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)]Universität Hamburg

September 2018

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 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 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. We also provide an empirical illustration of our maximum likelihood framework using wealth data from the 2013 Survey of Consumer Finances.

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), the 10th NNH-UiO Workshop on Economic Dynamics (Oslo, 2015) and seminar participants at various institutions 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 of 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 been used for the study of the distributional implications of monetary and fiscal policies (see Kaplan et al., 2016; Ozkan et al., 2016; Wong, 2016; Holm, 2017).

Currently, the main workhorse 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.

The standard approach to study the quantitative properties of these models is based on the calibration of their structural parameters. Hence, the parameter values are either fixed to those existing in the literature for which there exists a wide consensus, or chosen in such a way that they minimize the distance between a subset of moments implied by the model and the same moments obtained from the data, or by a combination of both. Accordingly, calibration can be defined as a partial or limited information approach in the sense that it only makes use of a subset of the model cross-equation restrictions. Kydland and Prescott (1982) introduced calibration into macroeconomics with subsequent developments made by Prescott (1986), Cooley and Prescott (1995) and Gomme and Rupert (2007). Recent examples that combine both types of calibration approaches, conditional on estimated values for the exogenous income process, can be found in Benhabib et al. (2015), Abbott et al. (2016) and Luo and Mongey (2017) for models without aggregate uncertainty.

On the other hand, full information methods which rely on the entire probability distribution of the model have received less attention. The first contribution of this paper is to provide a simple likelihood framework as an alternative to calibration, which exploits the information content in the cross-sectional distribution of wealth, to estimate the structural parameters of heterogeneous agent models of the Bewley-Hugget-Aiyagari type. Our approach relies on the ability to compute the model's implied stationary probability density function which can be later used to build the likelihood function of the model. Likelihood-based methods impose on the data the full set of restrictions implied by a particular economic model allowing the econometrician: (i) to assess the uncertainty surrounding the parameter values which ultimately provides a framework for inference and hypothesis testing, and (ii) to use standard tools for model selection and evaluation.

In general, the computation of the probability density function of the individual state variables in heterogeneous agent models is not straightforward as they turn out to be complicated endogenous and nonlinear objects that usually have to be numerically approximated either by Monte Carlo simulation or functional approximation techniques[1]. 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 or Kolmogorov's Forward equations equations for the analysis of endogenous distributions in macroeconomics. 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 Hamilton-Jacobi-Bellman equation that describes the optimal consumption-saving decisions of economic agents, they form a system of coupled partial differential equations that can be numerically solved with high degree of accuracy and computational efficiency on the entire state-space of the model using the finite difference methods described in Candler (1999) and Achdou et al. (2017). It is the ability to compute the probability density function of wealth in a non-parametric way through the use of Fokker-Planck equations together with the increased efficiency in the computation of the model's equilibrium provided by finite difference methods that makes our method novel and feasible for estimation purposes.

The approach proposed in this paper differs from those recently applied in Winberry (2016), Mongey and Williams (2017) and Williams (2017). In particular, these authors employ Bayesian-likelihood methods to estimate the parameters that govern the dynamics of aggregate exogenous macroeconomic shocks, conditional on calibrated values for the preference parameters which ultimately depend on the cross-sectional stationary distribution of individual states. Therefore, their estimation approach does not make any use of the model's implied probability density function. Closer to our approach is the work by Challe et al. (2017) where a subset of the preference parameters is included in the estimation step and hence some knowledge on the cross-sectional probability density function is required. While interesting and challenging, under their stated assumptions, the quantitative exercise reduces the probability distribution to a mass point to facilitate the estimation, departing this way from the approach proposed here where the entire distribution is used in the estimation process.

A condition for the maximum likelihood estimator to deliver consistent estimates of the

[1] See Heer and Maussner (2009) for a textbook treatment of the different approaches.

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 assess the identifiability 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 with aggregate certainty 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 probability distribution of wealth that can be used for estimation and/or identification analysis.

In Section 3 we show how to compute the model's likelihood function using the time-invariant density of wealth. We also introduce the concept of identification within our maximum likelihood framework, summarize the different types of identification issues that could potentially arise in heterogeneous agent models, and investigate the identifiability of the model parameters from the sample's objective function. While our framework can be easily extended to include cross-sectional data on a number of additional variables (e.g., income or consumption), the simple structure of the model we adopt restricts their use and hence we focus our description and analysis to the case where only a cross-section of wealth is available to the econometrician.

Section 4 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. Our results suggest that while the household's preference parameters are well identified and exhibit negligible biases, parameters related to the supply side of the economy exhibit partial identification problems that compromise the accuracy with which they can be estimated. The identification problems of the latter markedly contaminate the estimates of the capital-output ratio and the savings rate, even in large samples.

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. In Section 5 we investigate the consequences of following such an strategy and find that, even in the case where some parameters are miscalibrated, it improves the finite sample properties of those parameters being estimated. Section 6 provides an empirical illustration of our proposed framework by estimating the parameters of a Bewley-Hugget-Aiyagari model for the U.S. using the wealth data from the 2013 Survey of Consumer Finances. Section 7 concludes.

2. A prototypical heterogeneous agent model

For our study we consider a prototypical heterogeneous agent model *á la* Bewley-Hugget-Aiyagari 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:

$$\mathrm{d}a_t = (ra_t + we_t - c_t)\mathrm{d}t, \tag{2}$$

where $a_t \in \mathcal{A} \subset \mathbb{R}$ 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, $e_t \in \mathcal{E}$, making their labor income stochastic (see Castañeda et al., 2003). Following Heer and Trede (2003), the efficiency levels could also be understood as productivity shocks. Finally, w denotes the wage rate per efficiency unit which is the same across households and determined in general equilibrium together with the interest rate. 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 (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 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:

$$\phi_1(e_t) = \begin{cases} \phi_{hl} & e_t = e_h \\ 0 & e_t = e_l \end{cases}, \text{ and } \quad \phi_2(e_t) = \begin{cases} 0 & e_t = e_h \\ \phi_{lh} & e_t = e_l. \end{cases}$$

Finally, 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. Hence, $\mathcal{A} = [\underline{a}, \infty)$. 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 $Y = K^{\alpha} L^{1-\alpha}$, with $\alpha \in (0, 1)$, 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 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 $(a_t, e_t) \in \mathcal{A} \times \mathcal{E}$ 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:

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

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

An interesting feature of our continuous-time setup as opposed to the discrete-time case, is that (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):

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

It can be shown that (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, $K = \sum_{e_t \in \{e_l, e_h\}} \int_{\underline{a}}^{\infty} a_t g\left(a_t, e_t\right) \mathrm{d}a_t$, and the steady state aggregate labor is obtained by aggregating their efficiency labor units, $L = \sum_{e_t \in \{e_l, e_h\}} \int_{\underline{a}}^{\infty} e_t g\left(a_t, e_t\right) \mathrm{d}a_t$, providing this way 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 $V\left(a_t, e_l\right)$ and $V\left(a_t, e_h\right)$, individual policy functions for consumption $c\left(a_t, e_l\right)$ and $c\left(a_t, e_h\right)$, a time-invariant density of the state variables $g\left(a_t, e_l\right)$ and $g\left(a_t, e_h\right)$, constant prices of labor and capital $\{w, r\}$, and a vector of constant aggregates $\{K, L, Y, C\}$ such that:*

1. *the consumption functions $c\left(a_t, e_l\right)$ and $c\left(a_t, e_h\right)$ satisfy (7) and (8), i.e. they solve the household's allocation problem,*
2. *factor prices satisfy the first order condition in (10), i.e. they solve the firm's problem,*
3. *markets clear, i.e. (4) holds.*
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 $(a_t, e_t) \in \mathcal{A} \times \mathcal{E}$.*

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:

$$\int g(a_t, e_t)\, \mathrm{d}a_t \quad = \quad p(e_t), \tag{11}$$

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:

$$g(a_t) = g(a_t, e_h) + g(a_t, e_l). \tag{12}$$

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 the accompanying web appendix):

$$\lim_{t\to\infty} p(e_h, t) \equiv p(e_h) = \frac{\phi_{lh}}{\phi_{hl} + \phi_{lh}}, \text{ and } \lim_{t\to\infty} p(e_l, t) \equiv p(e_l) = \frac{\phi_{hl}}{\phi_{hl} + \phi_{lh}}. \tag{13}$$

Let $s(a_t, e_t) = ra_t + we_t - c(a_t, e_t)$ denote the optimal savings function for an individual with an efficiency endowment equal to e_t. As shown in Appendix B, the subdensities in (12) correspond to the solution of the following non-autonomous quasi-linear system of differential equations known as (stationary) Fokker-Planck equations:

$$s(a_t, e_l) \frac{\partial}{\partial a_t} g(a_t, e_l) \quad = \quad -\left(r - \frac{\partial}{\partial a_t} c(a_t, e_l) + \phi_{lh}\right) g(a_t, e_l) + \phi_{hl} g(a_t, e_h), \tag{14}$$

$$s(a_t, e_h) \frac{\partial}{\partial a_t} g(a_t, e_h) \quad = \quad -\left(r - \frac{\partial}{\partial a_t} c(a_t, e_h) + \phi_{hl}\right) g(a_t, e_h) + \phi_{lh} g(a_t, e_l), \tag{15}$$

where the partial derivatives with respect to a_t describe the cross-sectional dimension of the density function. The system of equations formed by (14) and (15) 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 globally approximated on a discretized state space for $\mathcal{A}$. 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 steps. In the first step it approximates the solution to the system of equations (7) and (8) which represents the household's allocation problem at the micro level. Given the optimal consumption function from step one, the second step approximates the stationary subdensities that solve the system of ordinary differential equations in (14) and (15). 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 the accompanying web appendix.

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. 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.

Let $\mathbf{a} = [a_1, \ldots, a_N]$ be a sample of N i.i.d observations on individual wealth. For the prototype model in Section 2, the probability density function of wealth can be obtained from the subdensity functions that solve (14) and (15). Then, using the identity in (12), the (marginal) stationary probability density function of wealth is given by:

$$g(a_n \mid \boldsymbol{\theta}) = g(a_n, e_l \mid \boldsymbol{\theta}) + g(a_n, e_h \mid \boldsymbol{\theta}) \tag{16}$$

for each $n = 1, \ldots, N$, where $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^{\mathcal{M}}$ is the $\mathcal{M} \times 1$ vector of structural parameters in the model, and where $\boldsymbol{\Theta}$ is the parameter space, assumed to be compact. Then, for a given sample $\mathbf{a}$, the log-likelihood function can be computed as:

$$\mathcal{L}_N(\boldsymbol{\theta} \mid \mathbf{a}) = \sum_{n=1}^{N} \log g(a_n \mid \boldsymbol{\theta}), \tag{17}$$

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

$$\hat{\boldsymbol{\theta}}_N = \underset{\boldsymbol{\theta} \in \boldsymbol{\Theta}}{\arg\max}\, \mathcal{L}_N(\boldsymbol{\theta} \mid a_1, \ldots, a_N). \tag{18}$$

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 $U_t = \mathbb{E}_t \left[\int_t^\infty e^{\rho(s-t)} u\left(c_s\right) \mathrm{d}s\right]$ where $u\left(c_t\right) = \frac{c_t^{1-\gamma}}{1-\gamma}$. 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

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 $\mathcal{A}$ as described in Section 2. In particular, the wealth lattice is discretized using $I \leq N$ grid points on a partially ordered set defined by $\mathcal{A} = [\min(\mathbf{a}), \max(\mathbf{a})]$. Once the density function of wealth has been approximated, the log-likelihood function is constructed in two steps: (i) For each $a_n \in \mathbf{a}$, we use a piecewise linear interpolation to evaluate $g\left(a_n \mid \boldsymbol{\theta}\right)$; (ii) Once $g\left(a_n \mid \boldsymbol{\theta}\right)$ has been evaluated for all $a_n \in \mathbf{a}$, the log-likelihood function is computed using (17).

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 $\mathcal{L}\left(\boldsymbol{\theta} \mid \mathbf{a}\right)$ 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 $\mathcal{L}\left(\boldsymbol{\theta} \mid \mathbf{a}\right) \neq \mathcal{L}\left(\boldsymbol{\theta}_0 \mid \mathbf{a}\right)$, for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. Recently, Canova and Sala (2009) documented the existence of identification issues in the context of linearized representative agent 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 can be classified as observational equivalence, partial identification, weak identification, and asymmetric weak identification.

More precisely, a model suffers from *observational equivalence* if two vector of parameters deliver the same likelihood function. *Partial identification* occurs when a subset of the parameters enters the objective function proportionally and hence cannot be recovered individually. On the other hand, if the objective function exhibits reduced curvature along some dimension of the parameter space, even in the absence of the former two problems, then we say the model exhibits *weak identification*. Finally, if the likelihood function is asymmetric in the neighborhood of the maximum, and its curvature is insufficient only in a portion of the parameter space, then these group of parameters exhibit *asymmetric weak identification*.

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. For this reason, we instead rely on graphical evidence as in Canova and Sala (2009) to analyze identification issues in our nonlinear framework.

In particular, we generate a random sample of wealth of size $N = 10000$ from the model's population probability density function $g\left(a \mid \boldsymbol{\theta}_0\right)$ and then compute the log-likelihood profile in each dimension of $\boldsymbol{\theta}$ by varying one parameter at a time in an economically reasonable neighborhood of its true value while keeping the remaining parameters at their population values.

The population values of the structural parameters of the model, $\boldsymbol{\theta}_0$, are provided in Table 1. These values are fairly standard in the literature. In the model, time is measured in years and parameter values should be interpreted accordingly. The labor efficiency process is set to match the long run employment-unemployment dynamics of the US economy. 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 (13).

Since in our Bewley-Hugget-Aiygari economy the labor efficiency endowment process is completely exogenous, we will focus our attention on the ability to identify the supply side and household's preference parameters, conditional on a fixed level of income risk. Hence, the parameters describing the endowment process will remain fixed to the values in Table 1. To avoid unnecessary additional notation, the parameter vector $\boldsymbol{\theta}$ will refer exclusively to $\boldsymbol{\theta} = \{\gamma, \rho, \alpha, \delta\}$[2].

The resulting log-likelihood profiles are reported in Figure 1. The figure displays two im-

[2]A sensitivity analysis on the full parameter set, i.e. including the parameters of the income process, can be found in Parra-Alvarez (2015).

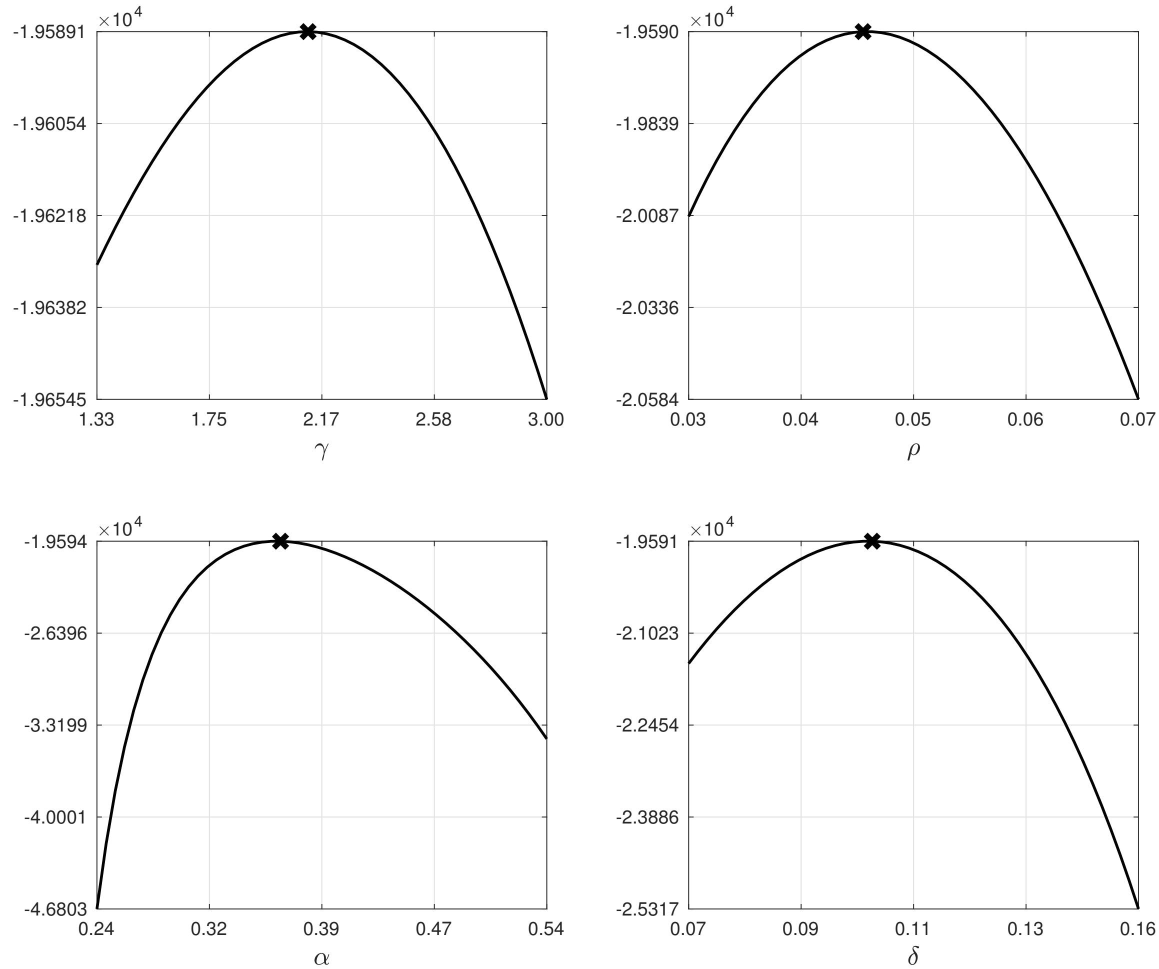

Figure 1. Log-likelihood profile. The graph shows the log-likelihood function $\mathcal{L}(\boldsymbol{\theta} \mid \mathbf{a})$ for selected parameters as a function of the parameter space using a random sample of wealth of size $N = 10000$. The cross "$\times$" marks the parameter value associated with the maximum value of the log-likelihood profile.

portant features. First, the log-likelihood function is uniquely maximized at $\boldsymbol{\theta}_0$, ruling out this way identification problems related to observational equivalence. Second, the objective function exhibits enough curvature in the neighborhood of $\boldsymbol{\theta}_0$, suggesting strong identification power in each dimension of the parameter space.

Figure 2 plots the contours of the log-likelihood profile for pairwise combinations of the model parameters while keeping the remaining ones at their true population values. The upper right panel reveals a ridge in the log-likelihood function in the α - δ space, the parameters associated to the supply side of the model. Thus, a proportional increase in both parameters produce almost observational equivalent objective functions, a clear indication of potential partial identification problems. Interestingly, this relationship is not linear. In fact, the ridge appears to be slightly concave with respect to α. Although much less severe,

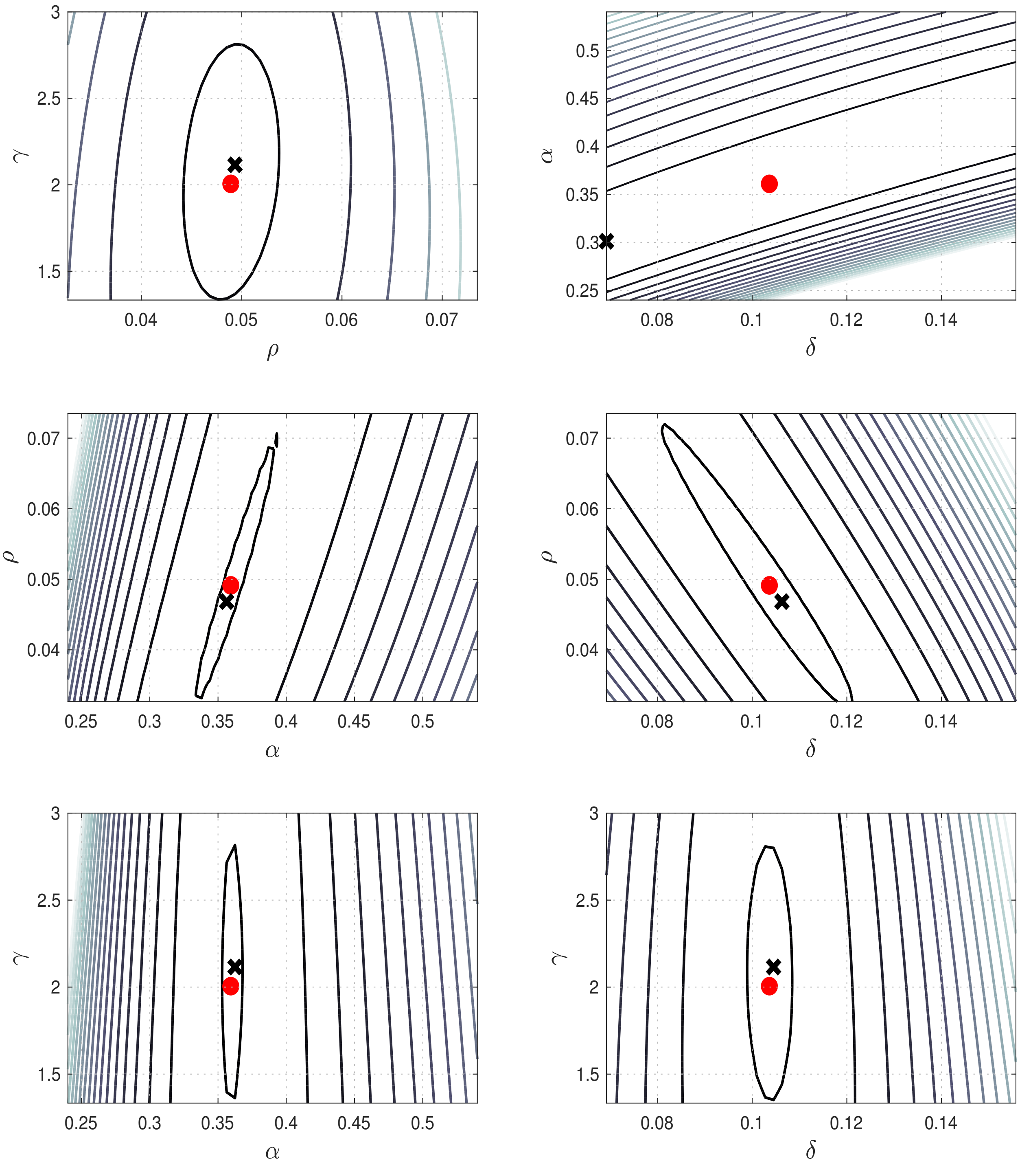

Figure 2. Log-likelihood profile contours for selected parameters. TThe graph shows the log-likelihood function $\mathcal{L}(\boldsymbol{\theta} \mid \mathbf{a})$ for selected parameters as a function of the parameter space using a random sample of wealth of size $N = 10000$. The cross "$\times$" marks the parameter value associated with the maximum value of the log-likelihood profile. The population values for the structural parameters, $\boldsymbol{\theta}_0$, are given in Table 1. A circle "●" indicates the true parameter values, and cross "$\times$" the combination of parameters that deliver the maximum log-likelihood.

the log-likelihood functions for the discount rate, ρ, and the supply side parameters, α and δ, also reveal a similar pattern. Nonetheless, the there is still enough curvature compared to the case of α and δ. The risk aversion parameter γ, on the other hand, is strongly identified in combination with any other parameter[3].

The tight relation between α and δ is an example of identification deficiencies that are rooted in the economic theory and will persist even in samples of finite size. As an example consider the steady state capital-output ratio, K/Y, from the standard neoclassical growth model. Assuming that the gap $r - \rho$ does not vary significantly with α and δ, and thus implicitly 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.

4. 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 on simulated cross-sectional data of individual wealth. The experiment is carried out by drawing 500 samples 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 (18).

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 3 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

[3] In the accompanying web appendix to this article we report the results from a sensitivity analysis of the model's implied probability density function to changes in the parameter values in order to diagnose irregularities that can translate into identification problems. The results are similar to the ones described here both for individual parameters as well as for pairwise combinations. Although not equivalent, such analysis compares to the population identification diagnosis carried out in Canova and Sala (2009) in that the conclusions that can be drawn are independent of the sample size, and could emerge directly from the economic theory.

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

	$N = 1000$			$N = 5000$			$N = 10000$			$N = 50000$		
$\boldsymbol{\theta}$	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

negligible in small samples.

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

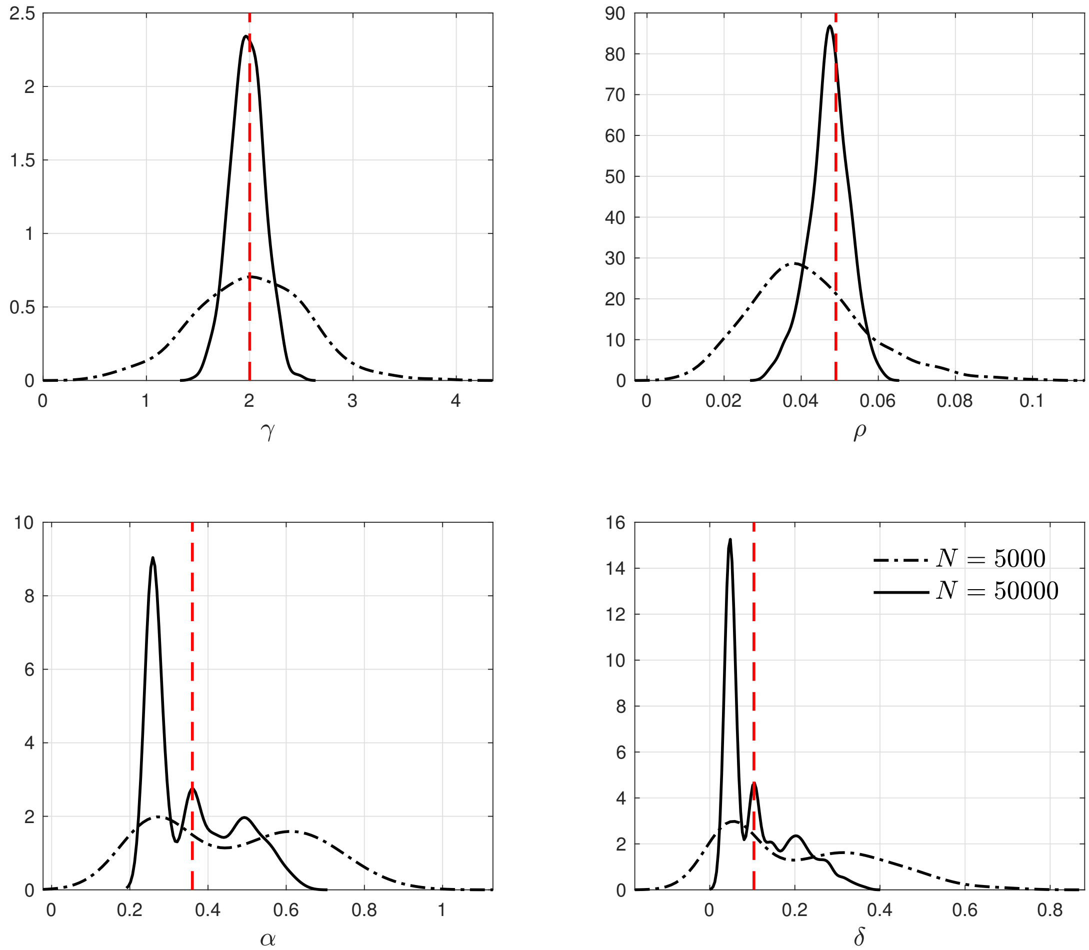

Figure 3. Finite sample distribution of parameter estimates. The graph plots the kernel density of estimated parameters across 500 random samples of size $N = 5000$ (dashed line) and $N = 50000$ (continuous line) generated from the true data generating process. The vertical line denotes the true parameter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

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.

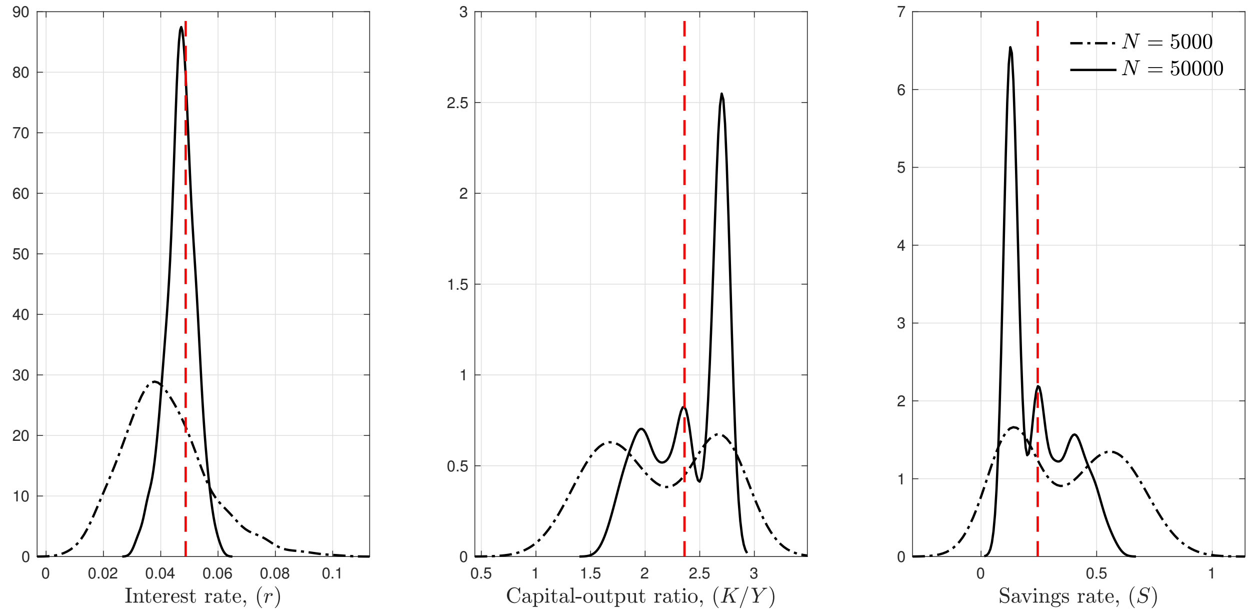

Figure 4. Steady state macroeconomic aggregates density estimates. The graph shows the implied distribution of the steady state values for the interest rate, the capital-output ratio, and the savings rate using the estimated parameters from the Monte Carlo experiment with $N = 5000$ and $N = 50000$.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

5. Calibration and estimation

Our findings indicate that across some dimensions of the parameter space the maximum likelihood estimator delivers biased and poorly identified estimates. 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 linearized representative agent 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.

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	-	-	-	-

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 a considerable bias in the parameters that were originally free 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

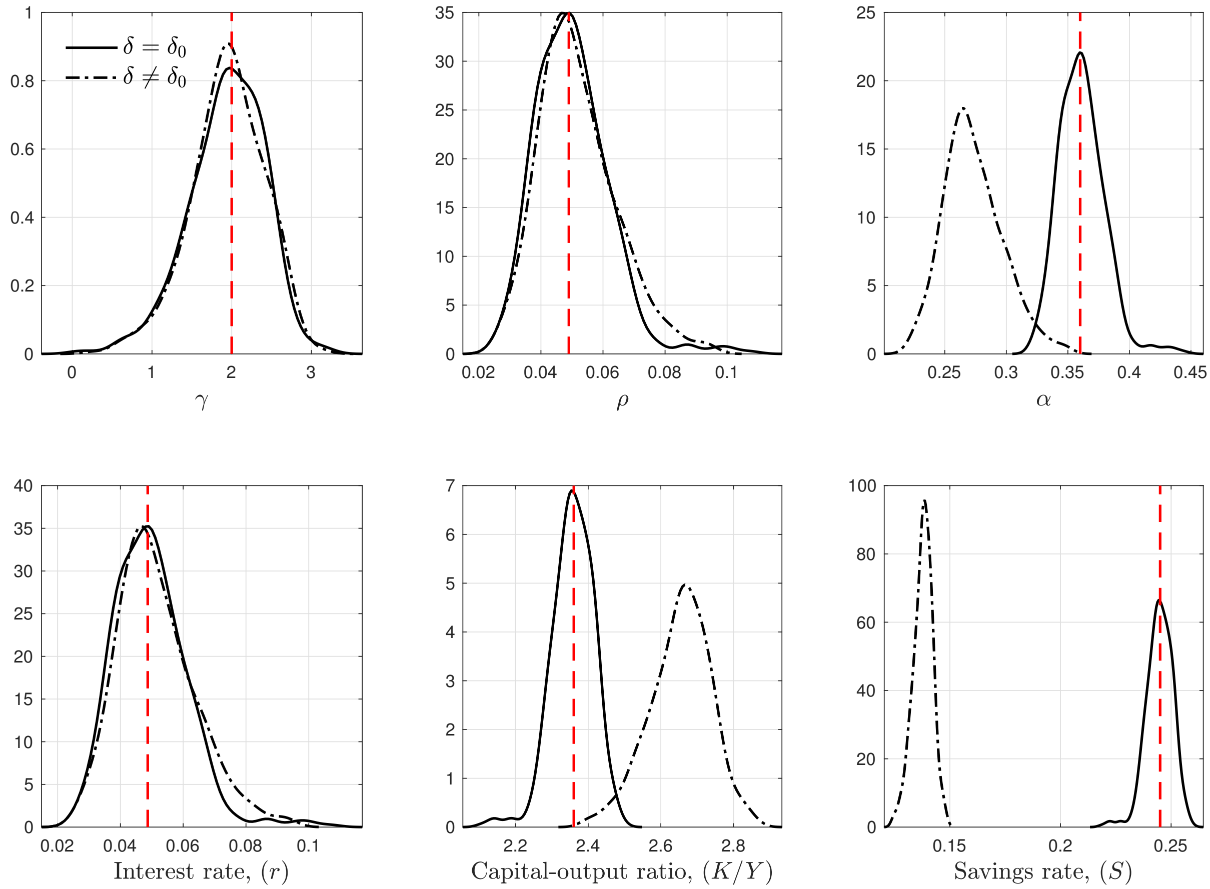

Figure 5. Parameter and macroeconomic aggregates densities with fixed δ. The graph shows the Monte Carlo implied distribution of the parameter estimates and the steady state interest rate, capital-output ratio, and savings rate when the depreciation rate is either fixed to its true value in the population,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

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 issues but cannot be precisely pinned down when δ is miscalibrated.

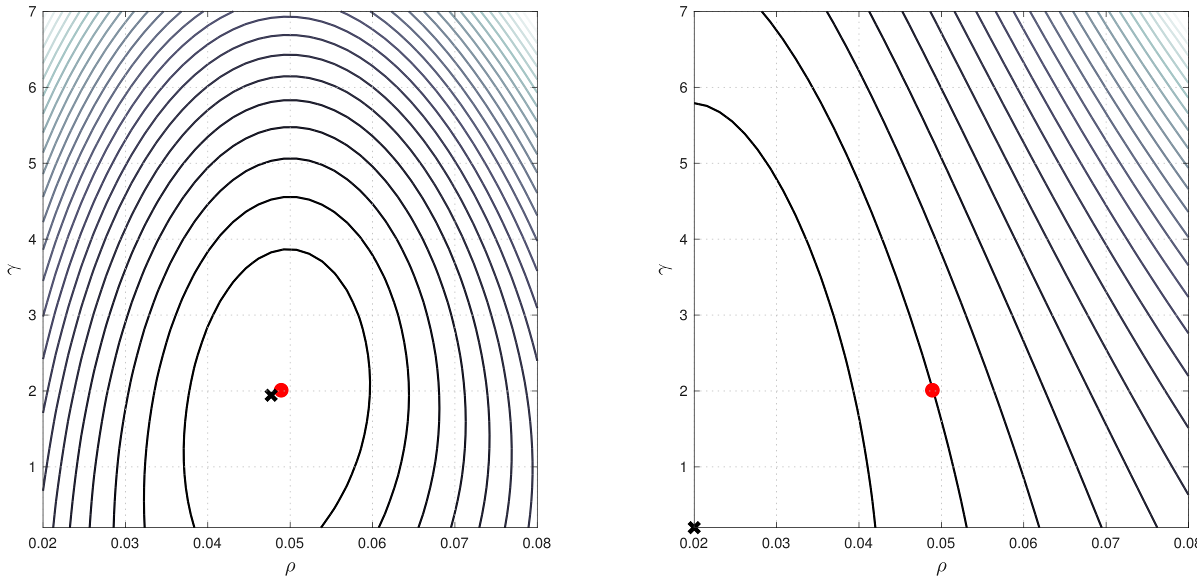

Figure 6. Log-likelihood function contour for fixed α and δ. The graph shows the contour of the log-likelihood function for combinations of γ and ρ within a reasonable economic range when a sample of size $N = 5000$ is generated from the true model. The contours to the left calibrate $\alpha = \alpha_0$ 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.

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 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

Table 4. Estimated parameters and labor efficiency levels.

The table reports the maximum likelihood estimates of the model parameters and the associated standard errors (in parenthesis).

γ	ρ	δ	e_1	e_2	e_3	e_4
3.7398	0.1455	0.2880	45.79	265.10	412.17	8553.5
(0.0387)	(0.0040)	(0.0020)	(0.6245)	(2.4009)	(4.5087)	(133.2604)

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

6. Empirical illustration

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

In order to match the high degree of wealth inequality observed in the data, we expand the number of labor efficiency states to four, similar to Castañeda et al. (2003)[4]. As our identification analysis suggests, we do not attempt to estimate α and δ jointly. Instead, we calibrate α to 0.33 and estimate all other remaining parameters of the model, including all parameters of the labor efficiency process. The estimation only includes households with positive net worth in order to compute meaningful empirical inequality measures. No further truncation or normalization has been applied to the data. Technical details on the estimation setup are provided in Appendix C. We also performed the estimation where i) none of the parameters were calibrated ii) δ was calibrated, iii) both α and δ were calibrated. The resulting log-likelihoods were much lower and unable to match the Lorenz curve of the data.

Tables 4 and 5 report the estimation results together with standard errors computed using a numerical hessian matrix in parenthesis. The estimated risk aversion parameter is 3.7398 and slightly larger than conventional calibrated values. This could induce a higher buffers stock saving. The estimated discount rate and depreciation rate are much higher than values reported in the literature. It should be noticed that the frequency of the estimated

[4]In the original working paper (Parra-Alvarez et al., 2017), we estimated different variations of a two labor efficiency states model and included additionally employment status from the SCF data to account for income. As expected, none of the results could match the high degree of the wealth inequality.

Table 5. Estimated transition probabilities of efficiency levels.

The table reports the maximum likelihood estimates of the transition probabilities and their associated standard errors (in parenthesis). The last probability value in the each row of the transition probability matrix is not estimated given the fact that all probabilities in each row have to sum up to one.

	to e'_j			
from e_i	e_1	e_2	e_3	e_4
e_1	0.7727 (0.0002)	0.1905 (0.0002)	0.0342 (0.0002)	0.0026
e_2	0.7439 (0.0006)	0.2035 (0.0007)	0.0487 (0.0007)	0.0039
e_3	0.3074 (0.0017)	0.1400 (0.0018)	0.5094 (0.0020)	0.0433
e_4	0.3330 (0.0082)	0.0010 (0.0049)	0.1388 (0.0096)	0.5272
	Implied stationary probabilities, $p(e_i)$			
	0.7281	0.1869	0.0728	0.0123

model is not pinned down *a priori*. Lower model frequency (e.g. from quarterly to annual) will induce both higher discount rates (lower discount factor) and higher depreciation rates. The estimated parameters of the labor efficiency process are mostly likely the key driving force for the high degree of the wealth inequality.

Table 6 compares some wealth statistics computed from the data from the estimated model. It reports the Gini coefficient and the percentage of 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. Strikingly, the estimated model can match the data on wealth inequality quite well considering the simplistic nature of the model. Similar to previous literature which successfully matches the wealth distribution by focusing on labor income, our estimates indicate that the data favors the inclusion of an ”awesome state” (see Benhabib and Bisin (2017) for a discussion.). In particular, they suggest that the most productive households are more than 20 times more productive than those households in the second most productive state, i.e., $e_4/e_3 > 20$. At the same time, the most productive households face a considerable high risk (1/3) of becoming the least productive ones. These forces generate a highly skewed stationary labor income process and contribute to the skewness of the implied wealth distribution. It should be stressed that we view our results not as evidence *for* this particular income process. As Benhabib and Bisin (2017) point out, the implication of the ”awesome state” in the labor income process is very likely to be counterfactual to the actual income data. The estimated income process simply captures all other relevant wealth inequality driving forces (bequest, entrepreneur risk, explosive wealth

Table 6. 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 model.

		% wealth in top		
	Gini Coefficient	1%	5%	20%
SCF 2013 data	0.8047	31.19	58.14	83.34
Model implied	0.8069	25.37	60.88	84.41

accumulation, etc.) which are not present in our the simple model.

7. Conclusions

In this paper we introduce a simple full information likelihood approach to estimate the structural parameters of heterogeneous agent models using cross-sectional data. Following the work of Bayer and Wälde (2011, 2013) and Achdou et al. (2014), the feasibility of our approach follows from 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 individual state variables which can then be used to build the model's likelihood function.

We also study the identification power of our maximum likelihood estimator based on 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 with aggregate certainty are well identified and can be accurately estimated, the parameters related to the supply side of the economy exhibit partial identification problems. In particular, the latter cannot be separately identified as increasing both parameters proportionally may leave the model's implied wealth distribution, and hence the likelihood function, unchanged. This partial identification problem is illustrated in finite samples 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-Hugget-Aiyagari model using the wealth data reported in the 2013 Survey of Consumer Finances. Despite the simplistic nature of the model, the estimated model matches the data quite well as measured by the implied Gini coefficient and the distribution of wealth across top percentiles.

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:

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

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:

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

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

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

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

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

Using (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:

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

The first-order condition for an interior solution reads:

$$u'(c_t) = V_a(a_t, e_t), \tag{19}$$

for any $t \in [0, \infty)$, making optimal consumption $c_t = c(a_t, e_t)$ 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:

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

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 $\mathbb{A}f(a_t, e_t) \equiv \frac{\mathbb{E}_t \mathrm{d}f(a_t, e_t)}{\mathrm{d}t}$:

$$\begin{aligned} \frac{\mathbb{E}_t \mathrm{d}f(a_t, e_t)}{\mathrm{d}t} &= f_a(a_t, e_t)\, s(a_t, e_t) \\ &\quad + (f(a_t, e_l) - f(a_t, e_h))\,\phi_{hl} + (f(a_t, e_h) - f(a_t, e_l))\,\phi_{lh}. \end{aligned} \tag{20}$$

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:

$$\mathbb{E}f(a_t, e_t) = \mathbb{E}f(a_s, e_s) + \int_s^t \mathbb{E}(\mathbb{A}f(a_\tau, e_\tau))\,\mathrm{d}\tau. \tag{21}$$

Differentiating (21) with respect to time:

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

that is:

$$\frac{\partial}{\partial t}\mathbb{E}f(a_t, e_t) = \underbrace{\int_{-\infty}^{\infty} \mathbb{A}f(a_t, e_h)\, g(a_t, e_h, t)\,\mathrm{d}a_t}_{\omega_{e_h}} + \underbrace{\int_{-\infty}^{\infty} \mathbb{A}f(a_t, e_l)\, g(a_t, e_l, t)\,\mathrm{d}a_t}_{\omega_{e_l}} \tag{22}$$

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 (20) we note that:

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

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:

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

and therefore:

$$\frac{\partial}{\partial t}\mathbb{E}f\left(a_t, e_t\right) = \int_{\underline{a}}^{\infty} f\left(a_t, e_h\right) \frac{\partial}{\partial t} g\left(a_t, e_h, t\right) \mathrm{d}a_t + \int_{\underline{a}}^{\infty} f\left(a_t, e_l\right) \frac{\partial}{\partial t} g\left(a_t, e_l, t\right) \mathrm{d}a_t. \tag{23}$$

Finally we equate (22) and (23) and collect terms to obtain:

$$\int_{\underline{a}}^{\infty} f\left(a_t, e_h\right) \varphi_{e_h} \mathrm{d}a_t + \int_{\underline{a}}^{\infty} f\left(a_t, e_l\right) \varphi_{e_l} \mathrm{d}a_t = 0 \tag{24}$$

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 (24) 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 (14) and (15).

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

$$G\left(a_t, e_t\right) = \int_{\underline{a}}^{a} g\left(x_t, e_t\right) \mathrm{d}x_t \tag{25}$$

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 $a \to \infty$, (11) implies that $\lim_{a_t \to \infty} G\left(a_t, e_t\right) = p\left(e_t\right)$. Similar to (12), the (unconditional) stationary probability distribution of wealth can be computed as:

$$G\left(a_t\right) = G\left(a_t, e_h\right) + G\left(a_t, e_l\right) \tag{26}$$

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

$$\mathcal{G} = \frac{1}{\mu} \int_{\underline{a}}^{\infty} G\left(a_t\right)\left(1 - G\left(a_t\right)\right) \mathrm{d}a_t \tag{27}$$

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

C. Data and estimation settings

The cross-section of individual wealth used in Section 6 for the estimation of the Bewley-Hugget-Aiyagari model is obtained from the 2013 Survey of Consumer Finances (SCF). In particular, our wealth data matches the net-worth reported in the Summary Extract Public Data provided by the SCF. The net-worth data is then resampled with associated weights to obtain an equally weighted sample of household wealth[5]. 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 $I = 500$ non-equally spaced points. The (negative) log-likelihood function is minimized using a constrained numerical minimizer routine provided by the Optimization Toolbox from MathWorks. Several algorithms (interior point, sequential quadratic programming and active-set) are employed to search for the minimum. The standard errors are based on the numerical Hessian computed with the DERIVESTsuite which employs a robust adaptive numerical differentiation based on a finite differencing scheme.

[5]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. 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.

Challe, E., J. Matheron, X. Ragot, and J. F. Rubio-Ramirez (2017): "Precautionary saving and aggregate demand," *Quantitative Economics*, 8, 435–478.

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

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 A. Maussner (2009): *Dynamic General Equilibrium Modeling*, Springer, 2nd ed.

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

Holm, M. B. (2017): "Monetary Policy Transmission with Income Risk," Available at ssrn: https://ssrn.com/abstract=3128510, •.

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

İmrohoroğlu, 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.

Parra-Alvarez, J. C. (2015): "Solution Methods and Inference in Continuous-Time Dynamic Equilibrium Economies," Ph.D. thesis, Aarhus University.

Parra-Alvarez, J. C., O. Posch, and M.-C. Wang (2017): "Identification and Estimation of Heterogeneous Agent Models: A Likelihood Approach," CREATES Research Papers 2017-35, Department of Economics and Business Economics, Aarhus University.

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

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.

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.