Estimation of Heterogeneous Agent Models: A Likelihood Approach

September 2022

Abstract

Using a Bewley-Hugget-Aiyagari model we show how to use the Fokker-Planck equation for likelihood inference in heterogeneous agent (HA) models. We study the finite sample properties of the maximum likelihood estimator (MLE) in Monte Carlo experiments using cross-sectional data on wealth and income. We use the Kullback-Leibler divergence to investigate identification problems that may affect inference. Unrestricted MLE leads to considerable biases of some parameters. Calibrating weakly identified parameters is shown to be useful to pin down the remaining structural parameters. We illustrate our approach by estimating the model for the U.S. economy using the Survey of Consumer Finances.

Word Count: 12,400 approximately.

Keywords: Fokker-Planck equations, Heterogeneous agents, Kullback-Leibler divergence, Maximum likelihood.

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

1. Introduction

Heterogeneous agent (HA) models have become an extensively used tool for the study and evaluation of macroeconomic policies and welfare implications. They have 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 of HA models during the last three 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 Ozkan et al., 2016; Holm, 2022; Kaplan et al., 2018; Wong, 2021).

Currently, the main workhorse of household heterogeneity is based on the models by 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 for which there exists a wide consensus in the literature, or chosen in such a way that they minimize the distance between a subset of moments obtained from the model and the same moments computed from the data, or by a combination of both. Accordingly, calibration can be classified 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. (2019), Luo and Mongey (2019), Abbott et al. (2019).

On the other hand, full information methods which rely on the entire probability distribution of the model have received less attention. Given the increased quality and quantity of household data, the first contribution of this paper is to introduce a likelihood framework that can be used to estimate the structural parameters of HA models using the information content on a sample of cross-sectional data, $\mathbf{x}_i$. The approach proposed here uses the fact that any HA model with parameter vector $\boldsymbol{\theta}$ induces a joint distribution of idiosyncratic state variables with probability density function $g(\boldsymbol{\theta})$ that can be used to compute the likelihood of the data, $\sum_i \log g(\boldsymbol{\theta} \mid \mathbf{x}_i)$. In this paper, we rely 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. Hence, our approach applies exclusively to the estimation of structural parameters that affect the steady state of macroeconomic aggregates from microeconomic data. Using a standard Bewley-Hugget-Aiyagari, we show how to combine the time-invariant equilibrium joint probability distribution of wealth and income with a sample of observations on individual's wealth and labor status to estimate different subsets of structural parameters of the model, e.g., household preferences, production technology, and/or idiosyncratic income dynamics. Likelihood-based methods provide a natural point for the investigation of efficiency and identification properties to the extent that the commonly used partial information methods, like the simulated method of moments (SMM), only rely on a subset of the full information provided by the likelihood function. Hence, the full information approach allows 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 state variables in HA models is not straightforward as it turns out to be a complicated endogenous and nonlinear object that usually has to be numerically approximated either by Monte Carlo simulation or functional approximation techniques (see Heer and Maussner, 2009). More recently, Bayer and Wälde (2010a,b, 2011), Achdou et al. (2014), and Gabaix et al. (2016) have suggested the use of Fokker-Planck equations, also known as Kolmogorov's Forward equations, for the analysis of endogenous distributions in macroeconomics. These partial differential equations (PDEs) 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 PDEs 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 in Candler (1999) and Achdou et al. (2022).

A condition for the maximum likelihood (ML) estimator to deliver consistent estimates of the model parameters, and for 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 its dimensions. Lack of identification leads to misleading statistical inference that may suggest the existence of some features in the data that are in fact absent. Therefore, it is important to verify the identification condition prior to estimation[1]. The recent contributions of Canova and Sala (2009), Iskrev (2010), Komunjer and Ng (2011), Qu and Tkachenko (2012), 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 our likelihood-based framework. 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 in Rothenberg (1971) for linear models cannot be applied. Instead, we propose to use the Kullback-Leibler (KL) divergence between two distribution functions (see Kullback and Leibler, 1951, Kullback, 1959 and McCulloch, 1989) to investigate the identification power of our ML estimator. The KL divergence is computed for the model's

[1]See Ireland (2004) and Mumtaz and Zanetti (2015) for an early discussion of potential identification issues in the estimation of representative agent DSGE models.

implied distribution and hence it is independent of the data. This allows the researcher to analyze the sensibility of the model's probability distribution function to changes in the parameter values even before any estimation takes place. Lack of variability along any dimension of the parameter space provides an early diagnosis for potential identification issues. In fact, irregular behavior in the density function found at this stage will impact the estimator's objective function and hence limit its ability to accurately identify the parameters of the model using the likelihood of the data.

The estimation approach proposed in this paper differs from related contributions by Mongey and Williams (2017), Williams (2017), and Winberry (2018). They 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 statistical methods do not make any use of the model's implied probability density function. Challe et al. (2017) extend this approach by including a subset of the preference parameters in the estimation step and hence some knowledge of the cross-sectional probability density function is in principle required. However, in their quantitative exercise they collapse the model's density function to a single mass point and therefore do not make use of the entire distribution function in the estimation process.

More recently, Bayer et al. (2020) make use of Bayesian methods to estimate the dynamics of the aggregate shocks and frictions in a New Keynesian HA model (HANK), while calibrating most of the structural parameters. Using the solution method in Reiter (2009) and Bayer and Luetticke (2020), the likelihood function of the linearized model is computed via the Kalman filter. They estimate the model using both data on macroeconomic aggregates together with cross-sectional information in the form of snapshots of the cross-sectional distribution, but not the entire cross-sectional distribution of the individual's state variables. A similar approach is followed by Auclert et al. (2021) using the moving average representation implied by the sequence-space Jacobian method. On the other hand, Fernández-Villaverde et al. (2020) use continuous-time methods in order to exploit the Fokker-Planck equation that describes the time evolution of the probability density function of aggregate macroeconomic variables in a non-linear HANK model. Using quarterly data on aggregate output in the U.S., they estimate the volatility parameter of the aggregate exogenous shock conditional on the calibrated values of all the remaining parameters. Unfortunately, the paper is silent about including microeconomic data in the estimation. In general, the existing frameworks only exploit the information content of macroeconomic aggregates in addition to some cross-sectional information in the form of summary statistics. In contrast, in this paper we use of the information available in the entire cross-sectional distribution of microeconomic data to estimate parameters that enter directly in the computation of the steady state of the economy. Closer to our approach is the work by Papp and Reiter (2020) and Liu and Plagborg-Møller (2021). They develop likelihood-based methods to estimate the parameters of HA models using simultaneously macroeconomic time series and microeconomic data in the form of snapshots (moments) of either repeated cross-sections or panel data. However, both papers only provide proof-of-concept examples where the focus is on a reduced number of parameters affecting primarily the exogenous processes of idiosyncratic states.

The rest of the paper is organized as follows. Section 2 shows how to compute the model's likelihood function from the stationary joint density function that solves the model's Fokker-Planck equation. To illustrate our approach, we introduce a 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 labor productivity. We characterize and solve for the stationary competitive equilibrium which equip us with a time-invariant joint probability distribution of wealth and income for estimation and/or identification purposes. While our framework can be easily extended to include cross-sectional data on a number of additional endogenous variables, e.g., individual consumption, our general equilibrium approach limits this possibility unless additional features such as measurement errors are considered.

Section 3 examines the finite sample properties of the MLE using a Monte Carlo experiment. We identify potential biases and the precision of the estimates along different dimensions of the parameter space. Our results suggest that estimating the complete set of parameters using cross-sectional data on income and wealth leads to considerable biases on the parameters describing the HA income process, but also on other parameters such as the coefficient of relative risk aversion. As a negative result, the biases in these parameters persists in large samples and their distributions tend to be wide and skewed. In Section 4 we compute the KL divergence associated to the model's implied distribution function and find that the poor performance of the MLE can be explained by the insensibility of the wealth-income distribution to changes in this subset of parameters. On the other hand, we find that cross-sectional data is informative for parameters related to the supply side, like the capital share in output and the depreciation rate.

A standard practice in applied macroeconomics when identification problems arise 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 a strategy and find that even in cases where some parameters are mis-calibrated the MLE, conditional on some parameters being calibrated, improves the finite sample properties of the 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. economy using individual data on wealth and income from the 2013 Survey of Consumer Finances. Section 7 concludes.

2. Structural estimation

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 on a sample of cross-sectional data. We use the fact that any HA model with parameter vector $\boldsymbol{\theta}$ induces a joint distribution of idiosyncratic state variables with stationary probability density function $g(\boldsymbol{\theta})$ that can be used to compute the likelihood of the data. As shown below, our approach makes use of the model's Fokker-Planck equation to approximate the stationary probability density function of the state variables.

More specifically, let $\{\mathbf{x}_i\}_{i=1}^{N}$ be a sample of N i.i.d observations of the state variables. Then, the log-likelihood function of any HA model can be computed as

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

where $\boldsymbol{\theta} \in \boldsymbol{\Theta} \subset \mathbb{R}^{\mathcal{M}}$ is the $\mathcal{M} \times 1$ vector of structural parameters, and where $\boldsymbol{\Theta}$ is the parameter space, assumed to be compact. The ML estimator, $\hat{\boldsymbol{\theta}}_N$ is defined as

$$\hat{\boldsymbol{\theta}}_N = \underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\max}\ \mathcal{L}_N\left(\boldsymbol{\theta} \mid \mathbf{x}_1 \ldots, \mathbf{x}_N\right). \tag{2}$$

A prototypical heterogeneous agent model

To show how the estimation approach works, we consider a prototypical HA model *á la* Bewley-Hugget-Aiyagari as in Achdou et al. (2022). In our economy there is no aggregate uncertainty, and we assume that all aggregate variables are constant and equal to their steady-state values, while at the individual level households face idiosyncratic uninsurable risk and variables change over time in a stochastic way.

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

$$U_0 \equiv \mathbb{E}_0 \int_0^{\infty} e^{-\rho t} u(c_t) \mathrm{d}t, \tag{3}$$

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

$$u\left(c_t\right) = \begin{cases} c_t^{1-\gamma}/(1-\gamma) & \text{for } \gamma \neq 1 \\ \log\left(c_t\right) & \text{for } \gamma = 1. \end{cases}$$

Here, $\gamma > 0$ denotes the coefficient of relative risk aversion (or the inverse of the elasticity of intertemporal substitution, EIS). At time $t = 0$, the agent knows his initial wealth and income level and chooses the optimal path of consumption $\{c_t\}_{t=0}^{\infty}$ subject to

$$\mathrm{d}a_t = (ra_t + we_t - c_t)\mathrm{d}t, \quad a_0 \in \mathcal{A}, \tag{4}$$

where $a_t \in \mathcal{A} \subset \mathbb{R}$ denotes the household's 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). 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.

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

$$\mathrm{d}e_t = -\Delta_e \mathrm{d}q_{1,t} + \Delta_e \mathrm{d}q_{2,t}, \quad \Delta_e \equiv e_h - e_l \quad \text{and} \quad e_0 \in \mathcal{E}, \tag{5}$$

where Δ_e can be interpreted as the labor efficiency gap. 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 nonnegativity of consumption. Hence, $\mathcal{A} = [\underline{a}, \infty)$.

Production possibilities and macroeconomic identity

Aggregate output in this economy, Y, is produced by identical firms owned by the households. The representative firm combines aggregate capital, K, and aggregate labor, L, through a constant return to scale production function $Y = K^{\alpha} L^{1-\alpha}$, with $\alpha \in (0, 1)$, to maximize its profits.

We 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

$$Y = C + \delta K \tag{6}$$

holds $\forall t$, where C and δK denote, respectively, aggregate consumption and aggregate investment. We have removed the temporal subscript t from all aggregate variables to indicate that the economy is in a stationary equilibrium[2].

Equilibrium

In this economy, households face uncertainty regarding their future level of 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-income 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 the density function is independent of time and thus it simplifies to $g(a_t, e_t)$.

Households. 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

[2] Introducing a time series dimension at the aggregate level through an aggregate productivity shock is straightforward from a modeling perspective (see Ahn et al., 2017). However, this increases the computational time dramatically making any extensive Monte Carlo investigation infeasible.

by the Hamilton-Jacobi-Bellman equation (HJB)

$$\rho V(a,e) = \max_{c\in\mathbb{R}^+} \Big\{ u(c) + V_a(a,e)(ra + we - c) \\ + (V(a,e_l) - V(a,e_h))\phi_1(e) + (V(a,e_h) - V(a,e_l))\phi_2(e) \Big\}, \quad (7)$$

where $V(a,e)$ denotes the value function of the agent[3]. The first-order condition for an interior solution reads

$$u'(c) = V_a(a,e) \quad (8)$$

for any $t \in [0,\infty)$, making optimal consumption a function only of the state variables and independent of time, $c = c(a,e)$.

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

$$\rho V(a,e_l) = u(c(a,e_l)) + V_a(a,e_l)(ra + we_l - c(a,e_l)) + (V(a,e_h) - V(a,e_l))\phi_{lh}, \quad (9)$$

$$\rho V(a,e_h) = u(c(a,e_h)) + V_a(a,e_h)(ra + we_h - c(a,e_h)) + (V(a,e_l) - V(a,e_h))\phi_{hl} \quad (10)$$

As argued in Achdou et al. (2022), Equation (8) holds for all $a > \underline{a}$ since the borrowing constraint never binds in the interior of the state space. Therefore, the system of equations formed by (9) and (10) does not get affected by the existence of the inequality constraint $a \geq \underline{a}$, and instead gives rise to a state-constraint boundary condition that ensures that the borrowing constraint is never violated.

Firms. The representative firm rents capital and labor from the households in perfectly competitive markets. Hence, in equilibrium the production factors are paid their respective marginal products

$$r = \alpha K^{\alpha-1}L^{1-\alpha} - \delta \quad \text{and} \quad w = (1-\alpha)K^{\alpha}L^{-\alpha}, \quad (11)$$

[3]A complete derivation of the HJB equation, the Fokker-Planck equations that described the subdensity functions of wealth, and the stationary probability distributions of the efficiency endowments can be found in the Online Appendix.

where $K = \sum_{e_t \in \{e_l, e_h\}} \int_{\underline{a}}^{\infty} a_t g\left(a, e\right) \mathrm{d}a$ and $L = \sum_{e \in \{e_l, e_h\}} \int_{\underline{a}}^{\infty} e_t g\left(a, e\right) \mathrm{d}a$. This form of aggregation provides a link between the dynamics and randomness that occurs at the micro level with the deterministic behavior at the macro level.

Distribution of endowments and wealth. Given its dependence on one continuous random variable and one discrete random variable, the stationary joint density function, $g\left(a, e\right)$, can be split into $g\left(a, e_h\right)$ and $g\left(a, e_l\right)$. Following Khieu and Wälde (2019), we refer to these individual probability functions as *subdensities*. For each $e \in \mathcal{E}$, it follows that $g\left(a, e\right) \equiv g\left(a \mid e\right) p\left(e\right)$, implying that

$$\int g\left(a, e\right) \mathrm{d}a \quad = \quad p\left(e\right), \tag{12}$$

where $p\left(e\right) \equiv \lim_{t \to \infty} \mathbb{P}\left(e_t = e\right)$ is the stationary probability of having an efficiency endowment equal to e. Then, the (marginal) stationary density function of wealth is

$$g\left(a\right) = g\left(a, e_h\right) + g\left(a, e_l\right). \tag{13}$$

Given our two state Markov process for the endowment of labor efficiency units it is possible to show that its stationary distribution is given by

$$p\left(e_h\right) = \frac{\phi_{lh}}{\phi_{hl} + \phi_{lh}}, \text{ and } p\left(e_l\right) = \frac{\phi_{hl}}{\phi_{hl} + \phi_{lh}}. \tag{14}$$

Let $s\left(a, e\right) = ra + we - c\left(a, e\right)$ denote the optimal savings function for an individual with an efficiency endowment equal to $e \in \mathcal{E}$. The subdensities in (13) correspond to the solution of the following non-autonomous quasi-linear system of differential equations known as (stationary) Fokker-Planck equations

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

$$s\left(a, e_h\right) \frac{\partial}{\partial a} g\left(a, e_h\right) = -\left(r - \frac{\partial}{\partial a} c\left(a, e_h\right) + \phi_{hl}\right) g\left(a, e_h\right) + \phi_{lh} g\left(a, e_l\right), \tag{16}$$

where the partial derivatives with respect to wealth describe the cross-sectional dimension of the density function. The system of equations formed by (15) and (16) takes as given the optimal policy functions for consumption. This feature creates a recursive structure within the model that facilitates its solution: households and firms meet at the marketplace 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.

Equilibrium. 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 income is time-invariant, and all markets clear. More specifically, while the distribution of wealth is constant for both the low and high efficiency workers and the number of low and high efficiency workers is also constant, the households are not characterized by constant wealth levels and efficiency status over time. Achdou et al. (2022) show that such stationary equilibrium is unique if the EIS is greater or equal than one, i.e. $1/\gamma \geq 1$. Closely related results for the case of discrete-time economies have been shown in Açikgöz (2018) and Light (2020).

The solution of our prototype economy is not available in closed form. Therefore, for a given set of parameter values, the stationary competitive equilibrium is numerically approximated on a discretized state space for $\mathcal{A}$ using the finite-difference approaches in Candler (1999) and Achdou et al. (2022). A detailed description of the algorithm and its implementation can be found in the Online Appendix.

The likelihood function

Let $\{\mathbf{x}_i\}_{i=1}^N = \{\mathbf{a}_i, \mathbf{e}_i\}_{i=1}^N = \{a_1, e_1 \ldots, a_N, e_N\}$ be a sample of N i.i.d observations on individual wealth and income, respectively. Then, the log-likelihood function of the prototype model can be computed as

$$\mathcal{L}_N\left(\boldsymbol{\theta} \mid \mathbf{a}, \mathbf{e}\right) = \sum_{i=1}^{N} \log g\left(a_i, e_i \mid \boldsymbol{\theta}\right). \tag{17}$$

Alternatively, by using identity (13), it is possible to obtain the marginal density of wealth as $g(a_i \mid \boldsymbol{\theta}) = g(a_i, e_l \mid \boldsymbol{\theta}) + g(a_i, e_h \mid \boldsymbol{\theta})$ for each $i = 1, \ldots, N$. Therefore, in situations where only data on individual wealth is available, we can rewrite the log-likelihood function as

$$\mathcal{L}_N(\boldsymbol{\theta} \mid \mathbf{a}) = \sum_{i=1}^{N} \log g(a_i \mid \boldsymbol{\theta}). \tag{18}$$

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 $\mathcal{A} \times \mathcal{E}$ as described in the Online Appendix. While the efficiency lattice only takes two possible values, $\mathcal{E} = \{e_l, e_h\}$, the wealth lattice is discretized using $I \leq N$ points on a partially ordered set defined by $\mathcal{A} = [\min(\mathbf{a}), \max(\mathbf{a})]$. Once the joint density function of wealth and income has been approximated, the log-likelihood function is constructed in two steps: (i) For each pair $(a_i, e_i) \in \mathbf{a} \times \mathbf{e}$, we use a piece-wise linear interpolation to evaluate $g(a_i, e_i \mid \boldsymbol{\theta})$; (ii) Once $g(a_i, e_i \mid \boldsymbol{\theta})$ has been evaluated for all $(a_i, e_i) \in \mathbf{a} \times \mathbf{e}$, the log-likelihood function is computed using (17)[4].

3. 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 artificially generated data of individual wealth and individual income (labor efficiency). More precisely, and since that $\boldsymbol{\theta}$ shapes the stationary distribution of wealth and income, we are interested in investigating whether the stationary distribution of wealth and income contains relevant information in finite samples about the full set of parameters in the HA model. Similar to many equilibrium models, estimates are often ill-behaved, so we study identification problems in Section 4 and/or assess the finite sample performance of the ML estimator

[4] In terms of implementation, our approach is similar to that proposed in Young (2010) for discrete-time models: (i) no use of simulation-based methods to approximate the model's density function, and (ii) the approximated density resembles a histogram along the state space. However, the weights given to each point in the histogram are different in both approaches. In particular, our method uses the cross-equation restrictions imposed by the equilibrium HJB and Fokker-Planck equations and does not require a dense and uniformly state-space grid to approximate the model's density.

under different subsets of restrictions on the parameters.

The parameter values for the data generating process (DGP), $\boldsymbol{\theta}_0$, are provided in Table 1. In the model, time is measured in years and parameter values should be interpreted accordingly. To ensure the existence of a unique stationary equilibrium we assume an economy with unitary EIS. 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 a monthly average job finding rate of 0.45, and the demotion rate is calibrated to match a monthly average separation rate of 0.034. The endowment level of high efficiency is normalized to one while that of low efficiency is set to one-fifth of the one for employed individuals. These values imply a labor efficiency gap, Δ_e, of 80%, which is consistent with the values used in Huggett (1993), and Imrohoroğlu (1989) and Krusell and Smith (1998). The transition rates for the Poisson processes are computed using (14). The remaining parameter values are standard in the literature, implying the capital-output ratio $K/Y = 2.36$, the interest rates $r = 0.05$, and the aggregate savings rate $(1 - C/Y) = 0.24$.

The Monte Carlo experiment is based on $M = 200$ samples drawn from the model's population stationary joint density function $g\,(a, e \mid \boldsymbol{\theta}_0)$, each of them of size $N \in \{$1,000, 5,000, 10,000$\}$[5]. We first sample the two state labor efficiency units using the marginal stationary distribution in (14). Given the draws on the efficiency units, we then approximate the population density of wealth, $g\,(a|\boldsymbol{\theta}_0)$, using $I = 500$ uniform grid points between $\underline{a} = 0$ and $a_{\max} = 100$, from which we sample values of individual wealth using a slice sampler. For each simulated sample, we proceed to estimate the model's full parameter set using the maximum likelihood estimator using only data on wealth, as well as data on both wealth and income. The numerical maximization of the log-likelihood function is carried out by means of a Global Search algorithm with 250 initial stage points and 500 trial points.

Table 2 summarizes the results. For each parameter $\theta \in \boldsymbol{\theta}$, it reports: (i) the Median Normalized Bias, MNB = median$\{(\hat{\theta}_m - \theta_0)/\theta_0\}$; and (ii) the Mean Absolute Normalized Errors, MANE = $(1/M)\sum_m |(\hat{\theta}_m - \theta_0)/\theta_0|$. We report normalized metrics to avoid scale

[5]Each Monte Carlo experiment takes up to 28 hours on a dedicated 32 cores Xeon server.

TABLE 1

Population parameters, $\boldsymbol{\theta}_0$. The exogenous endowment of efficiency units is given by $\mathrm{d}e = -\Delta_e \mathrm{d}q_1 + \Delta_e \mathrm{d}q_2$*, with $\Delta_e \equiv e_h - e_l$, where q_1 and q_2 are Poisson processes with intensity rates ϕ_{lh} and ϕ_{hl} respectively. The representative household has standard preferences defined by $U_t = \mathbb{E}_t\left[\int_t^\infty e^{\rho(s-t)} u(c)\,\mathrm{d}s\right]$ where $u(c) = c^{1-\gamma}/(1-\gamma)$. The macroeconomic identity in the stationary competitive equilibrium is given by $Y = C + \delta K$, where $Y = K^\alpha L^{1-\alpha}$. In the model, time is measured in years and parameter values should be interpreted accordingly.*

Parameter	*Value*
Relative risk aversion, γ	1.0000
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

problem when comparing estimates across parameters. Panel A present results when the only data used in the estimation is individual wealth, while Panel B reports the results when using data on both individual wealth and income.[6]

The simulation results reveal that some of the model parameters exhibit large biases that persist even as the sample size increases. This includes the coefficient of relative risk aversion, γ, the levels of labor efficiency units, $\{e_l, e_h\}$, and their transition rates, $\{\phi_{lh}, \phi_{hl}\}$. In particular, the ML estimates for these parameters imply extremely risk averse households as well as income levels and transition rates that are larger than their population values. For example, the average estimate of γ is between 2 to 7 times larger than the true value in the population when only data on wealth is used. Similarly, the average estimate for e_l is 4 times larger than in the population for a sample of wealth and income with N =1,000 observations, 2 times larger for N =5,000, and just below 1 for N =10,000. Moreover, the differences between the (absolute value of) mean and median biases suggest that the small sample parameter distributions are skewed. This is particularly the case when both wealth and income data are used in the estimation. On

[6]Given the two-state nature of income levels in the model we do not attempt to estimate the parameters using only data on individual income. In this case, the distribution of income alone will not be informative about the persistence and transitions of income shocks, $(e_h, e_l, \phi_{lh}, \phi_{hl})$. However, as shown in Section 4 the distribution of wealth provides information on the persistence parameters. Moreover, the use of panel data or repeated cross-sections on income could improve the estimation of the income parameters (see, e.g., Papp and Reiter, 2020 and Liu and Plagborg-Møller, 2021).

TABLE 2

Finite sample properties of the ML estimator. For each $\theta \in \boldsymbol{\theta}$, the table reports the Median Normalized Bias (MNB) and the Mean Absolute Normalized Errors (MANE) from a Monte Carlo experiment using $M = 200$ samples, each of size $N = \{1{,}000, 5{,}000, 10{,}000\}$. Results in Panel A are based on cross-sectional data on wealth only, while those in Panel B are based on cross-sectional data on both wealth and income.

	MNB	*MANE*	*MNB*	*MANE*	*MNB*	*MANE*
$\boldsymbol{\theta}$	N=1,000		N=5,000		N=10,000	
	Panel A: Wealth only					
γ	7.139	6.374	2.399	3.739	3.165	3.754
ρ	0.056	0.506	-0.058	0.383	0.058	0.303
α	-0.218	0.225	-0.243	0.239	-0.205	0.192
δ	0.118	0.212	0.114	0.189	0.083	0.180
e_l	4.689	4.337	4.126	3.723	3.651	3.299
e_h	0.892	0.867	0.926	0.890	0.715	0.719
ϕ_{lh}	0.071	0.468	-0.093	0.351	0.030	0.282
ϕ_{hl}	0.457	2.004	-0.061	1.547	0.257	1.039
	Panel B: Wealth and income					
γ	4.028	4.797	0.084	2.297	0.057	1.933
ρ	0.168	0.431	-0.011	0.312	0.223	0.308
α	-0.205	0.215	-0.229	0.221	-0.138	0.155
δ	0.125	0.200	0.135	0.194	0.091	0.201
e_l	4.378	4.153	2.016	2.905	0.650	2.114
e_h	0.940	0.848	0.898	0.858	0.567	0.618
ϕ_{lh}	-0.070	0.417	0.057	0.389	0.120	0.351
ϕ_{hl}	-0.070	0.449	0.046	0.398	0.139	0.355

the other hand, the biases on the discount rate, ρ, the capital share in production, α, and on the depreciation rate, δ, are within reasonable ranges, even in small samples[7].

In general, increasing the sample size reduces the estimated bias for most of the parameters. Panel A shows that augmenting the sample size from N=1,000 to N=5,000 observations reduces the MANE of largely biased parameters by around one to three orders of magnitude. For example, the MANE for γ falls around 70.5% from 6.374 to 3.739, while the MANE for e_l and ϕ_{hl} drop 16.5% and 29.5% from 4.337 to 3.723, and from 2.004 to 1.547, respectively. On the other hand, Panel B suggests that using data on both individual wealth and individual income generally delivers ML estimates with smaller biases, and relatively larger bias reductions as a function of the sample size. This is particularly evident again for γ, e_l and ϕ_{hl}. However, the additional information brought by the use

[7]We do not propose to estimate the capital share and/or the depreciation rate from the wealth data alone rather than from NIPA data. As illustrated by Figures 1 and 2 below even slight variations within conventional calibrated values of these parameters can have great impact on the wealth distribution.

of income data has limited effects on the estimation errors and biases for ρ, α and δ.

4. The Kullback-Leibler divergence

Overall, the finite sample results for the unrestricted ML estimator reveal substantial differences among parameter estimates that could suggest potential identification problems, particularly for those that exhibit considerable normalized errors. To investigate this possibility we take one step back and look at the model's implied *population distribution function*, $G_0 \equiv G(a, e \mid \boldsymbol{\theta}_0)$, and its associated *population density function*, $g_0 \equiv g(a, e \mid \boldsymbol{\theta}_0)$. Since the model's probability density function constitutes the building block of the maximum likelihood estimator in (17), examining its behavior will provide valuable information on whether it is possible to achieve identification of the model parameters using the likelihood of the data. In particular, we are interested in studying the sensitivity of the population distribution to small perturbations in the values of the model's structural parameters.

We propose to use the Kullback-Leibler (KL) divergence, or relative entropy, to measure the divergence between any two distributions (see Kullback and Leibler, 1951 and Kullback, 1959). A similar procedure was proposed in Qu and Tkachenko (2017), where the distance between any two spectral densities is used to study identification in the class of linearized representative agent DSGE models. Let $\widetilde{G} \equiv G(a, e \mid \boldsymbol{\theta})$ and $\widetilde{g} \equiv g(a, e \mid \boldsymbol{\theta})$ denote the model's implied wealth-income distribution and density functions for $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$. Then, the KL divergence from $\widetilde{G}$ to G_0 is defined as

$$\mathcal{D}_{KL}\left(G_0 \,||\, \widetilde{G}\right) = \sum_{e \in \mathcal{E}} \int_{a \in \mathcal{A}} g(a, e \mid \boldsymbol{\theta}_0) \log\left(\frac{g(a, e \mid \boldsymbol{\theta}_0)}{g(a, e \mid \boldsymbol{\theta})}\right) \mathrm{d}a.$$

The value of the KL divergence, $k \geq 0$, measures the information differences between the two distributions G_0 and $\widetilde{G}$. If $k = 0$, then if follows that $G_0 = \widetilde{G}$ almost everywhere in $\mathcal{A} \times \mathcal{E}$, despite the fact that $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$. For $k > 0$, however, the KL divergence does not help to assess whether the difference between the two distributions is large or small along $\mathcal{A} \times \mathcal{E}$. Following McCulloch (1989), we therefore map the KL divergence between wealth-income distributions to the KL divergence between the two Bernoulli distributions, $B(0.5)$ and $B(q)$, where the implicit probability q is chosen in such a way that $\mathcal{D}_{KL}(B(0.5) \,||\, B(q)) = \mathcal{D}_{KL}(G_0 \,||\, \widetilde{G})$[8]. As an example, suppose that the probability implied by the two distributions G_0 and $\widetilde{G}$ is $q = 0.51$. This corresponds to assigning a fair coin toss a probability of 0.51 when the true probability is 0.5. Interestingly, Akaike (1973) and White (1982) have shown that $\min_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \mathcal{D}_{KL}(G_0 \,||\, \widetilde{G}) = \max_{\boldsymbol{\theta}\in\boldsymbol{\Theta}} \mathcal{L}_N$ as $N \to \infty$. Therefore, from an asymptotic perspective, the KL divergence can be used as a device to explore the behavior of the log-likelihood function around the $\boldsymbol{\theta}_0$, and hence, to inform on the ability of the ML estimator to identify the model parameters.

Figure 1 plots the probability q implied by the KL divergence from $\tilde{G}$ to G as we vary each $\theta \in \boldsymbol{\theta}$ while keeping the remaining parameters at their population values. All the KL divergences, apart from that for γ, are constructed using parameter values that are 50% below and 50% above of the true parameter value. In the case of γ, we employ values that lie 100% below and 100% above its population value. A dotted vertical line denotes the value in the DGP. The results suggest that all things equal, small perturbations to ρ, α and δ have a large impact on the shape of the joint distribution of wealth and income. Therefore, conditional on a given sample being observed, the ML estimator should be able to identify this subset of parameters given that small differences in their values will produce significantly different density functions. On the contrary, the influence of γ and some of the income process parameters is small which suggest that the likelihood surface will be flatter along these dimensions of the parameter space reducing the ability of the ML estimator to identify them from a given sample. This lack of curvature could explain the poor performance of the unrestricted ML estimator along these dimensions of the parameter space described in Section 3. To further exemplify the previous argument, Figure 2 plots the model's implied density of wealth for different values of the coefficient of relative risk aversion and the share of capital in output. As suggested by the KL divergence, small perturbations in γ have virtually no effect on the wealth distribution,

[8] McCulloch (1989) shows that the KL divergence from $B(q)$ to $B(0.5)$ is given by $\mathcal{D}_{KL}(B(0.5) \,||\, B(q)) = -\log(4q(1-q))/2$. Thus, q measures the divergence of an arbitrary Bernoulli trial from a fair Bernoulli trial. Given the KL divergence k from $\widetilde{G}$ to G_0, it is straightforward to compute the implied probability q.

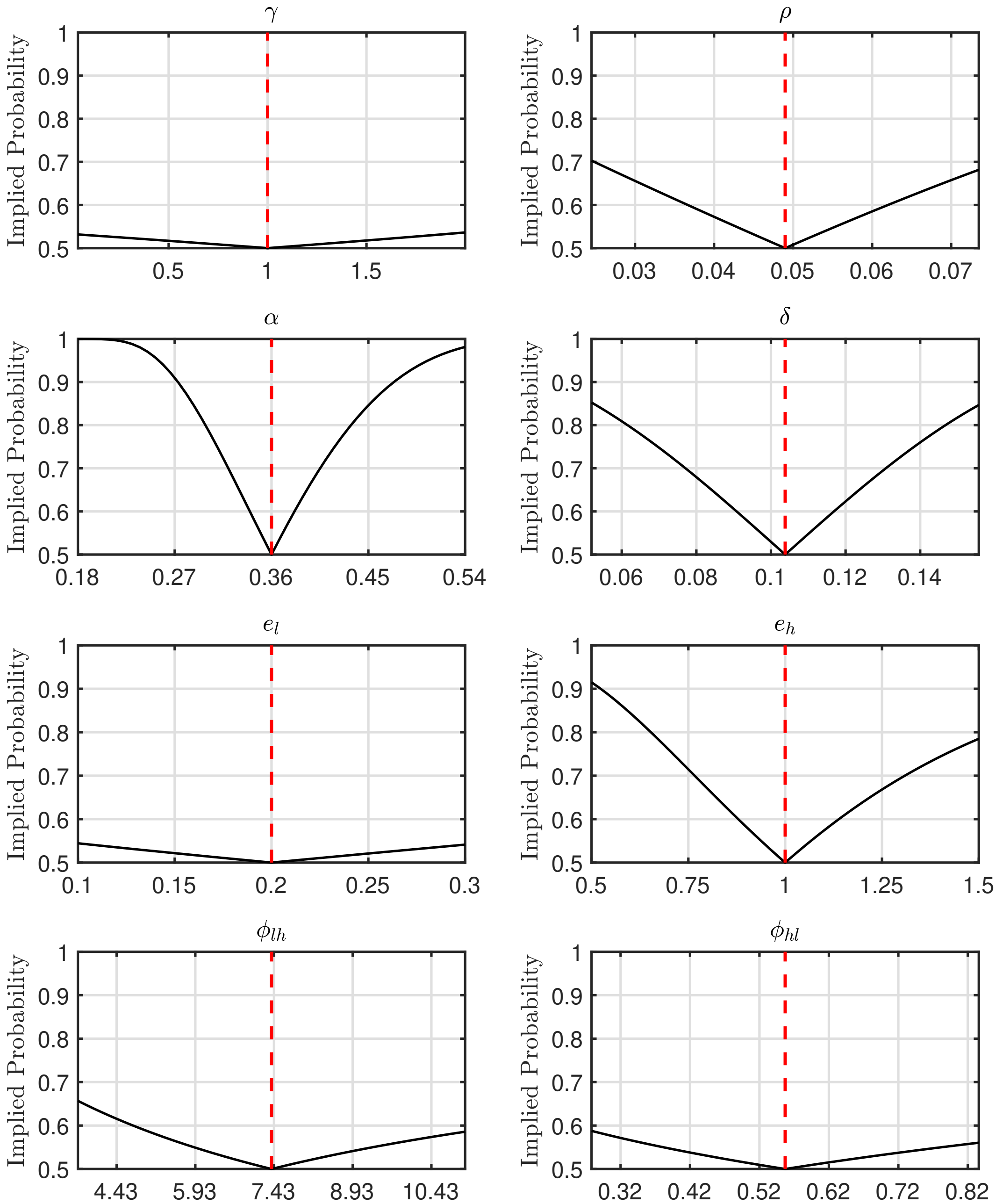

Figure 1. Kullback-Leibler divergence. The graph plots the implied probability q associated with the KL divergence from $G(a \mid \boldsymbol{\theta})$ to $G(a \mid \boldsymbol{\theta}_0)$ that results from varying each parameter at a time while keeping the remaining ones at their population value. The vertical dashed line denotes the true parameter value.

whereas small changes in α lead to substantial differences in the wealth distribution[9].

[9]The sensitivity of the wealth distribution to changes in the remaining parameters of the model can be found in the Online Appendix.

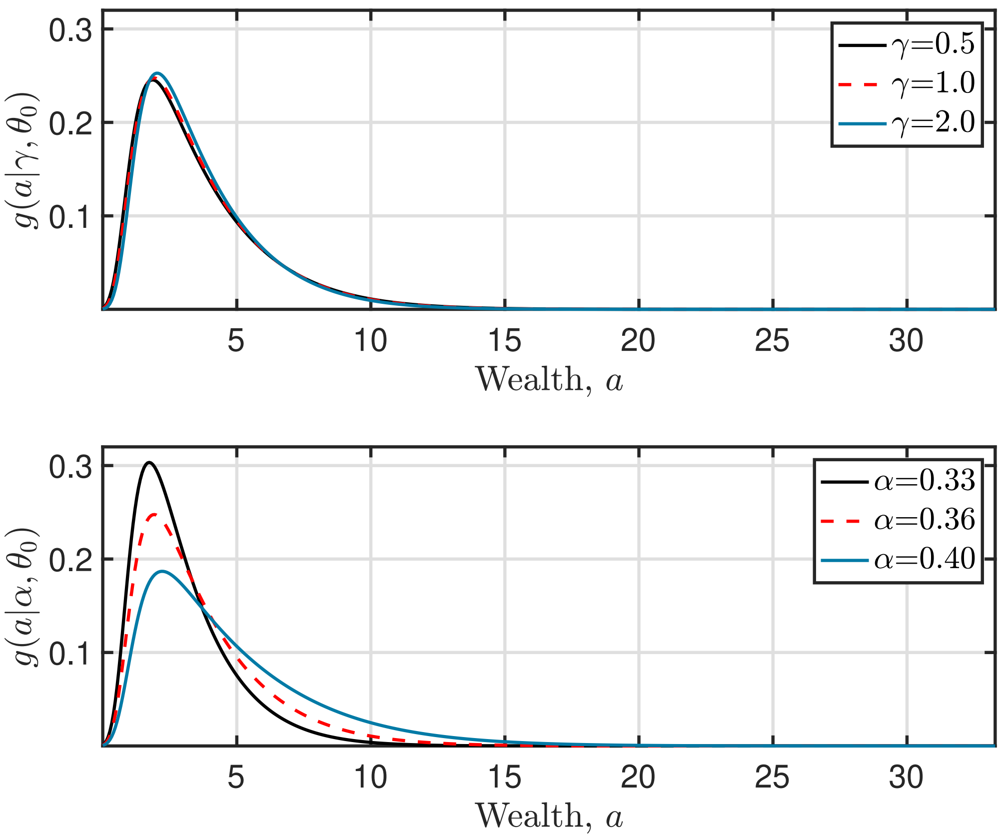

Figure 2. Sensitivity of the wealth distribution. The graph shows the sensitivity of the distribution of wealth, $g(a|\boldsymbol{\theta})$, for selected parameters. The dashed line denotes the population density of wealth. The continuous lines correspond to the density of wealth resulting from small perturbations in each parameter while keeping the remaining ones at their population value.

5. Calibration and estimation

Although the results from the previous section indicate that the parameter estimates approach their true values in the population as the sample size increases, they also suggest that the identification power of the likelihood function of wealth and income is reduced in some dimensions of the parameter space, particularly in small samples. For the prototype economy of Section 2, these inaccuracies are reflected in poor estimates of the parameters related to the exogenous income process and of the coefficient of relative risk aversion. A common practice among economists to get around this obstacle is to calibrate the parameters that are problematic and estimate the remaining ones.

To assess the consequences of following such a strategy, we investigate the finite sample behavior of the ML estimator when different subset of parameters are externally calibrated. We begin by considering the case where only γ and ρ are estimated. Then, we move on to a case where we also estimate α and δ. In both cases, the exogenous income process is fixed, a strategy that closely resembles the standard practice followed in the heterogeneous agent literature (cf. Benhabib et al., 2019, Abbott et al., 2019 and Luo and Mongey, 2019). Accordingly, Table 3 summarizes the results from a set of Monte Carlo experiments where we analyze the properties of the ML estimator conditional on the calibrated values of the corresponding structural parameters of the model. Each Monte Carlo simulation is based on $M = 200$ samples generated from the model's population stationary probability density function, each of them of size $N =$5,000.

We report the mean absolute normalized errors, MANE, the median normalized bias, MNB, and the implied aggregate capital-output ratio (K/Y), interest rate, and aggregate savings rate across simulations. Panel A reports the outcomes when only data on individual wealth is used in the estimation and Panel B reports the results when data on both individual wealth and income is used. The last column from each panel replicates the unrestricted ML estimation from Section 3 for comparison.

The results in Table 3 show that when we calibrate all the income parameters simultaneously, the ML estimator yields sharper estimates. In particular, we obtain smaller biases in the estimation of the relative risk aversion, the discount rate and the capital share in output, and a small deterioration in the precision with which we can estimate the depreciation rate. Considerable improvements, relative to the unrestricted estimation, are also obtained if we additionally calibrate the capital share in output and the depreciation rate. In general, the size of the finite sample biases are positively related to the number of parameters being estimated in that fixing a larger subset of parameters to their true values in the population delivers more precise parameter estimates. This directly translates to the model implied macroeconomic aggregates. In particular, notice that the capital-output ratio, the interest rate, and the savings rate are estimated with the most accuracy when only two parameters are estimated. These results hold across the different data sets used in the estimation.

Next, we study the finite sample properties of the ML estimator when the income process is only partially calibrated. We run Monte Carlo experiments conditional on the in-

TABLE 3

Conditional estimates: subset of parameters, including the income process, fixed to their population values. The table reports the Mean Absolute Normalized Error (MANE) and the Median Normalized Bias (MNB, in parenthesis) from a Monte Carlo experiment with $M = 200$ *samples, each of them of size* $N = 5{,}000$.

	Panel A: Wealth only			*Panel B: Wealth and income*		
$\boldsymbol{\theta}$	*Restricted*		*Unrestricted*	*Restricted*		*Unrestricted*
γ	0.8732 (0.8732)	0.487 (−0.100)	3.739 (2.399)	0.8703 (0.8703)	0.489 (−0.073)	2.297 (0.084)
ρ	0.0180 (0.0009)	0.341 (−0.354)	0.383 (−0.058)	0.0180 (0.0009)	0.343 (−0.357)	0.312 (−0.011)
α	α_0	0.173 (−0.153)	0.239 (−0.243)	α_0	0.185 (−0.147)	0.221 (−0.229)
δ	δ_0	0.318 (−0.130)	0.189 (0.114)	δ_0	0.345 (−0.138)	0.194 (0.135)
e_l	$e_{l,0}$	$e_{l,0}$	3.723 (4.126)	$e_{l,0}$	$e_{l,0}$	2.905 (2.016)
e_h	$e_{h,0}$	$e_{h,0}$	0.890 (0.926)	$e_{h,0}$	$e_{h,0}$	0.858 (0.898)
ϕ_{lh}	$\phi_{lh,0}$	$\phi_{lh,0}$	0.351 (−0.093)	$\phi_{lh,0}$	$\phi_{lh,0}$	0.389 (0.057)
ϕ_{hl}	$\phi_{hl,0}$	$\phi_{hl,0}$	1.547 (−0.061)	$\phi_{hl,0}$	$\phi_{hl,0}$	0.398 (0.046)
K/Y	0.0063 (0.0052)	0.0893 (0.0832)	0.3731 (−0.3114)	0.0063 (0.0051)	0.0942 (0.0777)	0.2859 (−0.3002)
Interest rate	0.0194 (−0.0160)	0.3406 (−0.3546)	0.5066 (0.0533)	0.0194 (−0.0159)	0.3430 (−0.3565)	0.3120 (−0.0132)
Savings rate	0.0063 (0.0052)	0.2437 (−0.0642)	0.2217 (−0.2152)	0.0063 (0.0051)	0.2601 (−0.0628)	0.2197 (−0.2246)

come levels, e_l and e_h, being fixed to their values in the population, and conditional on the transition rates, ϕ_{lh} and ϕ_{hl}, also fixed to their population values. The results are summarized in Table 4. In general, we find that calibrating the levels of the income process alone reduces both the absolute errors and the estimates' biases considerably. This is particularly the case for the coefficient of γ and α, and to a lesser extent, for the transition rates. Calibrating the transition rates of the income process alone has limited consequences. Although it helps to reduce the bias in γ by nearly half of that obtained from an unconstrained estimation, there is virtually no effect on the errors of the remaining parameters. Interestingly, we find that the use of income data, in addition to wealth data, does not provide any additional information that can help to identify the income levels. This is most likely due to the fact that the income data in this model is represented by a binary variable and hence it only contains information about the stationary probabilities of states.

To further understand the effects of calibrating the income process, either partially or

TABLE 4

Conditional estimates: income process fixed to their population values. The table reports the Mean Absolute Normalized Error (MANE) and the Median Normalized Bias (MNB, in parenthesis) from a Monte Carlo experiment with $M = 200$ samples, each of them of size $N = 5{,}000$.

	Panel A: Wealth only				*Panel B: Wealth and income*			
$\boldsymbol{\theta}$	*Income levels*	*Transition rates*	*All income*	*Unrestr.*	*Income levels*	*Transition rates*	*All income*	*Unrestr.*
γ	0.550 (−0.029)	1.518 (0.867)	0.487 (−0.100)	3.739 (2.399)	0.436 (−0.070)	1.423 (0.476)	0.489 (−0.073)	2.297 (0.084)
ρ	0.322 (−0.197)	0.403 (0.176)	0.341 (−0.354)	0.383 (−0.058)	0.284 (−0.2551)	0.403 (0.079)	0.343 (−0.357)	0.312 (−0.011)
α	0.080 (−0.027)	0.279 (−0.263)	0.173 (−0.153)	0.239 (−0.243)	0.067 (−0.033)	0.278 (−0.258)	0.185 (−0.147)	0.221 (−0.229)
δ	0.145 (0.004)	0.255 (0.060)	0.318 (−0.130)	0.189 (0.114)	0.124 (0.004)	0.277 (−0.031)	0.345 (−0.138)	0.194 (0.135)
e_l	$e_{l,0}$	3.476 (3.527)	$e_{l,0}$	3.723 (4.126)	$e_{l,0}$	3.040 (2.001)	$e_{l,0}$	2.905 (2.016)
e_h	$e_{h,0}$	1.068 (1.203)	$e_{h,0}$	0.890 (0.926)	$e_{h,0}$	0.918 (0.775)	$e_{h,0}$	0.858 (0.898)
ϕ_{lh}	0.350 (−0.127)	$\phi_{lh,0}$	$\phi_{lh,0}$	0.351 (−0.093)	0.211 (−0.092)	$\phi_{lh,0}$	$\phi_{lh,0}$	0.389 (0.057)
ϕ_{hl}	0.718 (−0.134)	$\phi_{hl,0}$	$\phi_{hl,0}$	1.547 (−0.061)	0.220 (−0.085)	$\phi_{hl,0}$	$\phi_{hl,0}$	0.398 (0.046)
K/Y	0.0459 (0.0158)	0.3626 (−0.3650)	0.0893 (0.0832)	0.3731 (−0.3114)	0.0358 (0.0211)	0.3308 (−0.2734)	0.0942 (0.0777)	0.2859 (−0.3002)
Interest rate	0.3234 (−0.2023)	0.4002 (0.1697)	0.3406 (−0.3546)	0.5066 (0.0533)	0.2841 (−0.2553)	0.4005 (0.0766)	0.3430 (−0.3565)	0.3120 (−0.0132)
Savings rate	0.1379 (0.0288)	0.3381 (−0.2819)	0.2437 (−0.0642)	0.2217 (−0.2152)	0.1149 (0.0293)	0.3452 (−0.2860)	0.2601 (−0.0628)	0.2197 (−0.2246)

completely, Figure 3 plots the finite sample distribution of parameter estimates when we use data on wealth and income in the estimation. A dotted vertical line represents the true parameter values. The figure confirms not only that a strategy based on calibrating the income levels delivers relatively unbiased parameter estimates, but also more precise estimates, as measured by the dispersion around the mean estimates. On the contrary, the simultaneous calibration of all income parameters produces sharp estimates of γ and ρ but it also generates distributions that exhibit multiple modes for α and δ that could suggest further identification problems. To investigate this claim we compute the KL divergence from $\tilde{G}$ to G_0, and the associated implied probability q, that results from varying both α and δ simultaneously while keeping the remaining parameters at their population values. Figure 4 plots the contour for $q(\alpha, \delta)$ which shows the presence of a ridge in the $\alpha - \delta$ space. In other words, a proportional increase in both parameters produces almost obser-

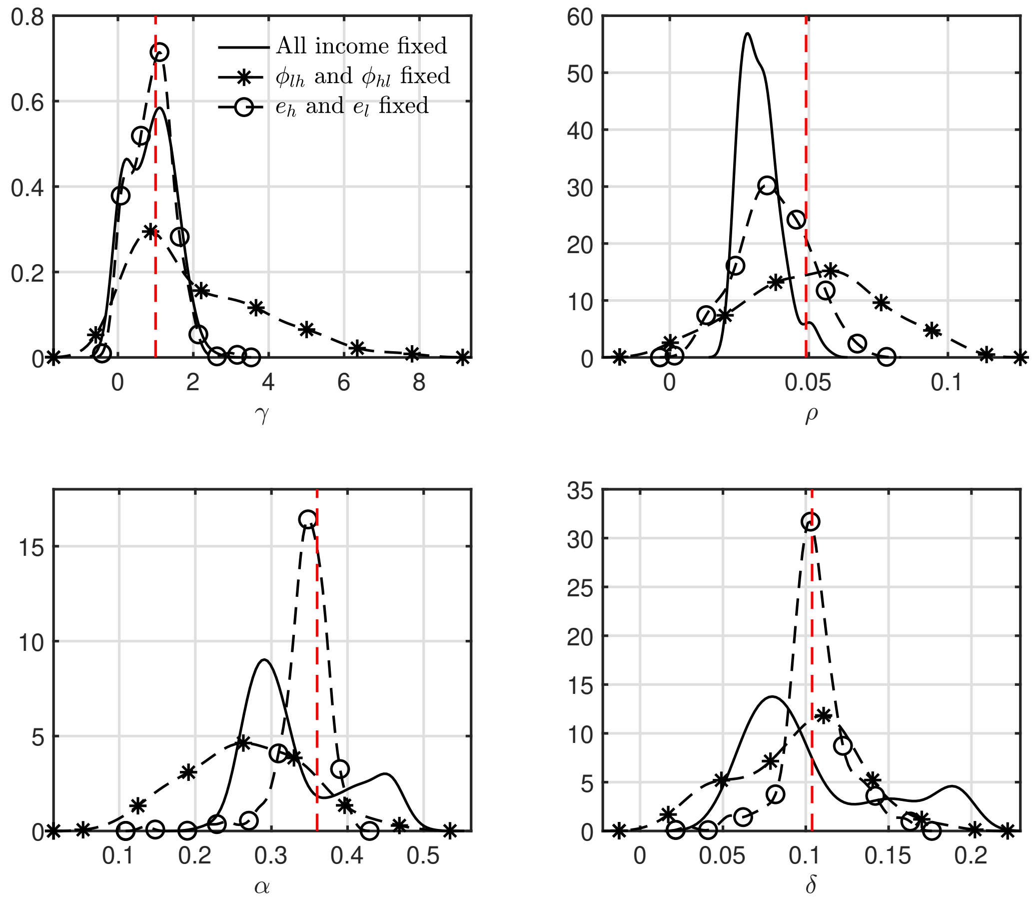

Figure 3. Finite sample distribution of parameter estimates. The graph plots the kernel density of estimated parameters across $M = 200$ random samples of size $N = 5{,}000$ generated from the true data generating process. The estimation uses data on individual wealth and income. The vertical line denotes the true parameter value.

vationally equivalent distribution functions, and therefore partial identification problems.

The tight relation between α and δ is an example of identification deficiencies that are rooted in the economic theory and could persist even in samples of finite size. As an example, consider the steady state capital-output ratio from the standard neoclassical growth model, K/Y. 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 4. The Monte Carlo evidence suggests that the multimodality in these two parameters, with their corresponding implications for identification, can be alleviated by calibrating the income process only partially. As

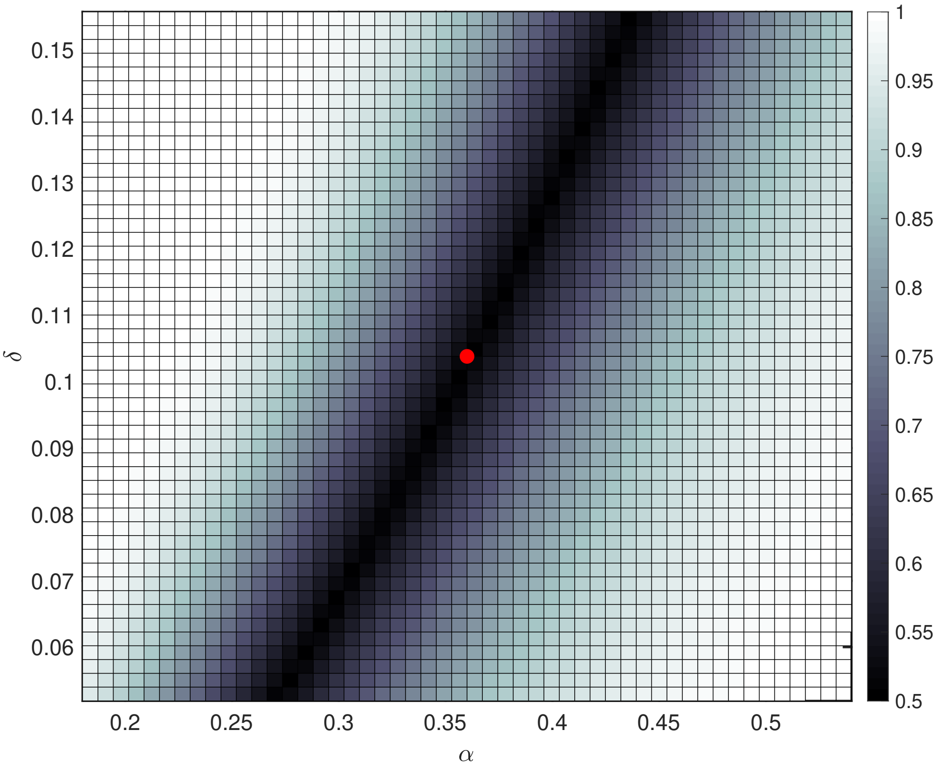

Figure 4. Kullback-Leibler divergence. The graph plots the contour of the implied probability q associated with the bivariate KL divergence that results from varying α and δ simultaneously while keeping the remaining parameters at their population value. The dot denotes the true parameter value.

shown in Figure 3 the calibration of the income levels or of the transition rates yields distribution of estimates for α and δ that are unimodal and at the same time do not affect the accuracy with which γ can be identified. Hence, allowing α and δ to interact with some of the income parameters during the estimation process provides a better identification. Note, however, that fixing the values for the income levels provide the best results in terms of bias reduction, correct identification, and reduced variability of parameter estimates.

In general, our calibration experiments point towards a strategy based on calibrating parameters that are weakly identified, as indicated by the KL divergence. This includes the income levels or the coefficient of relative risk aversion. However, this approach may not carry any improvement in the identification and estimation accuracy of the ML estimator if the calibrated values happen to be different from those in the population. Similar concerns have been raised previously in the context of linearized representative agent models (see Canova and Sala, 2009). Therefore, we investigate if our previous results are sensitive to mis-calibration. In particular, we consider the effects of calibrating

TABLE 5

Conditional estimates: alternative data generating process. The table reports the Mean Absolute Normalized Error (MANE) and the Median Normalized Bias (MNB, in parenthesis) from a Monte Carlo experiment with $M = 200$ samples, each of size $N = 5{,}000$, generated under alternative data generating process. In particular, Δ_e larger uses $e_h = 1.5$ and $e_l = 0.1$; Δ_e smaller uses $e_h = 0.5$ and $e_l = 0.1$; γ higher uses $\gamma = 2.0$; and γ lower uses $\gamma = 0.5$. The ML estimation uses data on individual wealth and income.

	Wealth and income					
	Δ_e *larger*	Δ_e *smaller*	$\Delta_e = 0.8$	γ *higher*	γ *lower*	$\gamma = 1.0$
K/Y	0.2434 (0.236)	0.3732 (−0.378)	0.0358 (0.021)	0.2888 (−0.283)	0.2631 (−0.275)	0.3081 (−0.299)
Interest rate	0.3008 (−0.278)	0.9752 (1.002)	0.2841 (−0.255)	0.5798 (0.604)	0.4342 (0.462)	0.3740 (0.169)
Savings rate	0.1717 (0.134)	0.1410 (0.118)	0.1149 (0.029)	0.1394 (−0.055)	0.1916 (−0.194)	0.1969 (−0.200)
Gini Coeff.	0.0550 (−0.055)	0.1204 (0.123)	0.0398 (−0.040)	0.0403 (−0.040)	0.0206 (0.021)	0.0396 (−0.039)
Bottom 50%	0.0507 (0.051)	0.1028 (−0.105)	0.0387 (0.038)	0.0368 (0.036)	0.0187 (−0.019)	0.0386 (0.038)
Top 10%	0.0390 (−0.040)	0.1088 (0.111)	0.0248 (−0.025)	0.0288 (−0.028)	0.0153 (0.015)	0.0243 (−0.025)

the risk aversion coefficient, γ, and the labor efficiency gap, Δe, to the values in Table 1 when in reality the true DGP is characterized by higher or lower values.

Table 5 reports the results from a Monte Carlo simulation with $M = 200$ samples of wealth and income, each of size $N = 5{,}000$. Due to the non-linear dependences among all structural parameters of the model, we report the MANE and MNB for some of the key macroeconomic statistics implied by the model. In particular, we analyze the effects of a higher and lower income level gap and relative risk aversion on the steady state levels of the capital-output ratio, K/Y, of the interest rate, r, of the aggregate savings rate, $(1 - C/Y)$, of the Gini coefficient, and of the Lorenz curve[10]. For comparison, the table also reports the case where Δ_e and γ are calibrated to the values in Table 1.

The mis-calibration of the income levels can have a substantial impact on the accuracy with which the implied aggregate statistics can be estimated, and thus lead to wrong inferences. In particular, we find that a calibrated income gap that is higher than its value in the population leads to considerable biases in the steady state capital-output ratio, the steady state interest rate and the overall implied wealth distribution. On the

[10]Similar statistics for all the Monte Carlo simulations described in the paper are available upon request.

other hand, the mis-calibration of the coefficient of relative risk aversion has a negligible effect on the implied macroeconomic quantities. This result is consistent with the fact that the shape of the wealth distribution is not sensitive to changes in γ, as suggested by the KL divergence, and documented in Figure 1.

In summary, our Monte Carlo evidence suggests that ρ, α and δ can be identified and accurately estimated with the use of cross-sectional data on individual wealth and income by means of our proposed ML estimator. On the other hand, the coefficient of relative risk aversion and the parameters describing the exogenous income process display some identification challenges that may lead to inferential problems that persist even in large samples. Following standard practice in macroeconomics, we find that a mixed strategy where a subset of the troublesome parameters is calibrated provides a considerable improvement in terms of statistical precision without affecting the overall results. Given the inherent uncertainty around the correct parameter values to use in the calibration, the results suggest that fixing the value of the relative risk aversion, and not the income levels, provides the best finite sample performance of the ML estimator.

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 for the U.S. economy using the wealth and wage income data reported in the Survey of Consumer Finances (SCF) for the year 2013.

To accommodate the high degree of wealth inequality observed in the data, we expand the number of labor efficiency states in the prototype economy to four so that the endowment process now follows a continuous-time Markov chain with state space $\mathcal{E} = \{e_1, e_2, e_3, e_4\}$, with $e_1 < e_2 < e_3 < e_4$, that evolves over time according to

$$\mathrm{d}e_t = \sum_i \sum_{j \neq i} (e_i - e_j)\, \mathrm{d}q_{ij,t}, \quad e_0 \in \mathcal{E}.$$

The Poisson processes $q_{ij,t}$ for all $i, j = 1, \ldots 4$ and $i \neq j$ count the frequency with which an agent moves from state i to state j. Associated with each efficiency level, we define $\phi_{ij} \geq 0$ to be the instantaneous transition rate from state i to state j. Since individuals cannot transit to state i while currently being in the same state, it follows that $\phi_{ii} = 0$, for all $i = 1, \ldots, 4$.

The estimation sample includes data on households with positive net worth and positive income (per hundred thousand) in order to be consistent with the model's non-negative borrowing constraint. The wealth data corresponds to the net-worth reported in the Summary Extract Public Data provided by the SCF. To obtain an equally weighted sample of household wealth, we resample the net-worth data using the weights provided by the SCF. The wage income data is recoded into four discrete states, where each state corresponds to income levels belonging to one of the following pre-defined quantile bins: 0-25, 25-50, 50-99, 99-100. The unequal spacing of the quantile bins tries to accommodate the high degree of income inequality in the data[11]. The final sample includes N =18,631 individuals.

The model's solution is approximated on a grid for wealth containing $I = 500$ equally spaced points. The resulting (negative) log-likelihood function is then minimized using a GlobalSearch algorithm with 1000 random trial points. We use a non-parametric bootstrap to compute confidence intervals for the parameter estimates using $M = 100$ bootstrap samples[12]. Following the Monte Carlo evidence of Section 5, we do not attempt to estimate the coefficient of relative risk aversion, γ. Instead, we calibrate it to 1.0 and estimate all the remaining parameters. Alternative calibrations result in lower log-likelihood values.

Table 6 reports the maximum likelihood estimates together with their 95% confidence intervals. In Panel A we present the results for the preference parameters, in Panel B for the income or labor efficiency levels, and in Panel C for the intensity rates associated with each of the count processes that describe the idiosyncratic income dynamics in the economy. Panel D reports the corresponding limiting distribution defined

[11]Using equally spaced quantile bins will produce much smaller overall estimates of the actual wealth-income inequality.

[12]The bootstrapping exercise is computationally demanding. Estimation with $M = 100$ samples takes about 58 hours on a dedicated 32 cores Xeon server.

TABLE 6

Maximum likelihood estimates. The table reports the maximum likelihood estimates (MLE) of the model parameters and their 95% confidence intervals computed from a non-parametric bootstrap with $M = 100$ samples. The estimation sample contains N =18,631 observations on individual wealth and income. The coefficient of relative risk aversion is calibrated to $\gamma = 1.0$.

Panel A: Preference parameters				
Parameter	γ	ρ	α	δ
Value	1.0	0.1120 [0.1077, 0.1138]	0.5601 [0.5507, 0.5624]	0.0306 [0.0256, 0.0349]

Panel B: Income levels, e_i				
Parameter	e_1	e_2	e_3	e_4
Value	0.1130 [0.1129, 0.1329]	0.1164 [0.1161, 0.1371]	0.2586 [0.2450, 0.2653]	5.2823 [4.9703, 5.4561]

Panel C: Intensity rates, ϕ_{ij} ($\times$ 100)				
$i \backslash j$	1	2	3	4
1	0	0.0006 [0.0005, 0.0008]	0.2275 [0.2119, 0.2480]	0.3055 [0.2334, 0.3073]
2	0.1427 [0.1089, 0.1503]	0	0.0001 [0.0001, 0.0002]	0.2128 [0.1859, 0.2452]
3	0.2045 [0.1686, 0.2124]	0.1842 [0.1653, 0.1972]	0	0.0000 [0.0000, 0.0000]
4	0.0012 [0.0010, 0.0016]	0.0013 [0.0008, 0.0015]	38.9288 [37.8903, 43.2152]	0

Panel D: Stationary probabilities (%)				
$p(e_i)$	25.50 [24.78, 26.15]	25.34 [24.76, 26.10]	48.83 [47.94, 49.26]	0.34 [0.28, 0.34]

as $p(e_i) \equiv \lim_{t\to\infty} p(e_i, t)$, where $p(e_i, t)$ denotes the unconditional probability of being in state e_i at time t. Our estimates capture a considerable and persistent degree of income inequality as suggested by the extreme estimate for e_4 in Panel C which is nearly 50 times the average income of the least efficient individual. It also suggests that the most productive households are about 20 times more productive than the second most productive households. With respect to the preference parameters in Panel A, we find that while the estimates for the discount rate and the capital share of output are somewhat above the values usually reported in the literature, the estimate for the depreciation rate is below. The estimates of the stationary probabilities closely match the allocation of households into the different income bins: the mass of agents in the first and second income/efficiency level is about 25%, the mass in the third level close to 50%, and finally, the mass of agents with extremely high-income levels does not exceed 1%

Finally, Table 7 compares some wealth statistics and macroeconomic aggregates to

TABLE 7

Wealth inequality and macroeconomic aggregates: data vs. model. The table reports the observed and estimated Gini coefficient, the distribution of wealth across top percentiles, the capital-output ratio, the interest rate, and the savings rate. It also reports 95% confidence intervals computed from a non-parametric bootstrap.

	Gini Coefficient	*% wealth in top*			*Aggregates*		
		5%	*10%*	*20%*	K/Y	Interest rate	Savings rate
Data	0.8048	57.73	70.27	83.44	3.1	0.08	0.089
Model	0.7912 [0.7817, 0.7939]	44.30 [42.98, 44.76]	63.82 [62.27, 64.31]	83.57 [82.27, 83.91]	4.1364 [4.0193, 4.2013]	0.1048 [0.1012, 0.1069]	0.1267 [0.1021, 0.1419]

those implied by the estimated model. It reports the Gini coefficient, the percentage of total wealth held by the top 5, 10 and 20 percentiles, the capital-output ratio, the interest rate, and the savings rate. The observed values for the wealth statistics are computed directly from the SCF data used in the estimation. The values for the capital-output ratio and the real interest rate are those reported in Barro (2021), while the aggregate savings rate corresponds to the historical average between 1959 and 2022 of the personal savings rate.

The estimated model can match the data 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". In particular, a high degree of income inequality is needed for the prototype model to generate a skewed wealth distribution. It should be stressed that we don't see our results as evidence for this particular income process. As pointed out in Benhabib and Bisin (2018), the implication of the "awesome state" in the labor income process is very likely to be a counterfactual to the actual income data. The estimated income process simply captures all other relevant wealth inequality driving forces (bequest, entrepreneur risk, explosive wealth accumulation, etc.) that are not present in our simple model. The estimated parameters of the labor efficiency process are mostly likely the key driving force for the high degree of the wealth inequality. Moreover, it is important to realize that the ability of the ML estimator to match the empirical data on wealth and income (or some of its moments) should not come as a surprise due to the one-to-one mapping between the model's likelihood function and the model's approximated joint probability density function of wealth and income. The fact that the parameter estimates in Table 6 do not match those usually reported elsewhere in the literature is indicative that our benchmark model is most likely misspecified[13]. Lastly, our results also suggest that the estimated model requires steady-state aggregates that exceed their observed values for the U.S. economy. In particular, the observed wealth distribution can only be matched if individuals in the model save a larger fraction of their income, implying a higher equilibrium real interest rate. As a consequence, the steady-state capital-output ratio will also be larger.

7. Conclusions

In this paper we introduce a likelihood approach to estimate the structural parameters of macroeconomic heterogeneous agent (HA) models using microeconomic data. Our approach makes use of the Fokker-Planck equations that describe the stationary probability density function of the model which is used to build the likelihood function.

Using a standard Bewley-Hugget-Aiyagari model as the data generating process, we perform extensive Monte Carlo experiments to study the finite sample properties of the proposed ML estimator. To investigate its identification power, we propose to use the Kullback-Leibler (KL) divergence as a tool to determine potential sources of irregular behavior in the likelihood function before any estimation is conducted.

The simulation results show that the parameters related to the supply side of the economy and the household's subjective discount rate can be identified and accurately estimated with the use of cross-sectional data on individual wealth. On the other hand, the parameters describing the exogenous income process and the coefficient of relative risk aversion pose some challenges that materialize in significant biases that persist even in large samples. The KL divergence indicates that changes in these parameters do not affect significantly the shape of the wealth distribution, and therefore imply flat likelihood surfaces in these dimensions of the parameter space. The lack of curvature translates into

[13] In the Online Appendix we report the estimation results from a modified version of the benchmark model that includes both income and discount factor heterogeneity along the lines of Krusell and Smith (1998).

weakly identified parameters that could lead to incorrect inferences. However, our results also suggest that including data on individual income in addition to the wealth data can help to reduce these biases.

Following standard practice, we instead calibrate some of the troublesome parameters and estimate all remaining ones. Simulation evidence suggest that this approach delivers significant improvements over the unrestricted ML estimation. However, given the risk of mis-calibrating some of these parameters, our results favor fixing the risk aversion coefficient over any of the income parameters. To illustrate our approach, we provide a small empirical application in which we estimate the parameters of an extended version of our benchmark Bewley-Hugget-Aiyagari model using household data on wealth and income from the Survey of Consumer Finances. Despite the simplistic nature of the model, our estimates match 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 methods in HA models. The increased quality and quantity of micro data should direct future research towards more elaborated models, like those studied in Krusell and Smith (1998), Cagetti and De Nardi (2006), Angeletos and Calvet (2006), Angeletos (2007) and Benhabib et al. (2011), among others, or more realistic income processes like those in Achdou et al. (2014) and Gabaix et al. (2016). The ML approach introduced here could then be extended by using the approximated solution to the time-varying Fokker-Plank equation instead of its stationary version. However, this will impose some computational challenges that need to be addressed if one wishes to continue using the entire cross-sectional data as done here. A potential way to overcome this difficulties is to use some dimensionality-reduction technique similar to those introduced in the recent HANK literature (see e.g., Bayer and Luetticke, 2020, Papp and Reiter (2020), Auclert et al. (2021), Liu and Plagborg-Møller, 2021, among others). This will help to extend the information set used in the estimation process, e.g., repeated cross-sections or panel data, potentially increase the identification power of the structural parameters, and eventually provide a better fit of the wealth distribution.

References

ABBOTT, B., G. GALLIPOLI, C. MEGHIR, AND G. L. VIOLANTE (2019): "Education Policy and Intergenerational Transfers in Equilibrium," *Journal of Political Economy*, 127, 2569–2624.

AÇIKGÖZ, O. T. (2018): "On the Existence and Uniqueness of Stationary Equilibrium in Bewley Economies with Production," *Journal of Economic Theory*, 173, 18–55.

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 (2022): "Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach," *The Review of Economic Studies*, 89, 45–86.

AHN, S., G. KAPLAN, B. MOLL, T. WINBERRY, AND C. WOLF (2017): "When Inequality Matters for Macro and Macro Matters for Inequality," in *NBER Macroeconomics Annual 2017, volume 32*, National Bureau of Economic Research, Inc, NBER Chapters, 1–75.

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

AKAIKE, H. (1973): *Information Theory and an Extension of the Maximum Likelihood Principle*, in B. N. Petrov and F. Csaki (Eds.), Proceedings of the 2nd International Symposium on Information Theory, Budapest: Akademiai Kiado.

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.

Auclert, A., B. Bardóczy, M. Rognlie, and L. Straub (2021): "Using the Sequence-Space Jacobian to Solve and Estimate Heterogeneous-Agent Models," *Econometrica*, 89, 2375–2408.

Barro, R. J. (2021): "Double Counting of Investment," *The Economic Journal*, 131, 2333–2356.

Bayer, C., B. Born, and R. Luetticke (2020): "Shocks, Frictions, and Inequality in US Business Cycles," CEPR Discussion Papers 14364, C.E.P.R. Discussion Papers.

Bayer, C. and R. Luetticke (2020): "Solving Discrete Time Heterogeneous Agent Models with Aggregate Risk and many Idiosyncratic States by Perturbation," *Quantitative Economics*, 11, 1253–1288.

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.

Benhabib, J. and A. Bisin (2018): "Skewed Wealth Distributions: Theory and Empirics," *Journal of Economic Literature*, 56, 1261–91.

Benhabib, J., A. Bisin, and M. Luo (2019): "Wealth Distribution and Social Mobility in the US: A Quantitative Approach," *American Economic Review*, 109, 1623–47.

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 (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.

Fernández-Villaverde, J., S. Hurtado, and G. Nuño (2020): "Financial Frictions and the Wealth Distribution," CESifo Working Paper Series 8482, CESifo.

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.

HOLM, M. B. (2022): "Monetary Transmission with Income Risk," *The Scandinavian Journal of Economics*.

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.

IRELAND, P. N. (2004): "A Method for Taking Models to the Data," *Journal of Economic Dynamics and Control*, 28, 1205–1226.

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

KAPLAN, G., B. MOLL, AND G. L. VIOLANTE (2018): "Monetary Policy According to HANK," *American Economic Review*, 108, 697–743.

KHIEU, H. AND K. WÄLDE (2019): "Capital Income Risk and the Dynamics of the Wealth Distribution," CESifo Working Paper Series 7970, CESifo Group Munich.

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.

KULLBACK, S. (1959): *Information Theory and Statistics*, John Wiley and Sons Inc.

KULLBACK, S. AND R. A. LEIBLER (1951): "On Information and Sufficiency," *The Annals of Mathematical Statistics*, 22, 79–86.

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

LIGHT, B. (2020): "Uniqueness of Equilibrium in a Bewley-Aiyagari Model," *Economic Theory*, 69, 435–450.

Liu, L. and M. Plagborg-Møller (2021): "Full-Information Estimation of Heterogeneous Agent Models Using Macro and Micro Data," mimeo.

Luo, M. and S. Mongey (2019): "Assets and Job Choice: Student Debt, Wages and Amenities," Working Paper 25801, National Bureau of Economic Research.

McCulloch, R. E. (1989): "Local Model Influence," *Journal of the American Statistical Association*, 84, 473–478.

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

Mumtaz, H. and F. Zanetti (2015): "Factor Adjustment Costs: A Structural Investigation," *Journal of Economic Dynamics and Control*, 51, 341–355.

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.

Papp, T. K. and M. Reiter (2020): "Estimating Linearized Heterogeneous Agent Models using Panel Data," *Journal of Economic Dynamics and Control*, 115.

Prescott, E. C. (1986): "Theory Ahead of Business Cycle Measurement," *Carnegie-Rochester Conference Series on Public Policy*, 25, 11–44.

Qu, Z. and D. Tkachenko (2012): "Identification and Frequency Domain Quasi-Maximum Likelihood Estimation of Linearized Dynamic Stochastic General Equilibrium Models," *Quantitative Economics*, 3, 95–132.

——— (2017): "Global Identification in DSGE Models Allowing for Indeterminacy," *Review of Economic Studies*, 84, 1306–1345.

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.

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

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

White, H. (1982): "Maximum Likelihood Estimation of Misspecified Models," *Econometrica*, 50, 1–25.

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

Winberry, T. (2018): "A Method for Solving and Estimating Heterogeneous Agent Macro Models," *Quantitative Economics*, 9, 1123–1151.

Wong, A. (2021): "Refinancing and The Transmission of Monetary Policy to Consumption," Unpublished.

Young, E. R. (2010): "Solving the Incomplete Markets Model with Aggregate Uncertainty using the Krusell-Smith Algorithm and Non-Stochastic Simulations," *Journal of Economic Dynamics and Control*, 34, 36–41.