficiency levels can be understood as productivity shocks following Heer and Trede (2003).

level. As an individual cannot move to a particular efficiency level while being in that same level, the arrival rates of both stochastic processes are state dependent. Let $\phi_1(e_t) \geq 0$ and $\phi_2(e_t) \geq 0$ denote the demotion and promotion rates respectively, with:

$$\phi_1(e_t) = \begin{cases} \phi_{hl} & e_t = e_h \\ 0 & e_t = e_l \end{cases}$$

and

$$\phi_2(e_t) = \begin{cases} 0 & e_t = e_h \\ \phi_{lh} & e_t = e_l. \end{cases}$$

Households in this economy cannot run their wealth below $\underline{a}$, where $a^n \leq \underline{a} \leq 0$, and $a^n = -{we_l}/{r}$ defines the natural borrowing constraint implied by the non-negativity of consumption. The effects of different values of $\underline{a}$ for the model implications are studied in Aiyagari (1994).

2.2 Production possibilities and macroeconomic identity

Aggregate output in this economy, Y, is produced by firms owned by the households. They combine aggregate capital, K, and aggregate labor, L, through a constant return to scale production function:

$$F(K, L) = K^{\alpha} L^{1-\alpha}, \quad \alpha \in (0, 1).$$

in order to maximize their profits.

We further assume that the aggregate capital stock in the economy depreciates at a constant rate, $\delta \in [0, 1]$. Since our focus is on the steady state, all the investment decisions in the economy are exclusively directed towards replacing any depreciated capital. Therefore the macroeconomic identity:

$$Y = C + \delta K \tag{4}$$

holds at every instant of time, where C denotes aggregate consumption, and δK aggregate investment. We have removed the temporal subscript t from all aggregate variables to indicate that the economy is in a stationary equilibrium.

2.3 Equilibrium

In this economy, households face uncertainty regarding their future labor efficiency. This makes their labor income and wealth also uncertain. Hence, the state of the economy at instant t is characterized by the wealth-efficiency process $(a_t, e_t) \in [\underline{a}, \infty) \times \mathcal{E}$ defined on a probability space $(\Omega, \mathcal{F}, G)$ with associated joint density function $g(a_t, e_t, t)$. In a stationary equilibrium this density is independent of time and thus it simplifies to $g(a_t, e_t)$.

As shown in Appendix A, for any given values of r and w, the optimal behavior of each of the households in the economy can be represented recursively from the perspective of time t by the

Hamilton-Jacobi-Bellman equation (HJB):

$$\begin{aligned}\rho V(a_t,e_t) &= \max_{c_t\in\mathbb{R}^+}\big\{u(c_t)+V_a(a_t,e_t)(ra_t+we_t-c_t)\\ &\quad +(V(a_t,e_l)-V(a_t,e_h))\phi_1(e_t)+(V(a_t,e_h)-V(a_t,e_l))\phi_2(e_t)\big\},\end{aligned} \tag{5}$$

where $V(a_t,e_t)$ denotes the value function of the agent. The first-order condition for an interior solution reads:

$$u'(c_t)=V_a(a_t,e_t) \tag{6}$$

for any $t\in[0,\infty)$, making optimal consumption a function only of the states and independent of time, $c_t=c(a_t,e_t)$. Equation (6) implies that in equilibrium, the instantaneous increase in utility due to marginally consuming more must be exactly equal to the increase in overall utility due to an additional unit of wealth.

Due to the state dependence of the arrival rates only one Poisson process will be active for each of the values in $\mathcal{E}$. This leads to a bivariate system of maximized HJB equations:

$$\rho V(a_t,e_l) = u(c(a_t,e_t))+V_a(a_t,e_l)(ra_t+we_l-c(a_t,e_t))+(V(a_t,e_h)-V(a_t,e_l))\phi_{lh}, \tag{7}$$

$$\rho V(a_t,e_h) = u(c(a_t,e_t))+V_a(a_t,e_h)(ra_t+we_h-c(a_t,e_t))+(V(a_t,e_l)-V(a_t,e_h))\phi_{hl}. \tag{8}$$

An interesting feature of our continuous-time setup as opposed to the discrete-time case, is that Equation (6) holds for all $a_t>\underline{a}$ since the borrowing constraint never binds in the interior of the state space. Therefore, the system of equations formed by (7) and (8) does not get affected by the existence of the inequality constraint $a_t\geq\underline{a}$, and instead gives rise to the following state-constraint boundary condition (see Achdou et al. (2017)):

$$V_a(\underline{a},e_t)\geq u'(r\underline{a}+we_t). \tag{9}$$

It can be shown that Equation (9) implies that $r\underline{a}+we_t-c(\underline{a},e_t)\geq 0$ and therefore the borrowing constraint is never violated.

On the other hand, the representative firm rents capital and labor from the household in perfectly competitive markets. Hence, in equilibrium the production factors are paid their respective marginal products:

$$r=\alpha K^{\alpha-1}L^{1-\alpha}-\delta \quad \text{and} \quad w=(1-\alpha)K^{\alpha}L^{-\alpha}, \tag{10}$$

where the steady state aggregate capital is obtained by aggregating the wealth held by every type of household, and similarly, the steady state aggregate labor is obtained by aggregating their efficiency labor units:

$$K = \sum_{e_t\in\{e_l,e_h\}}\int_{\underline{a}}^{\infty} a_t g(a_t,e_t)\,\mathrm{d}a_t, \tag{11}$$

$$L = \sum_{e_t \in \{e_l, e_h\}} \int_{\underline{a}}^{\infty} e_t g(a_t, e_t) \, \mathrm{d}a_t. \tag{12}$$

Equations (11) and (12) provide the link between the dynamics and randomness that occurs at the micro level with the deterministic behavior at the macro level.

A stationary equilibrium is defined as a situation where the aggregate variables and prices in the economy are constant, the joint distribution of wealth and efficiency units is time-invariant, and all markets clear. More specifically, while the distribution of wealth is constant for both the low and high efficient workers and the number of low and high efficient workers is also constant, the households are not characterized by constant wealth levels and efficiency status over time.

Definition 2.1 (Competitive stationary equilibrium) *A competitive stationary equilibrium is a pair of value functions $V(a_t, e_l)$ and $V(a_t, e_h)$, individual policy functions for consumption $c(a_t, e_l)$ and $c(a_t, e_h)$, a time-invariant density of the state variables $g(a_t, e_l)$ and $g(a_t, e_h)$, constant prices of labor and capital $\{w, r\}$, and a vector of constant aggregates $\{K, L, Y, C\}$ such that:*

1. *the consumption functions $c(a_t, e_l)$ and $c(a_t, e_h)$ satisfy Equations (7) and (8), i.e. they solve the household's allocation problem,*
2. *factor prices satisfy the first order condition in Equation (10), i.e. they solve the firm's problem,*
3. *markets clear, i.e. Equation (4) holds, with $C = \sum_{e_t} \int_{\underline{a}}^{\infty} c(a_t, e_t) g(a_t, e_t) \, \mathrm{d}a_t$, and production factors satisfy Equations (11) and (12),*
4. *the joint probability density function of the state variables is stationary, i.e. $\frac{\partial g(a_t, e_t)}{\partial t} = 0$ for all $(a_t, e_t) \in [\underline{a}, \infty) \times \mathcal{E}$.*

2.4 Distribution of endowments and wealth

Given its dependence on one continuous random variable and one discrete random variable, the stationary joint density function, $g(a_t, e_t)$, can be split into $g(a_t, e_h)$ and $g(a_t, e_l)$. Following Bayer and Wälde (2013), we refer to these individual probability functions as "subdensities". For each $e_t \in \mathcal{E}$, it follows that $g(a_t, e_t) \equiv g(a_t \mid e_t) p(e_t)$, implying that:

$$\int g(a_t, e_t) \, \mathrm{d}a_t = p(e_t), \tag{13}$$

where $p(e_t)$ is the stationary probability of having an efficiency endowment equal to e_t. Then, the (marginal) stationary density function of wealth can be computed as:

$$g(a_t) = g(a_t, e_h) + g(a_t, e_l). \tag{14}$$

Given our two state Markov process for the endowment of efficiency units it is possible to show that its stationary distribution is given by (see Appendix C):

$$\lim_{t\to\infty} p\left(e_h, t\right) \equiv p\left(e_h\right) = \frac{\phi_{lh}}{\phi_{hl}+\phi_{lh}}, \tag{15}$$

$$\lim_{t\to\infty} p\left(e_l, t\right) \equiv p\left(e_l\right) = \frac{\phi_{hl}}{\phi_{hl}+\phi_{lh}}. \tag{16}$$

Let $s\left(a_t, e_t\right) = ra_t + we_t - c\left(a_t, e_t\right)$ denote the optimal savings function for an individual with an efficiency endowment equal to e_t. As shown in Appendix B, the subdensities in Equation (14) correspond to the solution of the following non-autonomous quasi-linear system of differential equations known as (stationary) Fokker-Planck equations:

$$s\left(a_t, e_l\right) \frac{\partial}{\partial a_t} g\left(a_t, e_l\right) = -\left(r - \frac{\partial}{\partial a_t} c\left(a_t, e_l\right) + \phi_{lh}\right) g\left(a_t, e_l\right) + \phi_{hl} g\left(a_t, e_h\right), \tag{17}$$

$$s\left(a_t, e_h\right) \frac{\partial}{\partial a_t} g\left(a_t, e_h\right) = -\left(r - \frac{\partial}{\partial a_t} c\left(a_t, e_h\right) + \phi_{hl}\right) g\left(a_t, e_h\right) + \phi_{lh} g\left(a_t, e_l\right), \tag{18}$$

where the partial derivatives with respect to a_t describe the cross-sectional dimension of the density function. The system of equation (17)-(18) takes as given the optimal policy functions for consumption of individuals. This feature creates a recursive structure within the model that facilitates its solution: households and firms meet at the market place and make their choices taking prices as given. Prices in turn are determined in general equilibrium and hence depend on the entire distribution of individuals in the economy. Such distribution is determined by the optimal choices of households and the stochastic properties of the exogenous shocks.

2.5 Computation of the equilibrium

The solution of our prototype economy is not available in closed form. Therefore, for a given set of values of the structural parameters, the stationary competitive equilibrium in Definition 2.1 is numerically approximated on a discretized state space. The algorithm we use builds on earlier work by Candler (1999) and Achdou et al. (2017) and exploits the recursive nature of the model. It consists of two main blocks: (i) an outer block that takes the factor prices as given to compute in a recursive way the stationary equilibrium at the macro level; and (ii) an inner block that uses an implicit finite difference method in two stages. In the first stage it approximates the solution to the household's allocation problem at the micro level. Given the optimal consumption function obtained in stage one, the second stage approximates the stationary subdensities that solve the system of ordinary differential equations in Equations (17)-(18). Having approximated the density function, the factor prices are updated and the algorithm iterates until convergence. A detailed description of the algorithm and its implementation can be found in Appendix D.

3 Structural estimation: The likelihood function

While there is a broad consensus on the importance of heterogeneity in macroeconomics, there is less agreement on how these models should be taken to the data. To date, calibration is the standard approach used by researchers to map observations into parameter values of a structural model. Under this methodology, parameters are determined by minimizing the distance between a set of empirical moments and the same set of moments implied by the model[3], or by fixing the values of the parameters to those estimated in previous microeconomic studies, or to long-run averages of macroeconomic aggregates.

An alternative way to take structural models to the data is through formal econometric methods. In this section we show how to estimate the structural parameters of heterogeneous agent models using full information methods. The feasibility of our procedure is dictated, in general, by the use of continuous-time methods, and in particular by the Fokker-Planck equations that allow us to approximate the probability density function of wealth which can be then used to build the model's likelihood function. Our focus will be on the ability to estimate the model's parameters using the information content in the cross-sectional distribution of individual wealth. Appendix E shows how our framework can be extended to include information on individual income as proxied by the household's employment status.

Let $\mathbf{a} = [a_1, \ldots, a_N]$ be a sample of N i.i.d observations on individual wealth, and $\boldsymbol{\theta}$ a $\mathcal{K} \times 1$ vector of structural parameters to be estimated. In what follows, we assume that $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^{\mathcal{K}}$, where $\boldsymbol{\Theta}$ is the parameter space, assumed to be compact. For the model in Section 2, the likelihood function can be derived using the subdensity functions that solve Equations (17)-(18). According to the identity in Equation (14) the (marginal) stationary probability density function of wealth can be computed as:

$$g\left(a_n \mid \boldsymbol{\theta}\right) = g\left(a_n, e_l \mid \boldsymbol{\theta}\right) + g\left(a_n, e_h \mid \boldsymbol{\theta}\right) \tag{19}$$

for each $n = 1, \ldots, N$, where we have made explicit the dependence on the vector of parameter values, $\boldsymbol{\theta}$. For a given sample $\mathbf{a}$, the log-likelihood function can be then computed as:

$$\mathcal{L}_N\left(\boldsymbol{\theta} \mid \mathbf{a}\right) = \sum_{n=1}^{N} \log g\left(a_n \mid \boldsymbol{\theta}\right), \tag{20}$$

whereas the maximum likelihood (ML) estimator, $\hat{\boldsymbol{\theta}}_N$ is defined as:

$$\hat{\boldsymbol{\theta}}_N = \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{\arg\max}\ \mathcal{L}_N\left(\boldsymbol{\theta} \mid a_1, \ldots, a_N\right). \tag{21}$$

Since the density function of wealth, and hence the log-likelihood function, summarizes all the restrictions imposed by the economy model, our maximum-likelihood estimator belongs to the class of full information estimators

[3] See Castañeda et al. (2003) and Díaz-Giménez et al. (2014).

In practice, the ML estimation is carried out by means of an iterative procedure that requires solving the model for different values of the parameter vector $\boldsymbol{\theta}$. At each iteration the model is solved on the discretized state-space $\mathcal{A} \times \mathcal{E}$ using the algorithms described in Section 2. The wealth lattice is discretized using $I \leq N$ grid points on the partially ordered set $\mathcal{A} = [\min(\mathbf{a}), \max(\mathbf{a})]$. Once the density function of wealth has been approximated, the log-likelihood function is constructed in two steps: (i) For each $a_n \in \mathbf{a}$, use piece-wise linear interpolation to evaluate $g(a_n \mid \boldsymbol{\theta})$; (ii) Once $g(a_n \mid \boldsymbol{\theta})$ has been evaluated for all $a_n \in \mathbf{a}$, the log-likelihood function is computed using Equation (20).

A crucial assumption for the maximum likelihood estimator to deliver consistent estimates and valid asymptotic inference is that of identification. In general, a vector of parameters $\boldsymbol{\theta}$ is said to be identified if the objective function $\mathcal{L}(\boldsymbol{\theta} \mid \mathbf{a})$ has a unique maximum at its true value $\boldsymbol{\theta}_0$. Formally, the identification condition establishes that if $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$, then $\mathcal{L}(\boldsymbol{\theta} \mid \mathbf{a}) \neq \mathcal{L}(\boldsymbol{\theta}_0 \mid \mathbf{a})$, for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. Recently, Canova and Sala (2009) documented the existence of identification issues in the context of linearized DSGE models. These identification problems, which could also emerge in heterogeneous agent models of the type studied in this paper, are related to the shape and curvature of the objective function, and have been classified by the authors as follows[4]:

1. *Observational equivalence*: if two vectors of parameters, $\hat{\boldsymbol{\theta}}_1 \in \boldsymbol{\Theta}$ and $\hat{\boldsymbol{\theta}}_2 \in \boldsymbol{\Theta}$ deliver the same maximized objective function, they are said to be observational equivalent. In the maximum likelihood case, this occurs whenever $\mathcal{L}\left(\hat{\boldsymbol{\theta}}_1 \mid \mathbf{a}\right) = \mathcal{L}\left(\hat{\boldsymbol{\theta}}_2 \mid \mathbf{a}\right)$, and for any other $\boldsymbol{\theta} \in \boldsymbol{\Theta}$, $\mathcal{L}\left(\hat{\boldsymbol{\theta}}_j \mid \mathbf{a}\right) > \mathcal{L}\Big(\boldsymbol{\theta} \mid \mathbf{a}\Big)$, for $j = 1, 2$.

2. *Partial-identification*: if for some partition $\boldsymbol{\theta} = [\boldsymbol{\theta}_1, \boldsymbol{\theta}_2] \in \boldsymbol{\Theta}_1 \times \boldsymbol{\Theta}_2 = \boldsymbol{\Theta}$, $\mathcal{L}(\boldsymbol{\theta} \mid \mathbf{a}) = \mathcal{L}(f(\boldsymbol{\theta}_1, \boldsymbol{\theta}_2) \mid \mathbf{a})$, for all $\mathbf{a}$, and for all $\boldsymbol{\theta}_1 \in \boldsymbol{\Theta}_1$ and $\boldsymbol{\theta}_2 \in \boldsymbol{\Theta}_2$, where f is a continuous function, then $\boldsymbol{\theta}_1$ and $\boldsymbol{\theta}_2$ are said to be partially identified.

3. *Weak identification*: a subset of parameters in $\boldsymbol{\theta}$ is said to be weakly identified if the objective function, even though has a unique maximum, does not show enough curvature. In other words if there exists a $\hat{\boldsymbol{\theta}}$ such that $\mathcal{L}\left(\hat{\boldsymbol{\theta}} \mid \mathbf{a}\right) > \mathcal{L}\Big(\boldsymbol{\theta} \mid \mathbf{a}\Big)$ for all $\mathbf{a}$, and for all $\boldsymbol{\theta} \neq \hat{\boldsymbol{\theta}} \in \boldsymbol{\Theta}$. However, $\left\|\mathcal{L}\left(\hat{\theta}_i \mid \mathbf{a}\right) - \mathcal{L}\Big(\theta_i \mid \mathbf{a}\Big)\right\| < \epsilon$ for some $\theta_i \neq \hat{\theta}_i \in \boldsymbol{\Theta}, i = 1, \ldots, \mathcal{K}$.

4. *Asymmetric weak identification*: a group of parameters in $\boldsymbol{\theta}$ is said to exhibit asymmetric weak identification if the objective function is asymmetric in the neighborhood of the maximum, and its curvature is insufficient only in a portion of the parameter space. In other words if

[4] A fifth type of identification problem known as *under-identification* emerges in models where the solution is only locally valid, i.e. approximated using perturbation or linear quadratic methods. In that case, some of the model parameters disappear from the estimator's objective function because they are not present in the rational expectation solution of the model.

there exists a $\hat{\boldsymbol{\theta}}$ such that $\mathcal{L}\left(\hat{\boldsymbol{\theta}} \mid \mathbf{a}\right) > \mathcal{L}\left(\boldsymbol{\theta} \mid \mathbf{a}\right)$ for all $\mathbf{a}$, and for all $\boldsymbol{\theta} \neq \hat{\boldsymbol{\theta}} \in \boldsymbol{\Theta}$. However, $\left\|\mathcal{L}\left(\hat{\theta}_i \mid \mathbf{a}\right) - \mathcal{L}\left(\theta_i \mid \mathbf{a}\right)\right\| < \epsilon$ for some $\theta_i > \hat{\theta}_i \in \boldsymbol{\Theta}$ or for some $\theta_i < \hat{\theta}_i \in \boldsymbol{\Theta}, i = 1, \ldots, \mathcal{K}$.

Checking for identification in practice is difficult since the mapping from the structural parameters of the model to the objective function is highly nonlinear and usually not known in closed form. Therefore, the standard rank and order conditions used in linear models originally proposed in Rothenberg (1971) cannot be applied. In what follows, we use some of the simulation and graphical tools proposed in Canova and Sala (2009) to address the important question of identification in our nonlinear framework.

4 Population identification analysis

We begin our study of the proposed maximum likelihood estimator by analyzing the behavior of the density function of wealth when the sampling process is known. A basic prerequisite for carrying out valid inference about the parameter vector $\boldsymbol{\theta}$, is that distinct values of $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ imply distinct density functions. Therefore, this section investigates whether it is possible (or not) to distinguish the model's density function of wealth approximated using the true parameter values, $g\left(a \mid \boldsymbol{\theta}_0\right)$, from the density function approximated using a range of parameter values that differ from those in the population, $g\left(a \mid \boldsymbol{\theta}\right)$, with $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$. We refer to this approach as *population identification analysis* since it is independent of the data, and its conclusions remain valid even with samples of infinite size.

In what follows we assume that the population values of the structural parameters of the model, $\boldsymbol{\theta}_0$, are those given by the calibration in Table 1. These values are fairly standard in the literature. In particular, the labor efficiency process is calibrated to match the long run employment-unemployment dynamics of the US economy. In the model, time is measured in years and parameter values should be interpreted accordingly. Following Shimer (2005), the promotion rate is calibrated to match the monthly average job finding rate of 0.45, and the demotion rate is calibrated to match the monthly average separation rate of 0.034. The endowment level of high efficiency is normalized to one, and that of low efficiency unit is set to one-fifth of the employed, which lies in between the values used in Huggett (1993), and Imrohoroğlu (1989) and Krusell and Smith (1998). The transition rates for the Poisson processes are computed using Equations (15)-(16).

Since in our Bewley-Aiygari-Hugget economy the labor efficiency endowment process is completely exogenous, we will focus our attention on the ability to identify (and estimate) the supply side and household's preference parameters, while those parameters describing the endowment process will always remain fixed to the values in Table 1. To avoid unnecessary additional notation, the parameter vector $\boldsymbol{\theta}$ will refer exclusively to $\boldsymbol{\theta} = \{\gamma, \rho, \alpha, \delta\}$.

Formally, we say that the parameter vector $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ is identified if $g\left(a \mid \boldsymbol{\theta}\right) = g\left(a \mid \boldsymbol{\theta}_0\right)$. In order to make the identifiability condition operational we use the L_1 norm to measure the distance

Table 1. Population parameters, $\boldsymbol{\theta}_0$.

In the model, time is measured in years and parameter values should be interpreted accordingly. The endowment of efficiency units is given by:

$$\mathrm{d}e_t = -\Delta_e \mathrm{d}q_{1,t} + \Delta_e \mathrm{d}q_{2,t}, \quad \Delta_e \equiv e_h - e_l \quad \text{and} \quad e_0 \in \{e_h, e_l\},$$

where $q_{1,t}$ and $q_{2,t}$ are Poisson processes with intensity rates ϕ_{lh} and ϕ_{hl} respectively. The representative household has standard preferences defined by $U_t = \mathbb{E}_t \left[\int_t^\infty e^{\rho(s-t)} u(c_s)\, \mathrm{d}s\right]$ where $u(c_t) = \frac{c_t^{1-\gamma}}{1-\gamma}$. The macroeconomic identity in the stationary competitive equilibrium is given by:

$$Y = C - \delta K, \quad \text{where} \quad Y = K^\alpha L^{1-\alpha}.$$

Relative risk aversion, γ	2.0000
Rate of time preference, ρ	0.0490
Capital share in production, α	0.3600
Depreciation rate of capital, δ	0.1038
Endowment of high efficiency, e_h	1.0000
Endowment of low efficiency, e_l	0.2000
Demotion rate, ϕ_{hl}	0.5578
Promotion rate , ϕ_{lh}	7.3822

between two densities:

$$d(\boldsymbol{\theta}, \boldsymbol{\theta}_0) \equiv d\left(g(a \mid \boldsymbol{\theta}), g(a \mid \boldsymbol{\theta}_0)\right) = \sum_{i=1}^{I} \Big| g(a_i \mid \boldsymbol{\theta}) - g(a_i \mid \boldsymbol{\theta}_0) \Big| \tag{22}$$

where $g(a_i \mid \cdot)$ is the probability density function evaluated at grid point i for $i = 1, 2, \ldots, I$.

From a statistical point of view, the probability density function should contain all the relevant information about the value of the parameter vector $\boldsymbol{\theta}$. Therefore, if the distance function in Equation (22) features identification problems, we cannot hope to achieve identification of the model parameters using the likelihood of the data.

Figure 1 plots the shape of the distance function $d(\boldsymbol{\theta}, \boldsymbol{\theta}_0)$ for each of the elements of $\boldsymbol{\theta}$[5]. In each case, we vary one parameter at a time within an economically reasonable range while keeping all the remaining parameters at their population values. The population value of the parameter under analysis is represented by a dotted vertical line. The figure displays two important features. First, the distance function is uniquely minimized at $\boldsymbol{\theta}_0$, ruling out this way identification problems related to observational equivalence[6]. Second, the distance function exhibits enough curvature in the neighborhood of $\boldsymbol{\theta}_0$, suggesting strong identification power in each dimension of the parameter space.

Figure 1 only considers one dimension of the parameter space at a time which prevents us

[5] We also performed the population identification analysis on the full parameter set. Parameters associated with the labor endowment process are well identified among themselves, but poorly identified in association with other model parameters. The results are available upon request.

[6] Alternative metrics like the Chi-squared distance and the Earth's Moving Distance function give similar results. They are not displayed here for space considerations but are available upon request.

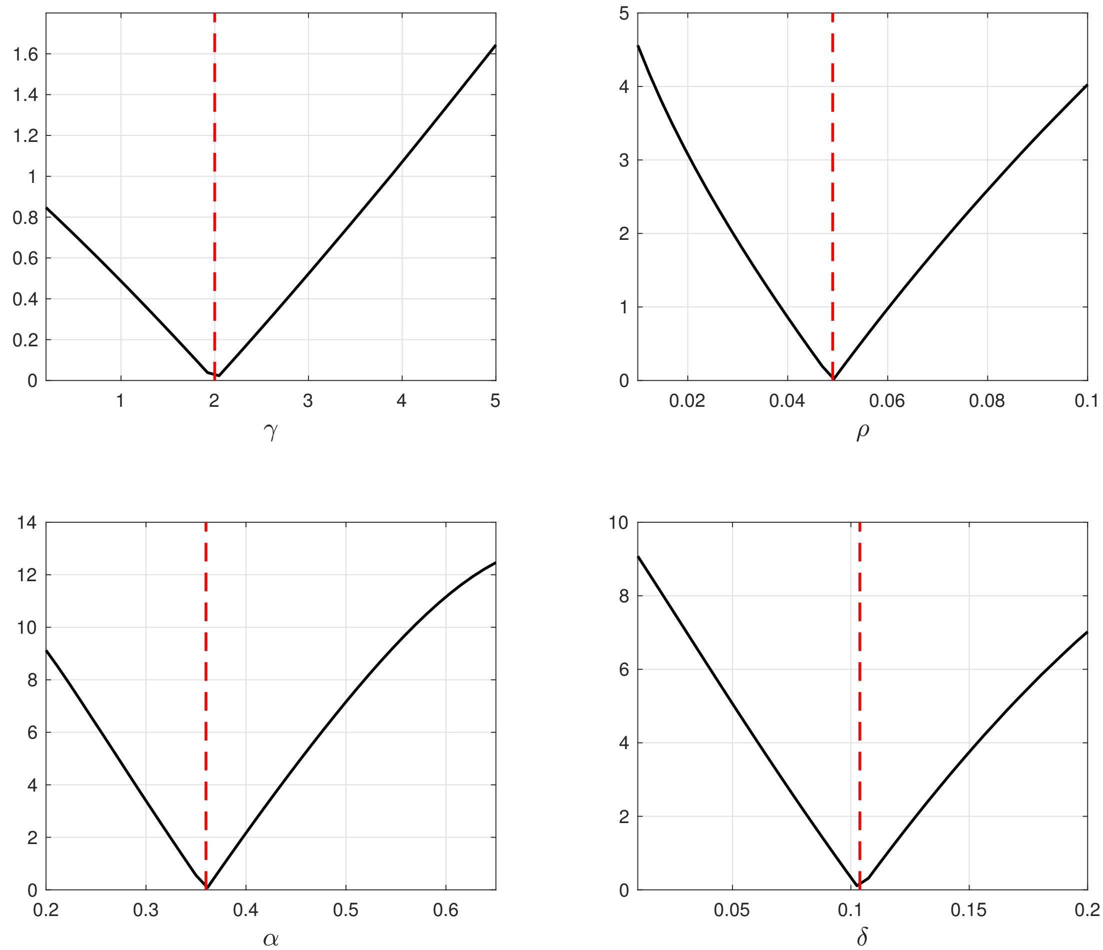

Figure 1. Distance function $d\left(g\left(a \mid \theta\right), g\left(a \mid \theta_0\right)\right)$**.** The graph shows the absolute deviation of the L_1 distance criterion as a function of the parameter space. The population values for the structural parameters, $\boldsymbol{\theta}_0$, are given in Table 1 and are represented by the dotted vertical line.

from detecting ridges in the objective function that might indicate partial identification problems. Therefore, Figure 2 plots the contours of the distance function for all pairwise combinations of parameters while keeping the remaining parameters fixed at their true value.

The upper right panel of Figure 2 reveals a ridge on the distance function for the supply side parameters, α and δ. This means that a proportional increase of both parameters may produce almost observational equivalent probability density functions of wealth, and therefore a clear indication of partial identification problems. In other words, small perturbations of the production function parameter and the depreciation rate have a large impact on the shape of the wealth distribution. Interestingly, the relationship is not linear. In fact, the ridge appears to be slightly concave with respect to α. The distance functions for the discount rate ρ and the supply side parameters, α and δ, also reveals a potential partial identification issue. However, the contours suggest that there is still enough curvature compared to the case of α against δ. The risk aversion parameter γ on the other hand, is strongly identified in combination with any other parameter.

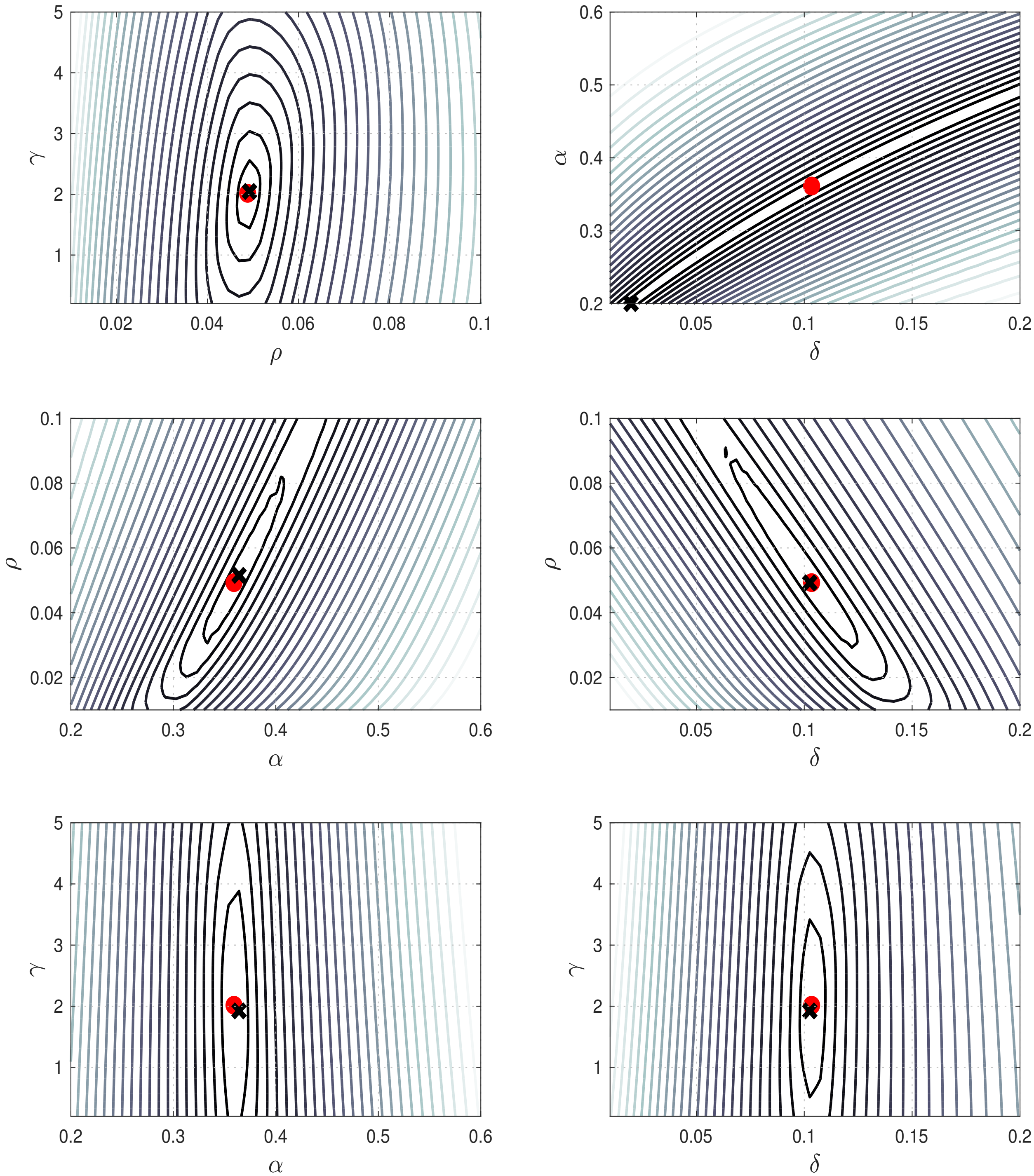

Figure 2. Distance surface for selected parameters. The graph shows the contour associated to the L_1 distance surface for selected combinations of parameters. The population values for the structural parameters, $\boldsymbol{\theta}_0$, are given in Table 1. A circle "●" indicates the true parameter values, and cross "×" the combination of parameters that deliver the minimum of the distance surface.

The partial identification issues just described are consistent with the identifiability problems that arise in the standard neoclassical growth model, as exemplified by its implied steady state

Table 2. Finite sample estimates.

The table reports finite sample estimates of the structural parameters of the model. Their absolute bias, their Monte Carlo standard errors (s.e.), and their mean squared errors (MSE) are obtained using 500 replications of the experiment. The sample size in each replication is given by N.

	$N=1000$			$N=5000$			$N=10000$			$N=50000$		
$\boldsymbol{\theta}$	Bias	s.e.	MSE	Bias	s.e.	MSE	Bias	s.e.	MSE	Bias	s.e.	MSE
γ	0.3345	1.3183	1.8464	0.0068	0.5374	0.2883	0.0196	0.3444	0.1188	−0.0251	0.1617	0.0267
ρ	−0.0048	0.0386	0.0015	−0.0071	0.0149	0.0003	−0.0041	0.0117	0.0002	−0.0018	0.0052	0.0000
α	0.1779	0.2454	0.0918	0.0964	0.1864	0.0440	0.0482	0.1645	0.0293	−0.0066	0.1100	0.0121
δ	0.2034	0.2754	0.1171	0.1052	0.1692	0.0396	0.0583	0.1394	0.0228	0.0066	0.0799	0.0064

capital-output ratio, K/Y. Assuming that the gap $(r-\rho)$ does not vary significantly with α and δ, and thus assumed to be relatively constant, the capital-output ratio of the Bewley-Hugget-Aiyagari economy is proportional to that of the neoclassical growth model, $K/Y \propto \alpha/(\rho+\delta)$. Therefore, for a given stationary capital-output ratio, and a given discount rate, the stationary equilibrium leads to a positive relation between α and δ similar to that depicted in Figure 2[7].

5 Finite sample properties

This section uses Monte Carlo simulations to investigate the properties of the ML estimator in finite samples by estimating the model of Section 2 using simulated cross-sectional data of individual wealth. The experiment is carried out by simulating 500 samples drawn from the model's population stationary probability density function $g(a \mid \boldsymbol{\theta}_0)$, each of them of size N, with $N \in \{1000, 5000, 10000, 50000\}$. For each sample, we estimate the model's parameters using the maximum likelihood estimator defined in Equation (21)[8].

The results of the Monte Carlo experiment are summarized in Table 2. For each N, it reports the absolute bias, the Monte Carlo standard errors, and the mean squared error (MSE). The Monte Carlo experiment reveals some important features that should be addressed. First, the bias on the risk aversion coefficient and the discount rate are within a reasonable range even in small samples. Their associated mean squared errors decrease by almost one order of magnitude as the sample size increase from $N=1000$ to $N=5000$. This in line with the results reported in Section 4 where it was shown that both parameters are well identified in the population. Second, the estimates of the capital share in production and the depreciation rate of capital exhibit a substantial positive bias that is far from negligible in small samples.

Figure 3 plots kernel density estimates of the ML estimates for both small ($N=5000$) and

[7] While the model's steady state capital-output ratio provides an intuitive way to understand the partial identification of α and δ, it should be kept in mind that Figure 2 depicts distance functions of the density of wealth which are not equal to the aggregate capital stock (see Equation (11)). Hence, although different combinations of α and δ lead to the same K/Y, the underlying stationary distributions of wealth need not to be the same.

[8] The initial value used in the estimation procedure corresponds to the true parameter vector.

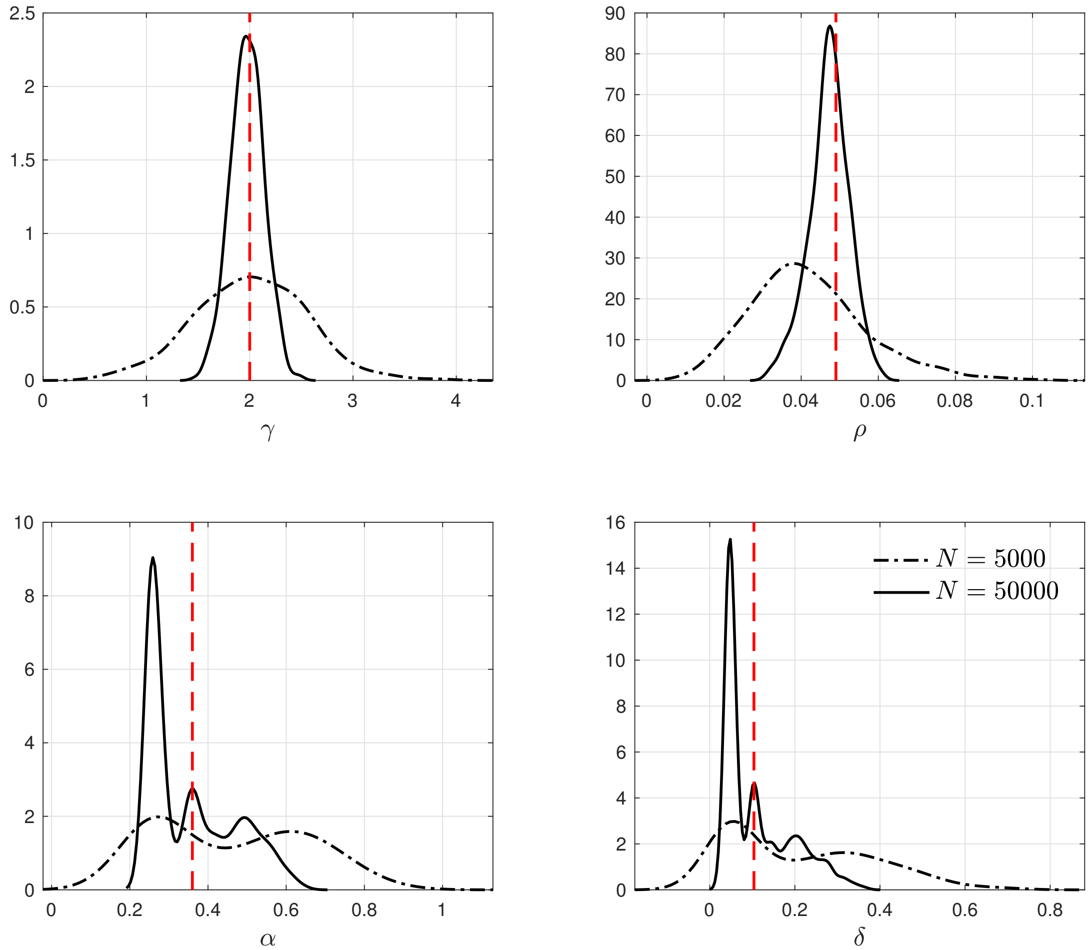

Figure 3. Finite sample distribution of parameter estimates. The graph plots the kernel density of estimated parameters across 500 random samples of size $N = 5000$ (dashed line) and $N = 50000$ (continuous line) generated from the true data generating process. The vertical line denotes the true parameter value.

large ($N = 50000$) samples. A dotted vertical line represents the true parameter value. The figure provides further evidence on the degree of accuracy with which γ can be identified, and the effects of using large samples on the bias reduction and correct identification of ρ, when the model is estimated on a cross-section of individual wealth. The figure also offers a clear picture of the meaningful biases in α and δ. Both parameters exhibit similar kernel density functions with multiple modes which reflect on the potential partial identification issues discussed in the previous section. These identification problems cannot be alleviated by increasing the sample size, as multiple modes still persist even for $N = 50000$.

While the results are somehow encouraging for large samples, as the parameter estimates approach their true values in the population, they suggest that the identification power of the maximum likelihood estimator in small samples is reduced in some dimensions of the parameter space when using data on a cross-section of individual wealth. In the case of the prototype economy of Section 2, the data deficiencies induced by the use of samples of reduced size are reflected in a

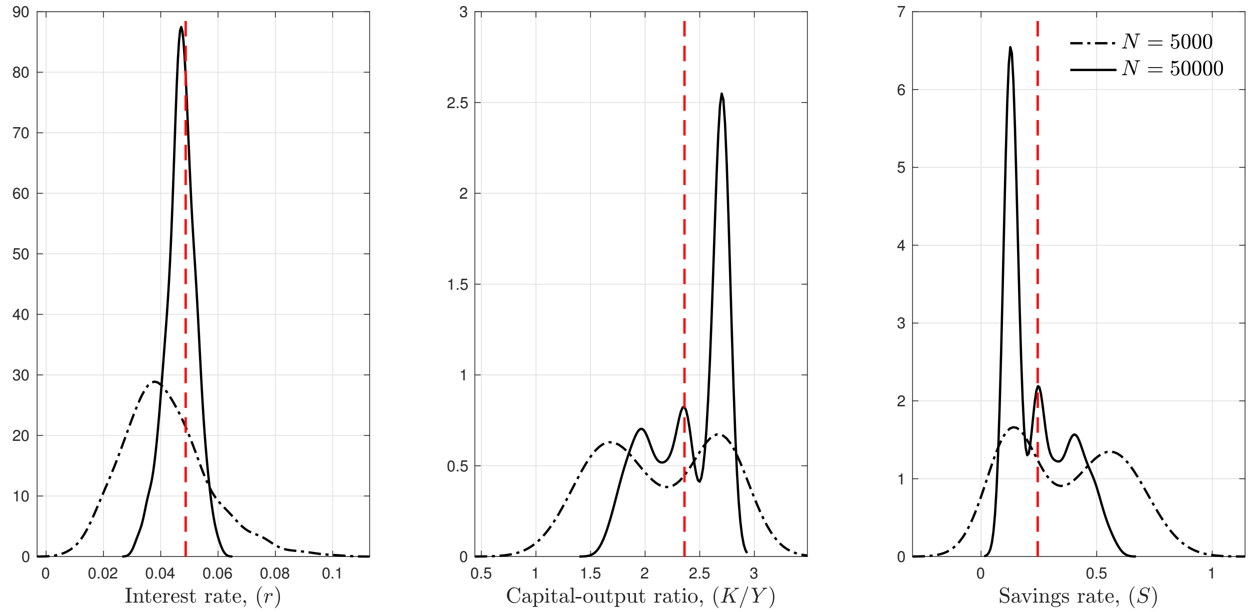

Figure 4. Steady state macroeconomic aggregates density estimates. The graph shows the implied distribution of the steady state values for the interest rate, the capital-output ratio, and the savings rate using the estimated parameters from the Monte Carlo experiment with $N = 5000$ and $N = 50000$. The savings rate is computed as $(Y - C)/Y$. The vertical line denotes the value in the population.

poor estimation of parameters related to the supply side of the economy. In particular, the results imply, on average, a higher use of capital in the production function, and a higher fraction of the depreciated capital stock.

What are the consequences of having biased estimates in some of the model parameters for the model implied macroeconomic aggregates? Figure 4 plots the implied distribution of the steady state interest rate, capital-output ratio and aggregate savings rate in both small and large samples. While the partial identification issues found between the supply side parameters hardly affect the model's implied steady state interest rate, they markedly contaminate the implied capital-output ratio and the aggregate savings rate, even in large samples. In the face of these partial identification issues, the estimated capital-output ratio and the savings rate cannot be correctly identified as suggested by the presence of multiple modes. Therefore, any economic interpretation or policy recommendation with regards to these two variables should be made with caution.

Overall, our Monte Carlo evidence suggests that while the parameters related to the household preferences can be identified and accurately estimated with the use of cross-sectional data on individual wealth, the parameters associated with the supply side of the economy cannot be separately identified leading to inferential problems that persist even in large samples. Following standard practice in macroeconomic, we next investigate the consequences of following an strategy where some of the troublesome parameters are calibrated at arbitrary values while estimating the remaining ones.

6 Calibration and estimation

Our findings indicate that across some dimensions of the parameter space the maximum likelihood estimator delivers biased and poorly identified estimates when the only data available is a finite cross-section on individual wealth. A common practice among economists to get around this obstacle is to calibrate those parameters that are problematic by fixing their value, and estimate the remaining ones. In the context of representative agent (linearized) DSGE models, Canova and Sala (2009) conclude that combining both approaches can lead to a biased inference and meaningless parameter estimates. To check whether this is the case in the heterogeneous agent model analyzed here, we use Monte Carlo simulations based on the presumption that the share of capital in the production function and the depreciation rate of capital cannot be separately identified from a cross-section of wealth.

Table 3 summarizes the results of our experiment when 500 samples of individual wealth have been drawn from the model's population stationary probability density function, each of them of size $N = 5000$. The table reports the absolute bias and MSE for each of the following scenarios: (i) no parametric restrictions; (ii) α is calibrated; (iii) δ is calibrated; (iv) α and δ are calibrated. The upper half of the table shows the results when the parameters in scenarios (ii)-(iv) are calibrated to their values in the population, while the lower half shows the results for the case in which they are miscalibrated to some fraction of their true value. The top panel of the table shows that, relative to the unrestricted model, fixing one of the supply side parameters (scenarios (ii) and (iii)) does not provide further improvements in the precision with which the model's preference parameters can be estimated. However, the estimation of partially identified parameters can be substantially improved. Most strikingly, the mean squared error of either α or δ is about two orders of magnitude lower compared to the case where both parameters are jointly estimated. For the case in which both troublesome parameters are calibrated to their true values, the biases and the MSE of the estimates of the preference parameters exhibit a considerably improvement.

Although the above results strongly suggest calibrating either one or both supply side parameters, such an approach may not carry an improvement in the identification and estimation accuracy of the model parameters if their calibrated values happen to be different from those in the population. The bottom panel of the table shows that even if one of the supply side parameters is miscalibrated to a fraction of its true value, the estimation of the non-calibrated parameter is still much more accurate compared to the case where both α and δ are jointly estimated. The results also indicate, as opposed to the findings in the top panel, that the biases in γ and ρ become much more severe when the supply side parameters are both miscalibrated. As argued in Canova and Sala (2009) this could be explained by the fact that the miscalibration changes the shape of the likelihood function inducing this way a considerable bias in the parameters that were originally free

Table 3. Conditional estimates.

The table reports finite sample estimates for a subset of the structural parameters of the model conditional on the calibrated values reported in the first row. The bias and the MSE are obtained using $M = 500$ samples, each of them of size $N = 5000$.

	No restrictions		$\alpha = \alpha_0$		$\delta = \delta_0$		$\alpha = \alpha_0$ and $\delta = \delta_0$	
$\boldsymbol{\theta}$	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
γ	0.0068	0.2883	−0.0636	0.2186	−0.0590	0.2236	−0.0034	0.0206
ρ	−0.0071	0.0003	0.0013	0.0002	0.0001	0.0001	0.0000	0.0000
α	0.0964	0.0440	-	-	0.0011	0.0003	-	-
δ	0.1052	0.0396	−0.0012	0.0002	-	-	-	-
	No restrictions		$\alpha = \frac{2}{3}\alpha_0$		$\delta = \frac{2}{3}\delta_0$		$\alpha = \frac{2}{3}\alpha_0$ and $\delta = \frac{2}{3}\delta_0$	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
γ	0.0068	0.2883	−0.0497	0.2119	−0.0992	0.2146	−1.9676	3.8713
ρ	−0.0071	0.0003	0.0027	0.0002	0.0012	0.0002	−0.0306	0.0009
α	0.0964	0.0440	-	-	−0.0572	0.0037	-	-
δ	0.1052	0.0396	−0.0686	0.0049	-	-	-	-

of identification issues.

Figure 5 complements our previous finding by plotting the kernel density estimates of the Monte Carlo ML estimates and their implied steady state interest rate, capital-output ratio and savings rate. Two cases are considered depending on whether the depreciation rate is fixed to its value in the population or miscalibrated to a fraction of the true value. The top panel shows that, regardless of the value chosen for δ, the preference parameters are correctly identified. Regarding the capital share, the results suggest that the identification problems discussed previously, which were associated to the presence of multiple modes as shown in Figure 3, disappear under this calibration strategy. However, although identifiable, the estimator of the capital share will exhibit a considerable bias when the depreciation rate is miscalibrated. The direction of this bias will follow that of the miscalibration. Similar conclusions are obtained for the density estimates of the implied macroeconomic aggregates. Independently of the calibration used, the interest rate remains well identified and accurately estimated, while the capital-output ratio and the savings rate are now free of identification issue but cannot be precisely pinned down when δ is miscalibrated.

Figure 6, on the other hand, provides evidence on the pervasive effects of calibrating both α and δ to a value different to that in the population. It plots the contour of the log-likelihood function for combinations of γ and ρ within a reasonable economic range using a random sample of size $N = 5000$ generated from the true model. The contour plot on the left column is generated when the capital share and the depreciation rate are fixed to $\alpha = \alpha_0$ and $\delta = \delta_0$, while the contour on the right column miscalibrates them to $\alpha = \frac{2}{3}\alpha_0$ and $\delta = \frac{2}{3}\delta_0$. For both cases, we have marked the combination of parameters that deliver the maximum of the log-likelihood function and the true values in the population. While fixing both troublesome parameters at the same time have no

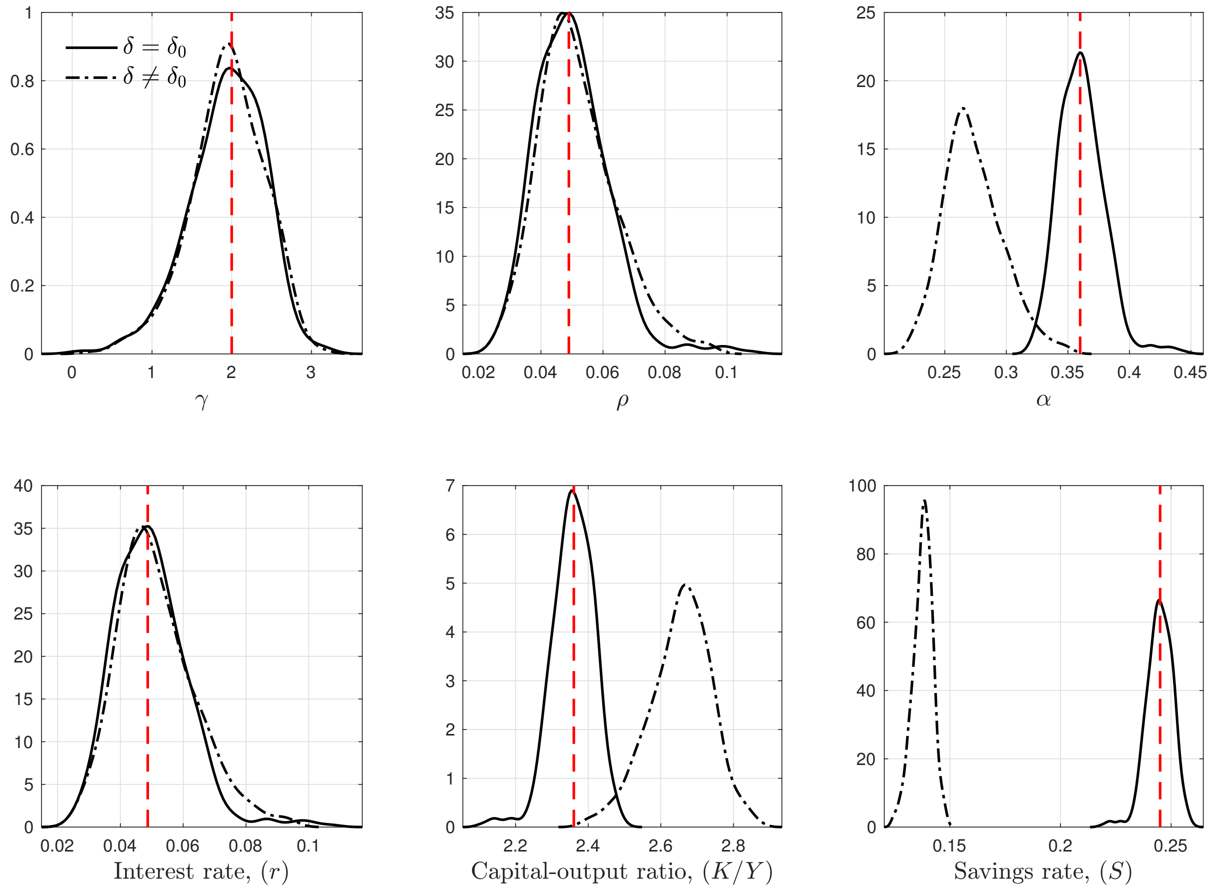

Figure 5. Parameter and macroeconomic aggregates densities with fixed δ. The graph shows the Monte Carlo implied distribution of the parameter estimates and the steady state interest rate, capital-output ratio, and savings rate when the depreciation rate is either fixed to its true value in the population, $\delta = \delta_0$, or miscalibrated to $\delta = \frac{2}{3}\delta_0$. The sample size is $N = 5000$. The vertical line denotes the value in the population.

effects on the estimation of the preference parameters if the calibration happens to coincide with their values in the population, the more realistic case in which they are miscalibrated to a fraction of their true values shows how the likelihood changes. Noticeably, the bivariate log-likelihood contour shifts dramatically downwards to the left, yielding much lower estimates for both γ and ρ. The bias induced by an strategy that calibrates both supply side parameters seems to follow the direction in which the fixed parameters where miscalibrated.

For finite samples, our results suggest to estimate the model parameters by fixing one of the supply side parameters *a priori*. This approach remains valid even when the parameter being fixed is miscalibrated given that those parameters related to the household's preferences remain well identified. While this strategy breaks down the inferential problems associated to the partial identification issues, it implies that the supply side parameter being freely estimated will face a similar bias to that introduced through the miscalibration, making any economic interpretation of its estimate difficult. Our results also advice not to calibrate both supply side parameters at the

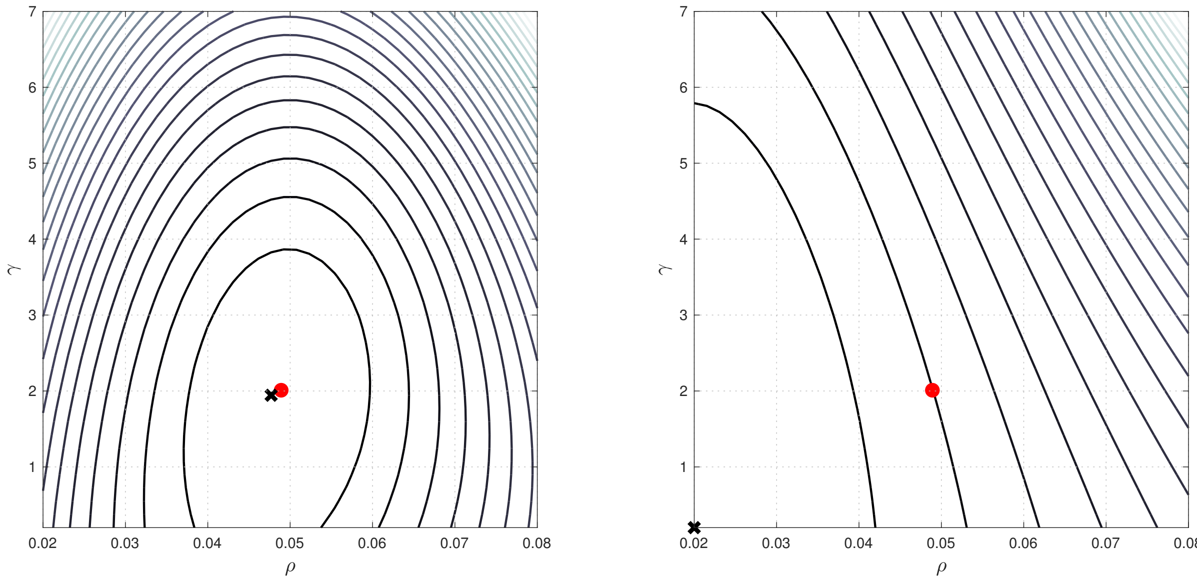

Figure 6. Log-likelihood function contour for fixed α and δ. The graph shows the contour of the log-likelihood function for combinations of γ and ρ within a reasonable economic range when a sample of size $N = 5000$ is generated from the true model. The contours to the left calibrate $\alpha = \alpha_0$ and $\delta = \delta_0$. The contours to the right calibrate $\alpha = \frac{2}{3}\alpha_0$ and $\delta = \frac{2}{3}\delta_0$. A circle "●" indicates the true combination of the parameter values, and a cross "×" the combination of parameters that deliver the maximum of the log-likelihood function.

same time since it creates non-negligible distortions in the distribution of parameter estimates that lead to serious biases on the remaining parameters.

7 Empirical illustration

This section provides an empirical illustration of our likelihood approach by estimating the parameters of the Bewley-Aiyagari-Hugget model of Section 2 using the wealth data reported in the 2013 Survey of Consumer Finances (SCF)[9].

Similar to our Monte Carlo experiments, we do not estimate the parameters of the income process directly since the model assumes earnings are completely exogenous, and hence there is no feedback from other features of the model that can be used to estimate its parameters. Furthermore, the discrete nature of our assumed income process does not allow us to exploit in a straight manner the information content in any cross-sectional data on individual income. Instead, we calibrate the parameters of the income process to match the 2013 employment-unemployment dynamics. In particular, we fix the endowment of high labor efficiency, e_h, to 1.76 to match the average annual hours worked (in thousands). The endowment of low labor efficiency, e_l, which should capture the unemployment income, is set to be 1/5 of e_h. The transition rates, ϕ_{lh} and ϕ_{hl}, are set,

[9] It should be stressed that our benchmark model is most likely misspecified as it cannot appropriately account for some of the main facts that characterize the wealth distribution in the U.S. Hence, the estimates reported are subject to considerably misspecification bias, and should be interpreted with caution.

Table 4. ML estimates with 2013 SCF data.

The table reports the maximum likelihood estimates of the model parameters for two different scenarios. It provides bootstrapped standard errors and confidence intervals based on 128 bootstrap repetitions.

	ML estimates	**Std. Error**	**95% CI**
	Scenario I: $\alpha = 0.36$		
γ	0.4840	0.1789	[0.1334, 0.8346]
ρ	0.0003	0.00003	[0.00029, 0.00039]
δ	0.0146	0.00016	[0.0143, 0.0150]
	Scenario II: $\delta = 0.1038$		
γ	0.4716	0.0498	[0.3739, 0.5692]
ρ	0.0010	0.00007	[0.0009, 0,0012]
α	0.6399	0.00064	[0.6387, 0.6412]

respectively, to 3.3353 and 0.2779, to match the 2013 monthly employment-unemployment transition probabilities of 0.24 (unemployment to employment) and 0.02 (employment to unemployment) obtained using the labor market statistics published by the U.S. Bureau of Labor Statistics. As our identification analysis suggests, we do not estimate α and δ jointly. Instead we provide estimation results for the two following scenarios: (i) α is calibrated to 0.36, while δ is freely estimated, and (ii) δ is calibrated to 0.1038, while α is freely estimated. Technical details on the estimation setup are provided in Appendix E.

Table 4 reports the estimation results together with bootstrapped standard errors and 95% confidence bands based on 128 bootstrap repetitions. The estimates of the risk aversion parameter, γ, are well below one. The point estimates are almost identical in both scenarios, a result that is in line with our previous findings that the identification of γ is not influenced by the calibration of the supply side parameters. The estimates of the discount rate are in both cases extremely low, suggesting that our benchmark model requires very patient households to be able to match the observed wealth of households. The point estimates of ρ exhibit substantial differences across scenarios with α fixed yielding a much lower estimate than the case of having δ fixed.

The estimates of the supply side parameters in each of the two scenarios reveal a very interesting fact. Relative to their calibrated values, the actual estimate of α is much higher, while the actual estimate of δ is much lower. Being this the case, a high estimate of α is accompanied by a high calibrated value of δ, and a low estimate of δ is accompanied by a low calibrated value of α. This result is in line with the partial identification issues reported earlier, where proportional increases (or decreases) on both parameters yield almost identical likelihood functions.

As mentioned above, our benchmark model is very unlikely to provide a correct specification of the underlying wealth data. Table 5 provides some wealth statistics computed from the data and the estimated model for each of the two scenarios. It reports the Gini coefficient and the percentage of

Table 5. Wealth Inequality: Data vs. Model.

The table compares the Gini coefficient and the distribution of wealth across top percentiles from the SCF data to those implied by the estimated models.

		% wealth in top		
	Gini Coefficient	1%	5%	20%
SCF 2013 data	0.8262	34.47	61.94	85.61
Scenario I: $\alpha = \alpha_0$	0.3781	5.39	17.56	44.44
Scenario II: $\delta = \delta_0$	0.4653	8.58	23.89	52.05

total wealth held by the top 1, 5 and 20 percentiles computed from the model's implied cumulative distribution function (CDF) as shown in Appendix B. As expected, the observed wealth is much more concentrated than what the estimated models predicts, with the case of δ fixed yielding a slightly higher concentration of wealth than the case of α fixed. The poor fit of the wealth statistics obtained from our estimated model is in line with the results usually reported for the type of heterogeneous agent model studied in this paper, as documented previously in Quadrini and Ríos-Rull (1997), Cagetti and De Nardi (2008), and Benhabib and Bisin (2017).

As a robustness check, we extend our estimation framework to exploit the stationary joint distribution of wealth and efficiency implied by the Bewley-Hugget-Aiyagari model, $g(a_t, e_t)$. We proxy the efficiency level e_t by the employment status of the household's head reported in the SCF[10]. The details on how to extend the log-likelihood function can be found in Appendix E. Table 6 reports the maximum likelihood estimates, the standard errors and 95% confidence bands, as well as the implied Gini coefficient and distribution of wealth across top percentiles for the scenario ii). The results suggest that the estimates in Table 4 are robust to the inclusion of a measure of income data. The parameter estimates remain statistically significant and do not change dramatically relative to the case where the model is estimated using just a cross-section on individual wealth. However, we suspect that the use of income data in such a simplistic way may not be completely informative to the estimation process and hence does not have a major impact on the model parameters identification. A future extension of our framework should include a continuous process for income as in Achdou et al. (2017) that allows to exploit in a better way the information content in a cross-section of individual earnings.

[10] It should be kept in mind that this is just a crude approximation to the model's discrete income dimension since the employment status data available in the SFC is for the household's head, whereas the data on wealth is for the entire household.

Table 6. ML estimates and wealth inequality statistics with 2013 SCF data.

The table reports the maximum likelihood estimates of the model parameters, and the implied Gini coefficient and distribution of wealth across top percentiles when the information used in the estimation contains a cross-section on individual wealth and a cross-section of employment status of the household's head. It provides bootstrapped standard errors and confidence intervals based on 128 bootstrap repetitions.

	ML estimates	**Std. Error**	**95% CI**
		$\delta = 0.1038$	
γ	0.6040	0.0582	[0.4899, 0.7180]
ρ	0.0012	0.0001	[0.0010, 0,0014]
α	0.6509	0.0006	[0.6497, 0.6522]

		% wealth in top		
	Gini Coefficient	1%	5%	20%
Data	0.8262	34.47	61.94	85.61
$\delta = 0.1038$	0.3945	5.85	18.37	45.45

8 Conclusions

Heterogeneous agent models constitute a powerful framework in macroeconomics not just for the study of inequality and the distribution of wealth but also for the understanding of macroeconomic aggregates. However, there is little agreement on how these models should be taken to the data. To date, calibration is the standard approach used by researchers to map observations into parameter values. Despite being very illustrative for the study of a model's implications, the use of econometric methods provide some important advantages by allowing: (i) to impose on the data the restrictions arising from the economic theory associated with a particular model; (ii) to assess the uncertainty surrounding the parameter values which ultimately provides a framework for hypothesis testing, (iii) for the use of standard tools of models selection and evaluation.

In this paper we introduce a simple full information likelihood approach to estimate the structural parameters of heterogeneous agent models using the information content in the cross-sectional distribution of wealth. Following the work of Bayer and Wälde (2011, 2013) and Achdou et al. (2014), the feasibility of our approach is dictated, in general, by the use of continuous-time methods, and in particular by the Fokker-Planck equations that allow us to approximate the stationary probability density function of wealth which can be used to build the model's likelihood function.

We also study the identification power of our maximum likelihood estimator based on data representing a large cross-section of individual wealth. Given that the mapping between the deep parameters of the model and the estimator's objective function is highly nonlinear, and not available in closed form, we follow Canova and Sala (2009) to assess in an indirect way whether the model's parameter are identified both in the population and in finite samples.

Our results indicate that while the parameters associated to the household preferences in heterogeneous agent models of the Bewley-Hugget-Aiyagari type are well identified and can be accurately estimated, the parameters closely related to the supply side of the economy exhibit partial identification problems. In other words, these parameters cannot be separately identified in the sense that increasing both parameters proportionally may leave the model's implied wealth distribution, and hence the likelihood function, unchanged. This partial identification problem is illustrated both at the population level, and in finite sample using Monte Carlo simulations. Our experiments suggest that the presence of partial identification issues lead to non-negligible biases. In particular, we find that the estimates of the capital share in production and the depreciation rate of capital exhibit a substantial positive bias that is far from negligible in small samples.

To overcome the partial identification problem between the capital share and the depreciation rate we propose and investigate the effects of following a strategy in which these parameters are calibrated, while the remaining ones are estimated. We conclude that a strategy in which only one of the two parameters is calibrated improves the finite sample properties of other one without affecting the identification, neither the estimation accuracy, of the preference parameters. This holds true even in the case where the underlying parameter is miscalibrated. While this strategy breaks down the inferential problems attached to partial identification issues, it implies that the supply side parameter being freely estimated will exhibit a similar bias to that introduced by the miscalibrated parameter. We also conclude that calibrating both parameters at the same time has pervasive consequences for the estimation of the preference parameters, and therefore such a strategy is not recommended when estimating the model on a cross-section of wealth.

We finally provide a small empirical illustration of our proposed framework by estimating the parameters of a Bewley-Aiyagari-Hugget model using the wealth data reported in the 2013 Survey of Consumer Finances. The estimates obtained provide supporting evidence that confirm our results on the identification power of the maximum likelihood estimator. As expected, the estimated model predicts reduced levels of wealth concentration relative to those observed in the U.S. economy. This results are shown to be robust to the inclusion of data on employment status as a proxy for income.

Our results are encouraging and suggest an important role for likelihood-based inference in heterogeneous agent models. With the increased availability of micro data on household characteristics and financial information, we expect that future research can consider more sophisticated models, like those in studied Krusell and Smith (1998), Cagetti and Nardi (2006), Angeletos and Calvet (2006), Angeletos (2007) and Benhabib et al. (2011), and more realistic income processes like the ones in Achdou et al. (2014) and Gabaix et al. (2016). This will allow to extend the information set used in the estimation process, potentially increase the identification power of the model structural parameters, and eventually provide a better fit of the wealth distribution.

A Hamilton-Jacobi-Bellman equations

Define the optimal value function:

$$V(a_0, e_0; w, r) = \max_{\{c_t\}_{t=0}^{\infty}} U_0 \quad s.t. \quad (2), (3)$$

in which the general equilibrium factor rewards r and w are taken as parametric.

Following the principle of optimality, the household's problem can be characterized by the Hamilton-Jacobi-Bellman equation:

$$\rho V(a_t, e_t; r, w) = \max_{c_t \in \mathbb{R}^+} \left\{ u(c_t) + \frac{1}{\mathrm{d}t} \mathbb{E}_t \mathrm{d}V(a_t, e_t; r, w) \right\}$$

for any $t \in [0, \infty)$.

Applying the change of variable formula (see Sennewald and Wälde, 2006) the continuation value is given by:

$$\mathrm{d}V(a_t, e_t; r, w) = V_a(a_t, e_t)\,\mathrm{d}a_t + (V(a_t, e_l) - V(a_t, e_h))\mathrm{d}q_{1,t} + (V(a_t, e_h) - V(a_t, e_l))\mathrm{d}q_{2,t}$$

where $V_a(a_t, e_t)$ denotes the partial derivative of the value function with respect to wealth.

Using Equation (2) together with the martingale difference properties of the stochastic integrals under Poisson uncertainty we have that for $s \leq t$:

$$\begin{aligned} \mathbb{E}_s \Big[\int_s^t (V(a_t, e_l) - V(a_t, e_h))\mathrm{d}q_{1,t} - \int_s^t (V(a_t, e_l) - V(a_t, e_h))\phi_1(e_t)\mathrm{d}t \Big] &= 0 \\ \mathbb{E}_s \Big[\int_s^t (V(a_t, e_h) - V(a_t, e_l))\mathrm{d}q_{2,t} - \int_s^t (V(a_t, e_h) - V(a_t, e_l))\phi_2(e_t)\mathrm{d}t \Big] &= 0. \end{aligned}$$

Then, the Hamilton-Jacobi-Bellman equation can be written as:

$$\begin{aligned} \rho V(a_t, e_t; r, w) = \max_{c_t \in \mathbb{R}^+} \big\{ & u(c_t) + V_a(a_t, e_t; r, w)(ra_t + we_t - c_t) \\ & + (V(a_t, e_l; r, w) - V(a_t, e_h; r, w))\phi_1(e_t) \\ & + (V(a_t, e_h; r, w) - V(a_t, e_l; r, w))\phi_2(e_t) \big\}. \end{aligned}$$

The first-order condition for an interior solution reads:

$$u'(c_t) = V_a(a_t, e_t; r, w), \tag{23}$$

for any $t \in [0, \infty)$, making optimal consumption $c_t^\star = c(a_t, e_t)$ a function only of the states and independent of calendar time, t.

Due to the state dependence of the arrival rates in the endowments of efficiency units, only one Poisson process will be active for each of the values of the discrete state variable, e_t. Using the first order condition we obtain a bivariate system of maximized HJB equations:

$$\begin{aligned} \rho V(a_t, e_h; r, w) &= u(c_t^\star) + V_a(a_t, e_h; r, w)(ra_t + we_h - c_t^\star) + (V(a_t, e_l; r, w) - V(a_t, e_h; r, w))\phi_{hl}, \\ \rho V(a_t, e_l; r, w) &= u(c_t^\star) + V_a(a_t, e_l; r, w)(ra_t + we_l - c_t^\star) + (V(a_t, e_h; r, w) - V(a_t, e_l; r, w))\phi_{lh}. \end{aligned}$$

B Fokker-Planck equations

Assume there exists a function f whose arguments are the stochastic processes a and e, and define the household's optimal savings function as $s(a_t, e_t) = ra_t + we_t - c(a_t, e_t)$. Using the change of variable formula, the evolution of f is given by:

$$\mathrm{d}f(a_t, e_t) = f_a(a_t, e_t)\, s(a_t, e_t)\,\mathrm{d}t + (f(a_t, e_l) - f(a_t, e_h))\,\mathrm{d}q_{1,t} + (f(a_t, e_h) - f(a_t, e_l))\,\mathrm{d}q_{2,t}.$$

Due to the state dependence of the arrival rates only one Poisson process will be active. Applying the expectations operator conditional on the information available at instant t and dividing by $\mathrm{d}t$ we obtain the infinitesimal generator of $f(a_t, e_t)$, denoted by $\mathcal{A}f(a_t, e_t) \equiv \frac{\mathbb{E}_t \mathrm{d}f(a_t, e_t)}{\mathrm{d}t}$:

$$\frac{\mathbb{E}_t \mathrm{d}f(a_t, e_t)}{\mathrm{d}t} = f_a(a_t, e_t)\, s(a_t, e_t) + (f(a_t, e_l) - f(a_t, e_h))\,\phi_{hl} + (f(a_t, e_h) - f(a_t, e_l))\,\phi_{lh}. \quad (24)$$

By means of Dynkin's formula, the expected value of the function $f(\cdot)$ at a point in time t is given by the expected value of the function at instant $s < t$ plus the sum of the expected future changes up to t:

$$\mathbb{E}f(a_t, e_t) = \mathbb{E}f(a_s, e_s) + \int_s^t \mathbb{E}(\mathcal{A}f(a_\tau, e_\tau))\,\mathrm{d}\tau. \quad (25)$$

Differentiating Equation (25) with respect to time:

$$\begin{aligned}
\frac{\partial}{\partial t}\mathbb{E}f(a_t, e_t) &= \frac{\partial}{\partial t}\left\{\mathbb{E}f(a_s, e_s) + \int_s^t \mathbb{E}(\mathcal{A}f(a_\tau, e_\tau))\,\mathrm{d}\tau\right\} \\
&= \frac{\partial}{\partial t}\left\{\mathbb{E}f(a_s, e_s) + \int_s^t \mathbb{E}\left(\frac{\mathbb{E}_\tau \mathrm{d}f(a_\tau, e_\tau)}{\mathrm{d}\tau}\right)\mathrm{d}\tau\right\} \\
&= \frac{\partial}{\partial t}\left\{\mathbb{E}f(a_s, e_s) + \int_s^t \mathbb{E}\mathrm{d}f(a_\tau, e_\tau)\right\} \\
&= \mathbb{E}(\mathcal{A}f(a_t, e_t)) \\
&= \sum_{e_t \in \{e_h, e_l\}} \int_{\underline{a}}^{\infty} \mathcal{A}f(a_t, e_t)\, g(a_t, e_t, t)\,\mathrm{d}a_t
\end{aligned}$$

that is:

$$\frac{\partial}{\partial t}\mathbb{E}f(a_t, e_t) = \underbrace{\int_{-\infty}^{\infty} \mathcal{A}f(a_t, e_h)\, g(a_t, e_h, t)\,\mathrm{d}a_t}_{\omega_{e_h}} + \underbrace{\int_{-\infty}^{\infty} \mathcal{A}f(a_t, e_l)\, g(a_t, e_l, t)\,\mathrm{d}a_t}_{\omega_{e_l}} \quad (26)$$

where $g(a_t, e_t, t)$ is the joint density function of wealth and endowment of efficiency units at instant t.

For illustration consider the case of $e_t = e_h$, i.e., ω_{e_h}. Using the definition of the infinitesimal operator together with Equation (24) we note that:

$$\mathcal{A}f(a_t, e_h) = f_a(a_t, e_h)\, s(a_t, e_h) + (f(a_t, e_l) - f(a_t, e_h))\, \phi_{hl}.$$

Hence,

$$\begin{aligned}\omega_{e_h} &= \int_{\underline{a}}^{\infty} \Big[f_a(a_t, e_h)\, s(a_t, e_h) + (f(a_t, e_l) - f(a_t, e_h))\, \phi_{hl}\Big] g(a_t, e_h, t)\, \mathrm{d}a_t \\ &= \int_{\underline{a}}^{\infty} f_a(a_t, e_h)\, s(a_t, e_h)\, g(a_t, e_h, t)\, \mathrm{d}a_t + \int_{\underline{a}}^{\infty} (f(a_t, e_l) - f(a_t, e_h))\, \phi_{hl} g(a_t, e_h, t)\, \mathrm{d}a_t.\end{aligned}$$

Using integration by part for the term associated with f_a:

$$\int_{\underline{a}}^{\infty} f_a(a_t, e_h)\, s(a_t, e_h)\, g(a_t, e_h, t)\, \mathrm{d}a_t = -\int_{\underline{a}}^{\infty} f(a_t, e_h)\, \frac{\partial}{\partial a_t}\left[s(a_t, e_h)\, g(a_t, e_h, t)\right] \mathrm{d}a_t$$

where:

$$\frac{\partial}{\partial a_t}\left[s(a_t, e_h)\, g(a_t, e_h, t)\right] = \left(r_t - \frac{\partial}{\partial a_t} c(a_t, e_h)\right) g(a_t, e_h, t) + s(a_t, e_h)\, \frac{\partial}{\partial a_t} g(a_t, e_h, t).$$

Hence,

$$\begin{aligned}\omega_{e_h} = \int_{\underline{a}}^{\infty} f(a_t, e_h) \left[-\left(r_t - \frac{\partial}{\partial a_t} c(a_t, e_h)\right) g(a_t, e_h, t) - s(a_t, e_h)\, \frac{\partial}{\partial a_t} g(a_t, e_h, t)\right] \mathrm{d}a_t \\ + \int_{\underline{a}}^{\infty} \Big[(f(a_t, e_l) - f(a_t, e_h))\, \phi_{hl}\Big] g(a_t, e_h, t)\, \mathrm{d}a_t\end{aligned}$$

and

$$\begin{aligned}\omega_{e_l} = \int_{\underline{a}}^{\infty} f(a_t, e_l) \left[-\left(r_t - \frac{\partial}{\partial a_t} c(a_t, e_l)\right) g(a_t, e_l, t) - s(a_t, e_l)\, \frac{\partial}{\partial a_t} g(a_t, e_2, t)\right] \mathrm{d}a_t \\ + \int_{\underline{a}}^{\infty} \Big[(f(a_t, e_h) - f(a_t, e_l))\, \phi_{lh}\Big] g(a_t, e_l, t)\, \mathrm{d}a_t.\end{aligned}$$

Note that the expected value of f can be written as:

$$\mathbb{E}f(a_t, e_t) = \int_{\underline{a}}^{\infty} f(a_t, e_h)\, g(a_t, e_h, t)\, \mathrm{d}a_t + \int_{\underline{a}}^{\infty} f(a_t, e_l)\, g(a_t, e_l, t)\, \mathrm{d}a_t$$

and therefore:

$$\frac{\partial}{\partial t}\mathbb{E}f\left(a_t,e_t\right) = \int\limits_{\underline{a}}^{\infty} f\left(a_t,e_h\right)\frac{\partial}{\partial t}g\left(a_t,e_h,t\right)\mathrm{d}a_t + \int\limits_{\underline{a}}^{\infty} f\left(a_t,e_l\right)\frac{\partial}{\partial t}g\left(a_t,e_l,t\right)\mathrm{d}a_t. \tag{27}$$

Finally we equate Equations (26) and (27) and collect terms to obtain:

$$\int\limits_{\underline{a}}^{\infty} f\left(a_t,e_h\right)\varphi_{e_h}\mathrm{d}a_t + \int\limits_{\underline{a}}^{\infty} f\left(a_t,e_l\right)\varphi_{e_l}\mathrm{d}a_t = 0 \tag{28}$$

where:

$$\varphi_{e_h} = -\left(r_t - \frac{\partial}{\partial a_t}c\left(a_t,e_h\right) + \phi_{hl}\right)g\left(a_t,e_h,t\right) - s\left(a_t,e_h\right)\frac{\partial}{\partial a_t}g\left(a_t,e_h,t\right) + \phi_{lh}g\left(a_t,e_l,t\right) - \frac{\partial}{\partial t}g\left(a_t,e_h,t\right)$$

and

$$\varphi_{e_l} = -\left(r_t - \frac{\partial}{\partial a_t}c\left(a_t,e_l\right) + \phi_{lh}\right)g\left(a_t,e_l,t\right) - s\left(a_t,e_l\right)\frac{\partial}{\partial a_t}g\left(a_t,e_l,t\right) + \phi_{hl}g\left(a_t,e_h,t\right) - \frac{\partial}{\partial t}g\left(a_t,e_l,t\right).$$

The Fokker-Planck equations that define these subdensities are obtained by setting:

$$\varphi_{e_l} = \varphi_{e_h} = 0$$

since that is that only way the integral equation (28) can be satisfied for all possible functions f. A formal proof can be found in Bayer and Wälde (2013). This results in a system of two non-autonomous quasi-linear partial differential equations in two unknowns $g\left(a_t,e_h,t\right)$, $g\left(a_t,e_l,t\right)$:

$$\frac{\partial}{\partial t}g\left(a_t,e_h,t\right) + s\left(a_t,e_h\right)\frac{\partial}{\partial a_t}g\left(a_t,e_h,t\right) = -\left(r_t - \frac{\partial}{\partial a_t}c\left(a_t,e_h\right) + \phi_{hl}\right)g\left(a_t,e_h,t\right) + \phi_{lh}g\left(a_t,e_l,t\right)$$

$$\frac{\partial}{\partial t}g\left(a_t,e_l,t\right) + s\left(a_t,e_l\right)\frac{\partial}{\partial a_t}g\left(a_t,e_l,t\right) = -\left(r_t - \frac{\partial}{\partial a_t}c\left(a_t,e_l\right) + \phi_{lh}\right)g\left(a_t,e_l,t\right) + \phi_{hl}g\left(a_t,e_h,t\right).$$

The stationary subdensities correspond to the case where the time derivatives $\partial g(a_t,e_t,t)/\partial t$ are zero for all $e_t \in \mathcal{E}$, which transforms the previous system of equations into one of ordinary differential equations as described by Equations (17) and (18).

Given the stationary subdensity function, the stationary probability "subdistributions" can be computed as:

$$G\left(a_t, e_t\right) = \int\limits_{\underline{a}}^{a} g\left(x_t, e_t\right) \mathrm{d}x_t \tag{29}$$

where $G\left(a_t, e_t\right)$ denotes the probability that an individual with endowment of efficiency equal to $e_t \in \mathcal{E}$ has a wealth level of at most a. When $a \to \infty$, Equation (13) implies that $\lim_{a_t \to \infty} G\left(a_t, e_t\right) = p\left(e_t\right)$. Similar to Equation (14), the (unconditional) stationary probability distribution of wealth can be computed as:

$$G\left(a_t\right) = G\left(a_t, e_h\right) + G\left(a_t, e_l\right) \tag{30}$$

which can be then used to compute the Gini coefficient in the economy:

$$\mathcal{G} = \frac{1}{\mu} \int\limits_{\underline{a}}^{\infty} G\left(a_t\right) \left(1 - G\left(a_t\right)\right) \mathrm{d}a_t \tag{31}$$

where we have defined $\mu = \mathbb{E}\left(a_t\right)$.

C Transition probabilities for the endowment of efficiency units

This appendix shows how to derive the limiting (stationary) probability distribution of the endowment of efficiency units defined in Equations (15) and (16) from the arrival rates of the stochastic process defined by Equation (3).

For illustration purposes consider an individual who is in state e_l at time s. Let $p(e_l, e_h, t) \equiv \mathbb{P}(e_t = e_h \mid e_s = e_l)$ for $s \leq t$ denote the probability that the individual jumps from state e_l at time s to state e_h at time t, and ϕ_{hl} and ϕ_{lh} the instantaneous transition rates at which the stochastic process jumps to state e_l from state e_h, and to state e_h from state e_l, respectively. Then the transition probabilities at time t can be computed from the solution to the following system of Backward Kolmogorov equations (see Ross, 2009):

$$\begin{aligned}\dot{p}(e_h, e_h, t) &= \phi_{hl}\left[p(e_l, e_h, t) - p(e_h, e_h, t)\right], \\ \dot{p}(e_l, e_h, t) &= \phi_{lh}\left[p(e_h, e_h, t) - p(e_l, e_h, t)\right]\end{aligned}$$

where $\dot{p}(e_i, e_j, t) = \lim_{s\to 0} \frac{1}{s}\left[p(e_i, e_j, t+s) - p(e_i, e_j, t)\right]$ for all $i, j \in \mathcal{E}$, and $p(e_h, e_h, s) = 1$ and $p(e_l, e_h, s) = 0$ are initial conditions. The solution to this system of ordinary differential equations is given by:

$$p(e_h, e_h, t) = \frac{\phi_{lh}}{\phi_{hl} + \phi_{lh}} + \frac{\phi_{hl}}{\phi_{hl} + \phi_{lh}} e^{-(\phi_{hl}+\phi_{lh})(t-s)} \tag{32}$$

$$p(e_l, e_h, t) = \frac{\phi_{lh}}{\phi_{hl} + \phi_{lh}} - \frac{\phi_{lh}}{\phi_{hl} + \phi_{lh}} e^{-(\phi_{hl}+\phi_{lh})(t-s)}. \tag{33}$$

Now let $p(e_h, s)$ denote the unconditional probability of being in state e_h at time s. The unconditional probability of being in the same state at time $t > s$ can be computed according to:

$$p(e_h, t) = p(e_h, s)\, p(e_h, e_h, t) + (1 - p(e_h, s))\, p(e_l, e_h, t). \tag{34}$$

In the limit as $t \to \infty$ the unconditional probability of having an endowment of high efficiency is given by:

$$\lim_{t\to\infty} p(e_h, t) = p(e_h) = \frac{\phi_{lh}}{\phi_{hl} + \phi_{lh}}. \tag{35}$$

A similar procedure can be used to show that the stationary and unconditional probability of having an endowment of low efficiency is:

$$\lim_{t\to\infty} p(e_l, t) = p(e_l) = \frac{\phi_{hl}}{\phi_{hl} + \phi_{lh}}. \tag{36}$$

The system of equations formed by (32) and (33) together with an appropriate choice of $(t - s)$ can be used to back out the instantaneous transition rates of the Poisson processes, ϕ_{hl} and ϕ_{lh} from any probability transition matrix. Given the annual frequency used in the calibration of the model of Section 2, we set $(t - s) = 1$ (one year).

D Computation of the stationary equilibrium

The computation of the stationary density of wealth is done following the method proposed in Achdou et al. (2017) which consists of two main blocks. The first block computes the stationary general equilibrium at the macro level by using the following fixed point algorithm in the time-invariant aggregate capital stock:

Algorithm D.1 (Stationary General Equilibrium) *Make an initial guess for the interest rate, $r^{(0)}$, and then for $j = 0, 1, \ldots$:*

1. *Compute the optimal consumption functions $c^{(j)}(a, e_h)$ and $c^{(j)}(a, e_l)$ and the subdensities $g^{(j)}(a, e_h)$ and $g^{(j)}(a, e_l)$.*
2. *Compute capital demand K^d and capital supply K^s.*
3. *Update $r^{(j+1)}$ using a combination of bisection, secant, and inverse quadratic interpolation methods.*
4. *If $\left\|K^s - K^d\right\| < \epsilon$ stop, otherwise return to step 1.*

Algorithm D.1 does not require to update the aggregate labor supply L at each iteration $j = 0, 1, \ldots$ since in our prototype economy the labor supply is assumed to be exogenous as can be seen from Equation (12).

The second block approximates both the solution to the household's problem at the micro level and to the Fokker-Planck equations using the finite difference methods suggested in Candler (1999) and Achdou et al. (2017). These solutions, which are required in step 2 of Algorithm D.1 for every iteration $j = 0, 1, \ldots$, are computed in two independent stages. The first stages approximates the policy functions for consumption that solve the HJB equations (7) and (8), while the second stage approximates the subdensity functions of wealth that solve the Fokker-Planck equations (17) and (18).

Solving the Hamilton-Jacobi-Bellman equations.

Consider first the solution to the HJB equations. For each $e_t \in \mathcal{E}$, the finite difference method approximates the function $V(a_t, e_t)$ on an equally spaced grid for wealth with I discrete points, $a_i, i = 1, \ldots, I$, where $a_i \in \mathcal{A} = [a_{min}, a_{max}]$ and $a_{min} = \underline{a}$. The distance between points is denoted by Δa and we introduce the short-hand notation $V_{e,i} \equiv V(a_i, e)$. The derivative $V_a(a_i, e) \equiv V'_{e,i}$ is computed with either a forward or a backward difference approximation:

$$V'^F_{e,i} \approx \frac{V_{e,i+1} - V_{e,i}}{\Delta a} \tag{37}$$

$$V'^B_{e,i} \approx \frac{V_{e,i} - V_{e,i-1}}{\Delta a}. \tag{38}$$

Following Candler (1999), the choice of difference operator is based on an upwind differentiation scheme. The correct approximation is based on the direction of the continuous state variable. Thus, if the saving function, $s(a_i, e) \equiv s_{e,i} = ra_i + we - (u')^{-1}\left(V'_{e,i}\right)$, is positive we use a forward operator and if it is negative we use the backward operator. This gives rise to the following upwind operator:

$$V'_{e,i} = V'^{F}_{e,i}\mathbf{1}_{\{s^F_{e,i}>0\}} + V'^{B}_{e,i}\mathbf{1}_{\{s^B_{e,i}<0\}} + \bar{V}'_{e,i}\mathbf{1}_{\{s^F_{e,i}<0<s^B_{e,i}\}} \tag{39}$$

where $\mathbf{1}_{\{\cdot\}}$ denotes the indicator function and, $s^F_{e,i}$ and $s^B_{e,i}$ the saving functions computed with the forward and difference operators respectively. Following Achdou et al. (2017), the concavity of the value function in the wealth dimension motivates the last term in Equation (39) since there could be grid points $a_i \in \mathcal{A}$ for which $s^F_{e,i} < 0 < s^B_{e,i}$. In those cases, they suggest to set savings to be equal to zero which implies that the derivative of the value function is equal to $\bar{V}'_{e,i} = u'(ra_i + we)$.

The finite difference approximation to the HJB equations is then given by[11]:

$$\rho V_{e,i} = u(c_{e,i}) + V'_{e,i}\left[ra_i + ew - c_{e,i}\right] + \phi_{-ee}\left[V_{-e,i} - V_{e,i}\right]$$

for each $e \in \mathcal{E}$, where optimal consumption is given by:

$$c_{e,i} = (u')^{-1}\left(V'_{e,i}\right).$$

The upwind representation of the HJB equation reads:

$$\rho V_{e,i} = u(c_{e,i}) + \frac{V_{e,i+1} - V_{e,i}}{\Delta a}(s_{e,i})^{+} + \frac{V_{e,i} - V_{e,i-1}}{\Delta a}(s_{e,i})^{-} + \phi_{-ee}\left[V_{-e,i} - V_{e,i}\right] \tag{40}$$

where:

$$(s_{e,i})^{+} = \max\left\{ra_i + we - (u')^{-1}\left(V'^{F}_{e,i}\right), 0\right\} \quad \text{and} \quad (s_{e,i})^{-} = \min\left\{ra_i + we - (u')^{-1}\left(V'^{B}_{e,i}\right), 0\right\}$$

denote the positive and negative parts of savings, respectively.

Equation (40) defines a highly non linear system of equations in $V_{e,i}$ that can only be solved by iterative methods. We follow Candler (1999) and set up an iterative procedure based on the time-dependent HJB equation, $V^l_{e,i} \equiv V(a_i, e, t)$. Then, from an arbitrary initial condition we integrate forward in time until the solution is no longer a function of the initial condition, i.e. until it converges to the time-independent HJB, $V_{e,i}$. The time-updating is carried out by means of an implicit scheme in which the value function at the next time step, $V^{l+1}_{e,i}$, is implicitly defined by the equation:

$$\frac{V^{l+1}_{e,i} - V^l_{e,i}}{\Delta} + \rho V^{l+1}_{e,i} = u\left(c^l_{e,i}\right) + \frac{V^{l+1}_{e,i+1} - V^{l+1}_{e,i}}{\Delta a}\left(s^l_{e,i}\right)^{+} + \frac{V^{l+1}_{e,i} - V^{l+1}_{e,i-1}}{\Delta a}\left(s^l_{e,i}\right)^{-} + \phi_{-ee}\left[V^{l+1}_{-e,i} - V^{l+1}_{e,i}\right] \tag{41}$$

[11] The state-constraint boundary condition in Equation (9) is enforced at the lower bound of the state space, a_{min}, by imposing $V'^{B}_{e,1} = u'(ra_1 + we)$.

where Δ is the time step size, $c^l_{e,i} = (u')^{-1}\left[\left(V^l_{e,i}\right)'\right]$, and $\left(V^l_{e,i}\right)'$ is given by Equation (39).

Equation (41) constitutes a system of $2 \times I$ linear equations in $V^{l+1}_{e,i}$ with the following matrix representation:

$$\mathbf{A}^l \mathbf{V}^{l+1} = \mathbf{b}^l \tag{42}$$

where $\mathbf{V}^{l+1} = \left(V^{l+1}_{e_l,1}, \ldots, V^{l+1}_{e_l,I}, V^{l+1}_{e_h,1}, \ldots, V^{l+1}_{e_h,I}\right)'$, $\mathbf{b}^l$ is a vector with elements $b^l_{e,i} = u\left(c^l_{e,i}\right) + V^l_{e,i}/\Delta$ and $\mathbf{A}^l$ is the block matrix:

$$\mathbf{A}^l = \begin{bmatrix} \mathbf{A}_{e_l} & -\Phi_{hl} \\ -\Phi_{lh} & \mathbf{A}_{e_h} \end{bmatrix}$$

with $\Phi_{-ee} = -\phi_{-ee}\mathbf{I}_I$ and

$$\mathbf{A}_e = \begin{bmatrix} y_{e,1} & z_{e,1} & 0 & \cdots & 0 & 0 \\ x_{e,2} & y_{e,2} & z_{e,2} & \cdots & 0 & 0 \\ 0 & x_{e,3} & y_{e,3} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & y_{e,I-1} & z_{e,I-1} \\ 0 & 0 & 0 & \cdots & x_{e,I} & y_{e,I} \end{bmatrix}.$$

where

$$x_{e,i} = \frac{\left(s^l_{e,i}\right)^-}{\Delta a}$$

$$y_{e,i} = \frac{1}{\Delta} + \rho + \frac{\left(s^l_{e,i}\right)^+}{\Delta a} - \frac{\left(s^l_{e,i}\right)^-}{\Delta a} + \phi_{-ee}$$

$$z_{e,i} = -\frac{\left(s^l_{e,i}\right)^+}{\Delta a}.$$

and $e \in \mathcal{E}$. The iterative algorithm used to find the solution to the HJB equation can be summarized as follows:

Algorithm D.2 (Solution of the HJB equation) *Guess $V^0_{e,i}$ for each $e \in \mathcal{E}$ and $i = 1, \ldots, I$. Then for $l = 0, 1, 2, \ldots$:*

1. *Compute $\left(V^l_{e,i}\right)'$ using Equation (39).*

2. *Compute $c^l_{e,i} = (u')^{-1}\left(V'_{e,i}\right)$.*

3. *Find $V^{l+1}_{e,i}$ by solving the system of equations defined in (42).*

4. *If $\left\|V^{l+1}_{e,i} - V^l_{e,i}\right\| < \epsilon$ stop. Otherwise, go to step 1.*

Solving the Fokker-Planck equations.

Once the optimal consumption has been computed from Algorithm D.2, we proceed to approximate the solution to the associated Fokker-Planck equations (17) and (18). As before, we use a finite difference method and apply it to:

$$0 = -\frac{\partial}{\partial a_t}\left[s\left(a_t, e_l\right) g\left(a_t, e_l\right)\right] - \phi_{hl} g\left(a_t, e_l\right) - \phi_{lh} g\left(a_t, e_h\right), \tag{43}$$

$$0 = -\frac{\partial}{\partial a_t}\left[s\left(a_t, e_h\right) g\left(a_t, e_h\right)\right] - \phi_{lh} g\left(a_t, e_h\right) - \phi_{hl} g\left(a_t, e_l\right) \tag{44}$$

which corresponds, as shown in Appendix B, to an alternative representation of Equations (17) and (18). We further need to restrict the solution to satisfy the integrability condition:

$$1 = \sum_{e_t \in \{e_l, e_h\}} \int_{-\infty}^{\infty} g\left(a_t, e_t\right) \mathrm{d}a. \tag{45}$$

The system of equations (43)-(45) is discretized as follows:

$$0 = -\left[s_{e,i} g_{e,i}\right]' - \phi_{-ee} g_{e,i} - \phi_{e,-e} g_{-e,i} \tag{46}$$

$$1 = \sum_{e_t \in \{e_l, e_h\}} \sum_{i=1}^{I} g_{e,i} \Delta a. \tag{47}$$

where $g_{e,i} \equiv g\left(a_i, e\right)$. To approximate the derivative $\left[s_{e,i} g_{e,i}\right]'$ we use the upwind differentiation scheme:

$$\left[s_{e,i} g_{e,i}\right]' = \frac{\left(s_{e,i}\right)^+ g_{e,i} - \left(s_{e,i-1}\right)^+ g_{e,i-1}}{\Delta a} + \frac{\left(s_{e,i+1}\right)^- g_{e,i+1} - \left(s_{e,i}\right)^- g_{e,i}}{\Delta a},$$

where $s_{e,i} = r a_i + w e - \left(u'\right)^{-1}\left(V'_{e,i}\right)$ is the optimal savings function obtained from the solution to the HJB equation. Equation (46) defines a system of $2 \times I$ linear equations in $g_{e,i}$ with matrix representation:

$$\mathbf{B}\mathbf{g} = \mathbf{0} \tag{48}$$

where $\mathbf{g} = \left(g_{e_l,1}, \ldots, g_{e_l,I}, g_{e_h,1}, \ldots, g_{e_h,I}\right)'$. The matrix $\mathbf{B}$ is defined as $\mathbf{B} = \tilde{\mathbf{A}}^\top$, where $\tilde{\mathbf{A}} = -\mathbf{A} + \left(\rho + \frac{1}{\Delta}\right)\mathbf{I}$. The matrix $\tilde{\mathbf{A}}$ captures the evolution of the continuous-time stochastic processes $\{a_t, e_t\}_{t=0}^{\infty}$. To impose the integrability condition in Equation (45) we follow Achdou et al. (2017) and fix $g_{e,i} = 0.1$ for an arbitrary i. Then solve the system of equations in (48) for some $\tilde{\mathbf{g}}$, and proceed to re-normalize $g_{e,i} = \tilde{g}_{e,i} / \left(\sum_{e,i} \tilde{g}_{e,i} \Delta a\right)$.

E Data, estimation settings and robustness check

The cross-section of individual wealth used in Section 7 for the estimation of the Bewley-Aiyagari-Hugget model is obtained from the 2013 Survey of Consumer Finances (SCF). In particular, our wealth data matches the net-worth reported in the SCF database. We re-sample the data based on the survey weights[12]. For the estimation we only include positive net-worth data (in thousands) to be consistent with the model's non-negative borrowing constraint.

Since the data on net-worth used in the estimation is highly skewed and contains a few very large outliers the use of an uniform grid on the wealth lattice for the approximation of the stationary probability density function of wealth would result inappropriate. Therefore, we use instead a non-uniform grid (a log-grid) and modify the solution step accordingly. Details of finite differencing with non-equally spaced grids can be found in Achdou et al. (2017).

The model's policy functions and wealth subdensity functions are approximated on a grid with $I = 2000$ non-equally spaced points. The maximization of the log-likelihood function is done by randomly selecting 200 starting values, $\boldsymbol{\theta}^{(0)}$, in order to prevent hitting and getting stuck in local maximum. The standard errors and confidence intervals are computed by means of the parametric bootstrap. Given a set of estimated parameters, we simulate 128 bootstrap data samples of wealth and re-estimate the model for each of the bootstrapped samples. The reported standard errors for the estimated coefficients correspond to the standard errors of the bootstrap estimates, and the confidence intervals are constructed using the Normal approximation.

As a robustness check we repeat our estimation exercise by including data on employment status of the household's head as a proxy for income. This allows us to exploit the model's implied joint density of individual wealth and efficiency endowment. Given our extended information set, the log-likelihood function in Equation (20) becomes:

$$\mathcal{L}_N\left(\boldsymbol{\theta}|\boldsymbol{a},\boldsymbol{e}\right)=\sum_{n=1}^{N}\left\{\mathbf{1}_{\{e_n=e_l\}}\log g\left(a_n,e_n|\boldsymbol{\theta}\right)+\mathbf{1}_{\{e_n=e_h\}}\log g\left(a_n,e_n|\boldsymbol{\theta}\right)\right\}$$

and the maximum likelihood estimator is given by:

$$\hat{\boldsymbol{\theta}}_N=\underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\max}\ \mathcal{L}_N\left(\boldsymbol{\theta}\mid\boldsymbol{a},\boldsymbol{e}\right)$$

where $\boldsymbol{e}=[e_1,\ldots,e_n]$ is a sample of N i.i.d observations on household's head employment status, $\mathbf{1}_{\{\cdot\}}$ is an indicator function for the type of efficiency endowment. The computation of the model's policy functions and the maximization of the log-likelihood function are carried out in the same way as for the case where the information set only contains a cross-section on individual wealth.

[12] We have several trials of estimates based on different re-sampled data, the overall estimation results are very similar.

References

Abbott, B., G. Gallipoli, C. Meghir, and G. L. Violante (2016): "Education Policy and Intergenerational Transfers in Equilibrium," Working Paper 18782, National Bureau of Economic Research.

Achdou, Y., F. J. Buera, J.-M. Lasry, P.-L. Lions, and B. Moll (2014): "PDE Models in Macroeconomics," *Proceedings of Royal Society A*, 1–15.

Achdou, Y., J. Han, J.-M. Lasry, P.-L. Lions, and B. Moll (2017): "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," Working Paper 23732, National Bureau of Economic Research.

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

Angeletos, G.-M. (2007): "Uninsured Idiosyncratic Investment Risk and Aggregate Saving," *Review of Economic Dynamics*, 10, 1–30.

Angeletos, G.-M. and L.-E. Calvet (2006): "Idiosyncratic Production Risk, Growth and the Business Cycle," *Journal of Monetary Economics*, 53, 1095–1115.

Bayer, C. and K. Wälde (2010a): "Matching and Saving in Continuous Time: Proofs," CESifo Working Paper Series 3026-A, CESifo Group Munich.

——— (2010b): "Matching and Saving in Continuous Time: Theory," CESifo Working Paper Series 3026, CESifo Group Munich.

——— (2011): "Describing the Dynamics of Distributions in Search and Matching Models by Fokker-Planck Equations," Unpublished.

——— (2013): "The Dynamics of Distributions in Continuous-Time Stochastic Models," Unpublished.

Benhabib, J. and A. Bisin (2017): "Skewed Wealth Distributions: Theory and Empirics," Working Paper 21924, National Bureau of Economic Research.

Benhabib, J., A. Bisin, and M. Luo (2015): "Wealth Distribution and Social Mobility in the US: A Quantitative Approach," Working Paper 21721, National Bureau of Economic Research.

Benhabib, J., A. Bisin, and S. Zhu (2011): "The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents," *Econometrica*, 79, 123–157.

Bewley, T. (Undated): "Interest Bearing Money and the Equilibrium Stock of Capital," Manuscript.

Cagetti, M. and M. De Nardi (2008): "Wealth Inequality: Data And Models," *Macroeconomic Dynamics*, 12, 285–313.

Cagetti, M. and M. D. Nardi (2006): "Entrepreneurship, Frictions, and Wealth," *Journal of Political Economy*, 114, 835–870.

Candler, G. (1999): *Finite-difference methods for continuous-time dynamic programming,*, in R. Marimon and A. Scott (Eds.): Computational Methods for the Study of Dynamic Economies, Oxford University Press.

Canova, F. and L. Sala (2009): "Back to Square One: Identification Issues in DSGE Models," *Journal of Monetary Economics*, 56, 431–449.

Castañeda, A., J. Díaz-Giménez, and J.-V. Ríos-Rull (2003): "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 111, 818–857.

Cooley, T. and E. Prescott (1995): *Economic growth and business cycles*, in T. Cooley (Ed.): Frontiers of Business Cycle Research, Princeton University Press.

Díaz-Giménez, J., R. Kirkby, and J. Pijoan-Mas (2014): "Flat Tax Reforms: Investment Expensing and Progressivity," Unpublished.

Gabaix, X., J.-M. Lasry, P.-L. Lions, and B. Moll (2016): "The Dynamics of Inequality," *Econometrica*, 84, 2071–2111.

Gomme, P. and P. Rupert (2007): "Theory, Measurement and Calibration of Macroeconomic Models," *Journal of Monetary Economics*, 54, 460–497.

Heathcote, J., K. Storesletten, and G. L. Violante (2009): "Quantitative Macroeconomics with Heterogeneous Households," *Annual Review of Economics*, 1, 319–354.

Heer, B. and M. Trede (2003): "Efficiency and Distribution Effects of a Revenue-Neutral Income Tax Reform," *Journal of Macroeconomics*, 25, 87–107.

Huggett, M. (1993): "The Risk-free Rate in Heterogeneous-Agent Incomplete-Insurance Economies," *Journal of Economic Dynamics and Control*, 17, 953–969.

Imrohoroğlu, A. (1989): "Cost of Business Cycles with Indivisibilities and Liquidity Constraints," *Journal of Political Economy*, 97, 1364–1383.

Iskrev, N. (2010): "Local Identification in DSGE Models," *Journal of Monetary Economics*, 57, 189–202.

Kaplan, G., B. Moll, and G. L. Violante (2016): "Monetary Policy According to HANK," NBER Working Papers 21897, National Bureau of Economic Research, Inc.

Komunjer, I. and S. Ng (2011): "Dynamic Identification of Dynamic Stochastic General Equilibrium Models," *Econometrica*, 79, 1995–2032.

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

Kydland, F. E. and E. C. Prescott (1982): "Time to Build and Aggregate Fluctuations," *Econometrica*, 50, 1345–1370.

Luo, M. and S. Mongey (2017): "Student Debt and Job Choice: Wages vs. Job Satisfaction," Unpublished.

Mongey, S. and J. Williams (2017): "Firm Dispersion and Business Cycles: Estimating Aggregate Shocks Using Panel Data," Unpublished.

Newey, W. K. and D. McFadden (1986): "Large Sample Estimation and Hypothesis Testing," in *Handbook of Econometrics*, ed. by R. F. Engle and D. McFadden, Elsevier, vol. 4 of *Handbook of Econometrics*, chap. 36, 2111–2245.

Ozkan, S., K. Mitman, F. Karahan, and A. Hedlund (2016): "Monetary Policy, Heterogeneity and the Housing Channel," 2016 Meeting Papers 663, Society for Economic Dynamics.

Prescott, E. C. (1986): "Theory Ahead of Business Cycle Measurement," *FRB MN Quarterly Review*.

Quadrini, V. and J.-V. Ríos-Rull (1997): "Understanding the U.S. Distribution of Wealth," *Federal Reserve Bank of Minneapolis Quarterly Review*, 21, 22–36.

Reiter, M. (2009): "Solving Heterogeneous-Agent Models by Projection and Perturbation," *Journal of Economic Dynamics and Control*, 33, 649–665.

Ríos-Rull, J. V. (1995): *Models with heterogeneous agents*, in T. Cooley (Ed.): Frontiers of Business Cycle Research, Princeton University Press.

——— (2001): *Computation of equilibria in heterogenous agent models*, in R. Marimon and A. Scott (Eds.): Computational Methods for the Study of Dynamic Economies, Oxford University. Press, 2nd ed.

Ríos-Rull, J.-V., F. Schorfheide, C. Fuentes-Albero, M. Kryshko, and R. Santaeulalia-Llopis (2012): "Methods versus Substance: Measuring the Effects of Technology Shocks," *Journal of Monetary Economics*, 59, 826–846.

Ross, S. M. (2009): *Introduction to Probability Models*, Academic Press, 10th ed.

Rothenberg, T. J. (1971): "Identification in Parametric Models," *Econometrica*, 39, pp. 577–591.

Sennewald, K. and K. Wälde (2006): "Itô's Lemma and the Bellman equation for Poisson processes: An applied view," *Journal of Economics*, 89, 1–36.

Shimer, R. (2005): "The Cyclical Behavior of Equilibrium Unemployment and Vacancies," *American Economic Review*, 95, 25–49.

Williams, J. (2017): "Bayesian Estimation of DSGE Models with Heterogeneous Agents," Unpublished.

Winberry, T. (2016): "A Toolbox for Solving and Estimating Heterogeneous Agent Macro Models," Unpublished.

Wong, A. (2016): "Population Aging and the Transmission of Monetary Policy to Consumption," 2016 Meeting Papers 716, Society for Economic Dynamics.