



DEPARTMENT OF ECONOMICS  
AND BUSINESS ECONOMICS  
AARHUS UNIVERSITY



## **Estimation of heterogeneous agent models: A likelihood approach**

**Juan Carlos Parra-Alvarez, Olaf Posch and Mu-Chun Wang**

**CREATEES Research Paper 2020-05**

# Estimation of heterogeneous agent models: A likelihood approach\*

Juan Carlos Parra-Alvarez<sup>†(a,b,e)</sup>, Olaf Posch<sup>(b,c)</sup> and Mu-Chun Wang<sup>(d)</sup>

<sup>(a)</sup>Aarhus University, <sup>(b)</sup>CREATES, <sup>(c)</sup>Universität Hamburg

<sup>(d)</sup>Deutsche Bundesbank, <sup>(e)</sup>Danish Finance Institute

May 2020

## Summary

We study the statistical properties of heterogeneous agent models. Using a Bewley-Hugget-Aiyagari model we compute the density function of wealth and income and use it for likelihood inference. We study the finite sample properties of the maximum likelihood estimator (MLE) using Monte Carlo experiments on artificial cross-sections of wealth and income. We propose to use the Kullback-Leibler divergence to investigate identification problems that may affect inference. Our results suggest that the unrestricted MLE leads to considerable biases of some parameters. Calibrating weakly identified parameters allows to pin down the other unidentified parameter without compromising the estimation of the remaining parameters. We illustrate our approach by estimating the model for the U.S. economy using wealth and income data from the Survey of Consumer Finances.

**Keywords:** Heterogeneous agent models, Continuous-time, Fokker-Planck equations, Kullback-Leibler divergence, Maximum likelihood.

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

---

\*This paper represents the author's personal opinions and does not necessarily reflect the views of the Deutsche Bundesbank. This version supersedes an earlier working paper entitled "Estimation and identification of heterogeneous agent models: A likelihood approach" (CREATES WP 2017-35). The authors 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-UoO Workshop on Economic Dynamics (Oslo, 2015), the IWH-CIREQ-GW Macroeconomic Workshop (Halle, 2016), the IAAE Annual Conference (Nikosia, 2019) and seminar participants at various institutions for their helpful comments and discussions. The first two authors acknowledge support from the Center for Research in Econometric Analysis of Time Series (DNRF78), CRESTES.

†Corresponding author: Department of Economics and Business Economics, Aarhus University, Fuglesangs Allé 4, 8210 Aarhus V, Denmark. Email address: [jparra@econ.au.dk](mailto: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 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, 2018](#); [Kaplan et al., 2018](#); [Wong, 2019](#)).

Currently, the main workhorse of household heterogeneity 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 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 provide a simple likelihood framework that exploits the information content in microeconomic data to estimate the structural parameters of macroeconomic models with heterogeneous agents. 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 state variables in heterogeneous agent 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 \(2009\)](#) for a textbook treatment.). However, [Bayer and Wälde \(2010a,b, 2011\)](#), [Achdou et al. \(2014\)](#), and [Gabaix et al. \(2016\)](#) have recently suggested the use of Fokker-Planck or Kolmogorov's Forward 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. \(2020\)](#). It is the ability to compute the model's non-parametric probability density function via the 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.

A condition for the maximum likelihood estimator to deliver consistent estimates of the model parameters, and 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 in fact absent. Therefore, it is important to verify the identification condition prior to estimation. 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, therefore, 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 maximum likelihood estimator. The KL divergence is computed for the model's 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 those used in [Mongey and Williams \(2017\)](#), [Williams \(2017\)](#) and [Winberry \(2018\)](#). In particular, 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. 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 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.

To illustrate our approach, Section 2 introduces a continuous-time version of an otherwise standard Bewley-Hugget-Aiyagari model in which a large number of households face idiosyncratic and uninsurable income risk in the form of exogenous shocks to their 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 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 stationary joint density function from the model. 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, e.g. measurement errors, are considered.

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 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 coefficient of relative risk aversion and on the parameters describing the income process. The biases in these parameters persists in large samples and their distributions tend to be wide and skewed. In Section 5 we compute the KL divergence associated to the model’s implied distribution function and find that the poor performance of the maximum likelihood estimator 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 very 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 6 we investigate the consequences of following such a strategy and find that even in cases where some parameters are mis-calibrated the maximum likelihood estimator, conditional on some parameters being calibrated, improves the finite sample properties of the parameters being estimated. Section 7 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 8 concludes.

## 2. A prototypical heterogeneous agent model

For our study we consider a prototypical heterogeneous agent model *à la* Bewley-Hugget-Aiyagari following Achdou et al. (2020). 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.

### 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

$$U_0 \equiv \mathbb{E}_0 \int_0^\infty e^{-\rho t} u(c_t) dt, \quad u' > 0, \quad u'' < 0, \tag{2.1}$$

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

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

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

$$da_t = (ra_t + we_t - c_t)dt, \quad a_0 \in \mathcal{A}, \quad (2.2)$$

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](#)). The latter could also be interpreted as productivity shocks (see [Heer and Trede, 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. 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 (2.2) 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

$$de_t = -\Delta_e dq_{1,t} + \Delta_e dq_{2,t}, \quad \Delta_e \equiv e_h - e_l \quad \text{and} \quad e_0 \in \mathcal{E}, \quad (2.3)$$

where  $\Delta_e$  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

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

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,  $\mathcal{A} = [\underline{a}, \infty)$ .

## 2.2 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  $\alpha \in (0, 1)$ , in order 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

$$Y = C + \delta K, \quad (2.4)$$

holds at every instant of time, where  $C$  denotes aggregate consumption, and  $\delta K$  denotes 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 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, as the one assumed throughout, 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  $t$  by the Hamilton-Jacobi-Bellman equation (HJB)

$$\begin{aligned} \rho V(a, e) &= \max_{c \in \mathbb{R}^+} \left\{ u(c) + V_a(a, e)(ra + we - c) \right. \\ &\quad \left. + (V(a, e_l) - V(a, e_h))\phi_1(e) + (V(a, e_h) - V(a, e_l))\phi_2(e) \right\}, \end{aligned} \quad (2.5)$$

where  $V(a, e)$  denotes the value function of the agent<sup>1</sup>. The first-order condition for an

---

<sup>1</sup>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 Appendices A and B.

interior solution reads

$$u'(c) = V_a(a, e) \quad (2.6)$$

for any  $t \in [0, \infty)$ , making optimal consumption a function only of the state variables and independent of time,  $c = c(a, e)$ . Equation (2.6) implies that in equilibrium, the instantaneous increase in utility from marginally consuming more must be exactly equal to the increase in overall utility brought by 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

$$\rho V(a, e_l) = u(c(a, e_l)) + V_a(a, e_l)(ra + we_l - c(a, e_l)) + (V(a, e_h) - V(a, e_l))\phi_{lh}, \quad (2.7)$$

$$\rho V(a, e_h) = u(c(a, e_h)) + V_a(a, e_h)(ra + we_h - c(a, e_h)) + (V(a, e_l) - V(a, e_h))\phi_{hl} \quad (2.8)$$

As argued in Achdou et al. (2020), an interesting feature of our continuous-time setup as opposed to the discrete-time case, is that (2.6) 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 (2.7) and (2.8) does not get affected by the existence of the inequality constraint  $a \geq \underline{a}$ , and instead gives rise to the following state-constraint boundary condition

$$V_a(\underline{a}, e) \geq u'(ra + we). \quad (2.9)$$

It can be shown that (2.9) implies that  $ra + we - c(\underline{a}, e) \geq 0$  and therefore 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

$$r = \alpha K^{\alpha-1} L^{1-\alpha} - \delta \quad \text{and} \quad w = (1 - \alpha) K^\alpha L^{-\alpha}, \quad (2.10)$$

where the steady state aggregate capital is obtained by aggregating the wealth held by every type of household,  $K = \sum_{e_t \in \{e_l, e_h\}} \int_{\underline{a}}^{\infty} a_t g(a, e) da$ , and the steady state aggregate labor is obtained by aggregating their efficiency labor units,  $L = \sum_{e \in \{e_l, e_h\}} \int_{\underline{a}}^{\infty} e_t g(a, e) da$ . 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,  $g(a, e)$ , can be split into  $g(a, e_h)$  and  $g(a, e_l)$ . Following Khieu and Wälde (2019), we refer to these individual probability functions as *subdensities*. For each  $e \in \mathcal{E}$ , it follows that  $g(a, e) \equiv g(a | e) p(e)$ , implying that

$$\int g(a, e) da = p(e), \quad (2.11)$$

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

$$g(a) = g(a, e_h) + g(a, e_l). \quad (2.12)$$

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

$$p(e_h) = \frac{\phi_{lh}}{\phi_{hl} + \phi_{lh}}, \text{ and } p(e_l) = \frac{\phi_{hl}}{\phi_{hl} + \phi_{lh}}. \quad (2.13)$$

Let  $s(a, e) = ra + we - c(a, e)$  denote the optimal savings function for an individual with an efficiency endowment equal to  $e \in \mathcal{E}$ . The subdensities in (2.12) correspond to the solution of the following non-autonomous quasi-linear system of differential equations known as (stationary) Fokker-Planck equations

$$s(a, e_l) \frac{\partial}{\partial a} g(a, e_l) = - \left( r - \frac{\partial}{\partial a} c(a, e_l) + \phi_{lh} \right) g(a, e_l) + \phi_{hl} g(a, e_h) \quad (2.14)$$

$$s(a, e_h) \frac{\partial}{\partial a} g(a, e_h) = - \left( r - \frac{\partial}{\partial a} c(a, e_h) + \phi_{hl} \right) g(a, e_h) + \phi_{lh} g(a, e_l), \quad (2.15)$$

where the partial derivatives with respect to wealth describe the cross-sectional dimension of the density function. The system of equations formed by (2.14) and (2.15) 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.

*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 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. [Achdou et al. \(2020\)](#) 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 [Açıkgoz \(2018\)](#) and [Light \(2020\)](#).

**Definition 2.1 (Competitive stationary equilibrium)** *A competitive stationary equilibrium is given by pairs of value functions  $V(a, e_l)$  and  $V(a, e_h)$ , individual policy functions for consumption  $c(a, e_l)$  and  $c(a, e_h)$ , time-invariant density functions  $g(a, e_l)$  and  $g(a, e_h)$ , a vector of constant prices  $[w, r]$ , and a vector of constant aggregates  $[K, L, Y, C]$  such that:*

1. the consumption functions  $c(a, e_l)$  and  $c(a, e_h)$  satisfy (2.7) and (2.8), i.e. they solve the household's consumption-saving problem,
2. factor prices satisfy the first order condition in (2.10), i.e. they solve the firm's allocation problem,
3. markets clear, i.e. (2.4) holds.
4. the joint probability density function of wealth and income is stationary, that is,  $\partial g(a, e, t) / \partial t = 0$  for all  $(a, e) \in \mathcal{A} \times \mathcal{E}$ , and the associated subdensities satisfy (2.14) and (2.15).

## 2.4 Computation of the equilibrium

The solution of our prototype economy is not available in closed form. Therefore, for a given set of parameter values, the stationary competitive equilibrium in Definition 2.1 is approximated on a discretized state space for  $\mathcal{A}$ . The algorithm we use to approximate the solution builds on earlier work by Candler (1999) and Achdou et al. (2020) which 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 (2.7) and (2.8) 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 (2.14) and (2.15). Having approximated the density function, the factor prices from the outer block are updated and the algorithm iterates until convergence. A detailed description of the algorithm and its implementation can be found in the Online Appendix D .

## 3. Structural estimation: The likelihood function

While there is a broad consensus on the importance of heterogeneity in macroeconomics, there is less agreement on how these models should be taken to the data. 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 the state variables which can be then used to write the model's likelihood function.

Let  $\mathbf{a} = [a_1, \dots, a_N]$  and  $\mathbf{e} = [e_1, \dots, e_N]$  be a sample of  $N$  i.i.d observations on individual wealth and income, respectively. Then, the log-likelihood function of the prototype model in Section 2 can be computed as

$$\mathcal{L}_N(\boldsymbol{\theta} | \mathbf{a}, \mathbf{e}) = \sum_{i=1}^N \log g(a_i, e_i | \boldsymbol{\theta}), \quad (3.1)$$

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

$$\hat{\boldsymbol{\theta}}_N = \arg \max_{\boldsymbol{\theta} \in \Theta} \mathcal{L}_N(\boldsymbol{\theta} | a_1, e_1, \dots, a_N, e_N). \quad (3.2)$$

Alternatively, by using identity (2.12), it is possible to obtain the marginal density of wealth as  $g(a_i | \boldsymbol{\theta}) = g(a_i, e_l | \boldsymbol{\theta}) + g(a_i, e_h | \boldsymbol{\theta})$  for each  $i = 1, \dots, N$ . Therefore, in situations where only data on individual wealth is available, we can rewrite the log-likelihood function as

$$\mathcal{L}_N(\boldsymbol{\theta} | \mathbf{a}) = \sum_{i=1}^N \log g(a_i | \boldsymbol{\theta}). \quad (3.3)$$

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 Section 2. 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  $g(a_i, e_i | \boldsymbol{\theta})$ ; (ii) Once  $g(a_i, e_i | \boldsymbol{\theta})$  has been evaluated for all  $(a_i, e_i) \in \mathbf{a} \times \mathbf{e}$ , the log-likelihood function is computed using (3.1).

## 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 artificially generated data of individual wealth and individual income (labor efficiency). 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

**Table 1. Population parameters,  $\theta_0$ .** The exogenous endowment of efficiency units is given by  $de = -\Delta_e dq_1 + \Delta_e dq_2$ , with  $\Delta_e \equiv e_h - e_l$ , where  $q_1$  and  $q_2$  are Poisson processes with intensity rates  $\phi_{lh}$  and  $\phi_{hl}$  respectively. The representative household has standard preferences defined by  $U_t = \mathbb{E}_t [\int_t^\infty e^{\rho(s-t)} u(c) ds]$  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, $\gamma$	1.0000
Rate of time preference, $\rho$	0.0490
Capital share in production, $\alpha$	0.3600
Depreciation rate of capital, $\delta$	0.1038
Endowment of high efficiency, $e_h$	1.0000
Endowment of low efficiency, $e_l$	0.2000
Demotion rate, $\phi_{hl}$	0.5578
Promotion rate, $\phi_{lh}$	7.3822

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,  $\Delta_e$ , of 80%, which is consistent with the values used in [Huggett \(1993\)](#), and [Imrohoroglu \(1989\)](#) and [Krusell and Smith \(1998\)](#). The transition rates for the Poisson processes are computed using (2.13). The remaining parameter values are fairly standard in the literature and are consistent with the long-run averages observed for the capital-output ratio and interest rates.

The Monte Carlo experiment is based on  $M = 200$  samples from the model's population stationary joint density function  $g(a, e | \boldsymbol{\theta}_0)$ , each of them of size  $N \in \{1,000, 5,000, 10,000\}$ <sup>2</sup>. We first sample the two state labor efficiency units using the marginal stationary distribution in (2.13). 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  $a = 0$  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 proposed maximum likelihood estimator when the econometrician only has access to data on wealth, and when also has data on 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,  $MNB = \text{median}\{(\hat{\theta}_m - \theta_0)/\theta_0\}$ ; (ii) the InterQuartile Range,  $IQR = Q_3 - Q_1$ , where  $Q_1$  and  $Q_3$  are the 25th and 75th percentiles of the distribution of the parameter es-

<sup>2</sup>Each Monte Carlo experiment takes up to 28 hours on a dedicated 32 cores Xeon server.

**Table 2. Finite sample properties of the ML estimator.** For each  $\theta \in \Theta$ , the table reports the Median Normalized Bias (MNB), the InterQuartile Range (IQR), 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	IQR	MANE	MNB	IQR	MANE	MNB	IQR	MANE
$\theta$	$N=1,000$			$N=5,000$			$N=10,000$		
Panel A: Wealth only									
$\gamma$	7.139	7.148	6.374	2.399	6.487	3.739	3.165	5.960	3.754
$\rho$	0.056	0.051	0.506	-0.058	0.035	0.383	0.058	0.025	0.303
$\alpha$	-0.218	0.055	0.225	-0.243	0.068	0.239	-0.205	0.077	0.192
$\delta$	0.118	0.021	0.212	0.114	0.022	0.189	0.083	0.021	0.180
$e_l$	4.689	0.498	4.337	4.126	0.745	3.723	3.651	0.823	3.299
$e_h$	0.892	0.674	0.867	0.926	0.746	0.890	0.715	0.858	0.719
$\phi_{lh}$	0.071	7.464	0.468	-0.093	4.576	0.351	0.030	3.424	0.282
$\phi_{hl}$	0.457	1.704	2.004	-0.061	1.142	1.547	0.257	0.746	1.039
Panel B: Wealth and income									
$\gamma$	4.028	8.076	4.797	0.084	3.535	2.297	0.057	2.498	1.933
$\rho$	0.168	0.040	0.431	-0.011	0.024	0.312	0.223	0.022	0.308
$\alpha$	-0.205	0.089	0.215	-0.229	0.060	0.221	-0.138	0.062	0.155
$\delta$	0.125	0.020	0.200	0.135	0.019	0.194	0.091	0.028	0.201
$e_l$	4.378	0.864	4.153	2.016	0.996	2.905	0.650	0.810	2.114
$e_h$	0.940	0.739	0.848	0.898	0.714	0.858	0.567	0.730	0.618
$\phi_{lh}$	-0.070	5.968	0.417	0.057	5.270	0.389	0.120	3.798	0.351
$\phi_{hl}$	-0.070	0.470	0.449	0.046	0.430	0.398	0.139	0.296	0.355

timate; and (iii) the Mean Absolute Normalized Errors,  $\text{MANE} = (1/M) \sum_m |(\hat{\theta}_m - \theta_0)/\theta_0|$ . We report normalized metrics in order to avoid scale 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 data on both individual wealth and income are used. A careful analysis of the (normalized) mean and median biases, and the interquartile range provides information on the small sample distributions of the parameter estimates. In particular, substantial differences between (the absolute value of) MNB and MANE point towards small sample parameter distributions that are skewed, while large values for the IQR would imply high levels of uncertainty around the parameter's estimate.

The simulation results suggest 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,  $\gamma$ , the levels of labor efficiency units,  $\{e_l, e_h\}$ , and their transition rates,  $\{\phi_{lh}, \phi_{hl}\}$ . In particular, their ML estimates imply extremely risk averse households as well as income levels and transition rates that are up to 2 and 4 times higher than their population values.

Their distributions are skewed and show considerable variability. On the other hand, the biases on the discount rate,  $\rho$ , the capital share in production,  $\alpha$ , and on the depreciation rate,  $\delta$ , are within reasonable ranges, even in small samples.

In general, increasing the sample size reduces the estimated bias for most of the parameters as well as the uncertainty around them. 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 two to three orders of magnitude. 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 for  $\gamma$ ,  $e_l$  and  $\phi_{hl}$ . However, the additional information brought by the use of income data has limited effects on the estimation errors and biases for  $\rho$ ,  $\alpha$  and  $\delta$ .

## 5. 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*,  $G_0 \equiv G(a, e | \boldsymbol{\theta}_0)$ , and its associated *population density function*,  $g_0 \equiv g(a, e | \boldsymbol{\theta}_0)$ . Since the model's probability density function constitutes the building block of the maximum likelihood estimator in (3.1), 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](#)). Let  $\tilde{G} \equiv G(a, e | \boldsymbol{\theta})$  and  $\tilde{g} \equiv g(a, e | \boldsymbol{\theta})$  denote the model's implied wealth-income distribution and density functions for  $\boldsymbol{\theta} \neq \boldsymbol{\theta}_0$ . Then, the KL divergence from  $\tilde{G}$  to  $G_0$  is defined as

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

The value of the KL divergence,  $k \geq 0$ , measures the information differences between the two distributions  $G_0$  and  $\tilde{G}$ . If  $k = 0$ , then it follows that  $G_0 = \tilde{G}$  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  $\mathcal{D}_{KL}(B(0.5) \parallel B(q)) = \mathcal{D}_{KL}(G_0 \parallel \tilde{G})$ <sup>3</sup>. As an example, suppose that the probability implied by the two distributions  $G_0$  and  $\tilde{G}$  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  $\min_{\theta \in \Theta} \mathcal{D}_{KL}(G_0 \parallel \tilde{G}) = \max_{\theta \in \Theta} \mathcal{L}_N$  as  $N \rightarrow \infty$ . 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  $\theta_0$ , and hence, to inform on the ability of the ML estimator to identify the model parameters.

[Figure 1 about here.]

Figure 1 plots the probability  $q$  implied by the KL divergence from  $\tilde{G}$  to  $G$  as we vary each  $\theta \in \Theta$  while keeping the remaining parameters at their population values. All the KL divergences, with the exception of that for  $\gamma$ , are constructed using parameter values that are 50% below and 50% above of the true parameter value. In the case of  $\gamma$ , 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  $\rho$ ,  $\alpha$  and  $\delta$  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 these subset of parameters given that small differences in their values will produce significantly different density functions. On the contrary, the influence of  $\gamma$  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 4. 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  $\gamma$  have virtually no effect on the wealth distribution, whereas small changes in  $\alpha$  lead to substantial differences in the wealth distribution<sup>4</sup>.

[Figure 2 about here.]

---

<sup>3</sup>McCulloch (1989) shows that the KL divergence from  $B(q)$  to  $B(0.5)$  is given by  $\mathcal{D}_{KL}(B(0.5) \parallel B(q)) = -\log(4q(1-q))/2$ . Thus,  $q$  measures the divergence of an arbitrary Bernoulli trial from a fair Bernoulli trial. Given the KL divergence  $k$  from  $\tilde{G}$  to  $G_0$ , it is straightforward to compute the implied probability  $q$ .

<sup>4</sup>The sensitivity of the wealth distribution to changes in the remaining parameters of the model can be found in the Online Appendix E.

## 6. 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 begin by investigating the finite sample behavior of the ML estimator when the exogenous income process is externally calibrated. This strategy 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](#)). We conduct a set of Monte Carlo experiments where we analyze the properties of the ML estimator conditional on the following restrictions on the income parameters: (i) the income levels,  $e_l$  and  $e_h$ , are fixed to their population values; (ii) the transition rates,  $\phi_{lh}$  and  $\phi_{hl}$ , are fixed to their population values; (iii) all income parameters are fixed to their population values. 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  $N = 5,000$ . Table 3 summarizes the results by reporting the mean absolute normalized errors, MANE, and the median normalized bias, MNB. 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 4 for comparison.

The results show 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 relative risk aversion and the capital share in output, and to a lesser extent, for the transition rates. On the other hand, calibrating the transition rates of the income process alone has limited consequences. Although it helps to reduce the bias in  $\gamma$  by nearly half of that obtained from an unconstrained estimation, there is virtually no effect on the errors of the remaining parameters. If instead, we calibrate all the income parameters simultaneously, the ML estimator yields sharper estimates for the relative risk aversion but deteriorates the precision with which the depreciation rate can be estimated. Interestingly, note 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

**Table 3. Conditional estimates: income process calibrated 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  $N = 5,000$ . The ML estimation is conditional on the income parameters being calibrated to their population values as indicated in the second row.

$\theta$	Panel A: Wealth only				Panel B: Wealth and income			
	Income levels	Transition rates	All income	No restr.	Income levels	Transition rates	All income	No restr.
$\gamma$	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)
$\rho$	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)
$\alpha$	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)
$\delta$	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)
$\phi_{lh}$	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)
$\phi_{hl}$	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)

data in this model is represented by a binary variable and hence it only contains information about the stationary probabilities of states.

[Figure 3 about here.]

To further understand the effects of partially, or completely, calibrating the income process, Figure 3 plots the finite sample distribution of parameter estimates for the cases (i), (ii), and (iii) 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  $\gamma$  and  $\rho$  but it also generates distributions that exhibit multiple modes for  $\alpha$  and  $\delta$  that could suggest further identification problems<sup>5</sup>. 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  $\alpha$  and  $\delta$  simultaneously while keeping the remaining parameters at their population values.

<sup>5</sup>An earlier version of this paper shows that the multimodality in these parameters' density remains even in large samples (see, Parra-Alvarez et al., 2017).

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 observationally equivalent distribution functions, and therefore partial identification problems.

[Figure 4 about here.]

The tight relation between  $\alpha$  and  $\delta$  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  $\alpha$  and  $\delta$ , 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  $\alpha$  and  $\delta$  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 shown in Figure 3 the calibration of the income levels or of the transition rates yields distribution of estimates for  $\alpha$  and  $\delta$  that are unimodal and at the same time do not affect the accuracy with which  $\gamma$  can be identified. Hence, allowing  $\alpha$  and  $\delta$  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.

We next investigate the finite sample implications for our ML estimator of calibrating some of the household's and firm's structural parameters, while estimating all of the remaining parameters, including those describing the income process. We conduct a Monte Carlo simulation under the following restrictions: (i)  $\gamma$  is calibrated to its value in the population; (ii)  $\rho$  is calibrated to its true value in the population; (iii)  $\alpha$  is calibrated to its value in the population; (iv)  $\delta$  is calibrated to its value in the population; and (v)  $\alpha$  and  $\delta$  are calibrated to their values in the population. Table 4 summarizes our results by reporting the MANE and the MNB when using data on individual wealth and individual income in the estimation. The last column reports the results for the unrestricted ML estimation from Section 4 for comparison. The results suggest that calibrating  $\gamma$  reduces the biases and estimation errors in some dimensions of the parameter space. On the other hand, calibrating  $\rho$ ,  $\alpha$  and  $\delta$  separately hardly affects inferences at all. Interestingly, calibrating  $\alpha$  and  $\delta$  simultaneously does not generate an overall improvement over alternative calibrations.

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

**Table 4. Conditional estimates: preference parameters calibrated to their population values.** The table reports the Mean Absolute Normalized Error (MANE) and the Mean Normalized Bias (MNB, in parenthesis) from a Monte Carlo experiment with  $M = 200$  samples, each of them of size  $N = 5,000$ . The ML estimation is conditional on the preference parameters being calibrated to their population values as indicated in the second row. The ML estimation uses data on individual wealth and income.

$\theta$	Wealth and income					
	$\gamma$	$\rho$	$\alpha$	$\delta$	$\alpha$ and $\delta$	No restrictions
$\gamma$	$\gamma_0$	2.4793 (0.633)	3.0433 (0.065)	2.6989 (0.055)	2.5315 (0.010)	2.2969 (0.084)
$\rho$	0.3745 (0.170)	$\rho_0$	0.3604 (-0.234)	0.3291 (-0.027)	0.4754 (-0.484)	0.3125 (-0.011)
$\alpha$	0.1817 (-0.178)	0.2686 (-0.269)	$\alpha_0$	0.2545 (-0.267)	$\alpha_0$	0.2206 (-0.229)
$\delta$	0.1783 (0.143)	0.1982 (0.068)	0.5313 (0.612)	$\delta_0$	$\delta_0$	0.1941 (0.135)
$e_l$	1.4518 (0.945)	3.3686 (2.605)	2.1285 (1.296)	2.9364 (1.831)	0.7651 (0.148)	2.9053 (2.016)
$e_h$	0.8410 (0.873)	0.9460 (0.974)	0.5563 (0.575)	0.8214 (0.882)	0.2340 (-0.241)	0.8579 (0.898)
$\phi_{lh}$	0.3681 (0.342)	0.4152 (-0.092)	0.4476 (-0.059)	0.4136 (0.078)	0.4391 (-0.323)	0.3891 (0.057)
$\phi_{hl}$	0.3830 (0.320)	0.4157 (-0.083)	0.4556 (-0.089)	0.4182 (0.064)	0.4471 (-0.352)	0.3979 (0.046)

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 the risk aversion coefficient,  $\gamma$ , and the labor efficiency gap,  $\Delta e$ , 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,  $(Y - C)/Y$ , of the Gini coefficient, and of the Lorenz curve<sup>6</sup>. For comparison, the table also reports the

<sup>6</sup>Similar statistics for all the Monte Carlo simulations described in the paper are available upon request.

**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,  $\Delta_e$  larger uses  $e_h = 1.5$  and  $e_l = 0.1$ ;  $\Delta_e$  smaller uses  $e_h = 0.5$  and  $e_l = 0.1$ ;  $\gamma$  higher uses  $\gamma = 2.0$ ; and  $\gamma$  lower uses  $\gamma = 0.5$ . The ML estimation uses data on individual wealth and income.

	Wealth and income					
	$\Delta_e$ larger	$\Delta_e$ smaller	$\Delta_e = 0.8$	$\gamma$ higher	$\gamma$ 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)

case where  $\Delta_e$  and  $\gamma$  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 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  $\gamma$ , as suggested by the KL divergence and documented in Figure 1.

In summary, our Monte Carlo evidence suggests that  $\rho$ ,  $\alpha$  and  $\delta$  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 are 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.

## 7. 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<sup>7</sup>.

In order to accommodate the high degree of wealth inequality observed in the data, we follow [Castañeda et al. \(2003\)](#) and 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  $\mathcal{E} = \{e_1, e_2, e_3, e_4\}$  that evolves over time according to

$$de_t = \sum_i \sum_{j \neq i} (e_i - e_j) dq_{ij,t}, \quad e_0 \in \mathcal{E},$$

and where  $e_1 < e_2 < e_3 < e_4$ . The Poisson processes  $q_{ij,t}$  for all  $i, j = 1, \dots, 4$  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  $i = 1, \dots, 4$ .

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<sup>8</sup>. 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

---

<sup>7</sup>It should be stressed that our benchmark model is most likely misspecified as it does not account for some of the main drivers that characterize the wealth distribution in the U.S. Hence, the estimates reported are subject to considerably misspecification bias, and should be interpreted with caution.

<sup>8</sup>Using equally spaced quantile bins will produce much smaller overall estimates of the actual wealth-income inequality.

**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	$\gamma$	$\rho$	$\alpha$	$\delta$
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, $\phi_{ij}$ ( $\times 100$ )				
$i \setminus j$	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

to compute confidence intervals for the parameter estimates using  $M = 100$  bootstrap samples<sup>9</sup>. Following the Monte Carlo evidence of Section 6, we do not attempt to estimate the coefficient of relative risk aversion,  $\gamma$ . 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. 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.

To provide a better understanding of the estimated intensity rates in Panel B of Table

<sup>9</sup>The bootstrapping exercise is computationally demanding. Estimation with  $M = 100$  samples takes about 58 hours on a dedicated 32 cores Xeon server.

**Table 7. Transition probability matrix and stationary probabilities of the endowment of efficiency levels.** Panel A reports the implied maximum likelihood estimates (MLE) of the transition probabilities and their 95% confidence intervals computed from a non-parametric bootstrap. Panel B reports the implied maximum likelihood estimates (MLE) of the stationary probabilities of labor efficiency units and their 95% confidence intervals computed from a non-parametric bootstrap.

Panel A: Transition probability matrix, $p(e_i, e_j, t)$ (%)				
	$e_1$	$e_2$	$e_3$	$e_4$
$e_1$	99.47 [99.46, 99.55]	0.00 [0.00, 0.00]	0.28 [0.26, 0.30]	0.25 [0.19, 0.25]
$e_2$	0.14 [0.11, 0.15]	99.65 [99.63, 99.69]	0.04 [0.03, 0.04]	0.18 [0.15, 0.20]
$e_3$	0.20 [0.17, 0.21]	0.18 [0.17, 0.20]	99.61 [99.61, 99.66]	0.00 [0.00, 0.00]
$e_4$	0.04 [0.03, 0.04]	0.03 [0.03, 0.04]	32.18 [31.48, 35.02]	67.75 [64.91, 68.46]
Panel B: 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]

<sup>6</sup>, we compute the associated transition probabilities, i.e., the probability that the income process currently in state  $i$  transits to state  $j$  at an instant later in time. We denote these probabilities by  $p(e_i, e_j, t) \equiv \mathbb{P}(e_{t+s} = e_j | e_s = e_i)$  for all  $s \leq t$ <sup>10</sup>. Panel A of Table 7 reports the implied transition probabilities for the endowment process of efficiency units. Panel B reports the corresponding limiting distribution defined as  $p(e_i) \equiv \lim_{t \rightarrow \infty} p(e_i, t)$ , where  $p(e_i, t)$  denotes the unconditional probability of being in state  $i$  at time  $t$ .

Panel A indicates that the ML estimates of the intensity rates imply a very persistent income process, specially for the first three efficiency states. The table shows that individuals whose current efficiency endowment is either  $e_1$ ,  $e_2$  or  $e_3$  will rarely move to a different state. On the other hand, the most productive individuals have a relatively high probability to move to the second most productive level of efficiency. The stationary probabilities reported in Panel B 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 8 compares some wealth statistics computed from the data to those implied by the estimated model. In particular, it reports the Gini coefficient and the percentage of total wealth held by the top 5, 10 and 20 percentiles. 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

<sup>10</sup>A complete derivation of the transition probabilities can be found in the Online Appendix C.

**Table 8. Wealth Inequality: data vs. model.** The table reports the observed and estimated Gini coefficient and the distribution of wealth across top percentiles. It also reports 95% confidence intervals computed from a non-parametric bootstrap.

		% wealth in top		
	Gini Coefficient	5%	10%	20%
SCF 2013 data	0.8048	57.73	70.27	83.44
Model implied	0.7912 [0.7817, 0.7939]	44.30 [42.98, 44.76]	63.82 [62.27, 64.31]	83.57 [82.27, 83.91]

on labor income, our estimates indicate that the data favors the inclusion of an “awesome state” (see [Benhabib and Bisin \(2018\)](#) for a discussion.). 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 view our results not as evidence *for* this particular income process. As [Benhabib and Bisin \(2018\)](#) 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 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.

## 8. Conclusions

In this paper we introduce a likelihood approach to estimate the structural parameters of macroeconomic models with heterogeneous agent using microeconomic data. The feasibility of our approach follows from the use of continuous-time methods and in particular from the use of the Fokker-Planck equations that describe the stationary probability density function of the model which can then be used to build the model’s 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 maximum likelihood 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 and income. However, the coefficient of relative risk aversion and the parameters describing the exogenous income process display some identification challenges that materialize in significant biases that persist even

in large samples. The KL divergence indicates that changes in these parameters do not affect 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 weakly identified parameters that could lead to incorrect inferences.

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 maximum likelihood 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 a 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 heterogeneous agent macroeconomic models. The increased quality and quantity of micro data should direct future research towards more sophisticated models, like those studied in Krusell and Smith (1998), Cagetti and Nardi (2006), Angeletos and Calvet (2006), Angeletos (2007) and Benhabib et al. (2011), among others, which also allow for more realistic income processes like those in Achdou et al. (2014) and Gabaix et al. (2016). This will help to extend the information set used in the estimation process, potentially increase the identification power of the structural parameters, and eventually provide a better fit of the wealth distribution.

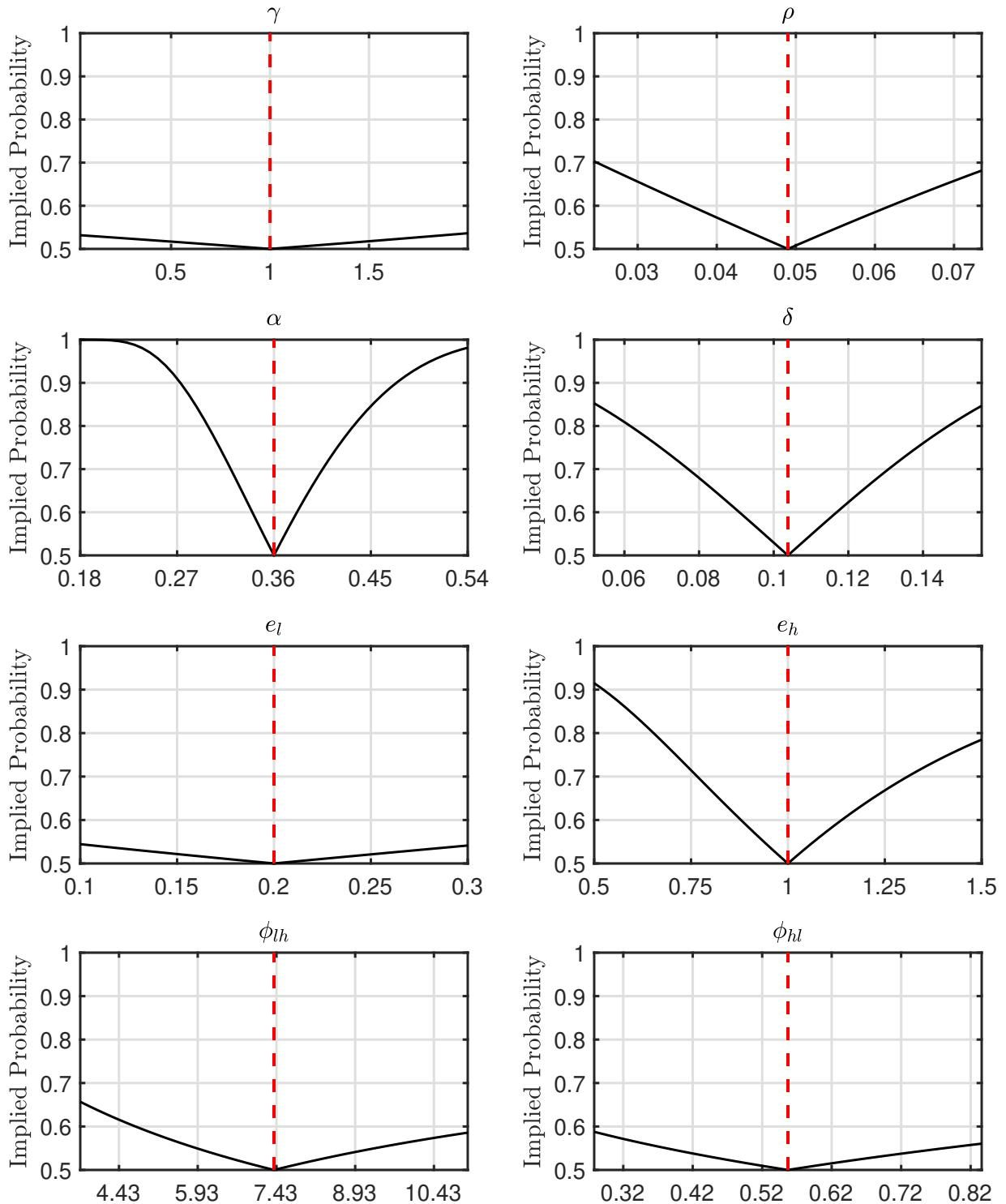
## References

- 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 (2020): “Income and Wealth Distribution in Macroeconomics: A Continuous-Time Approach,” Manuscript, available at <https://benjaminmoll.com>.
- 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.
- 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.

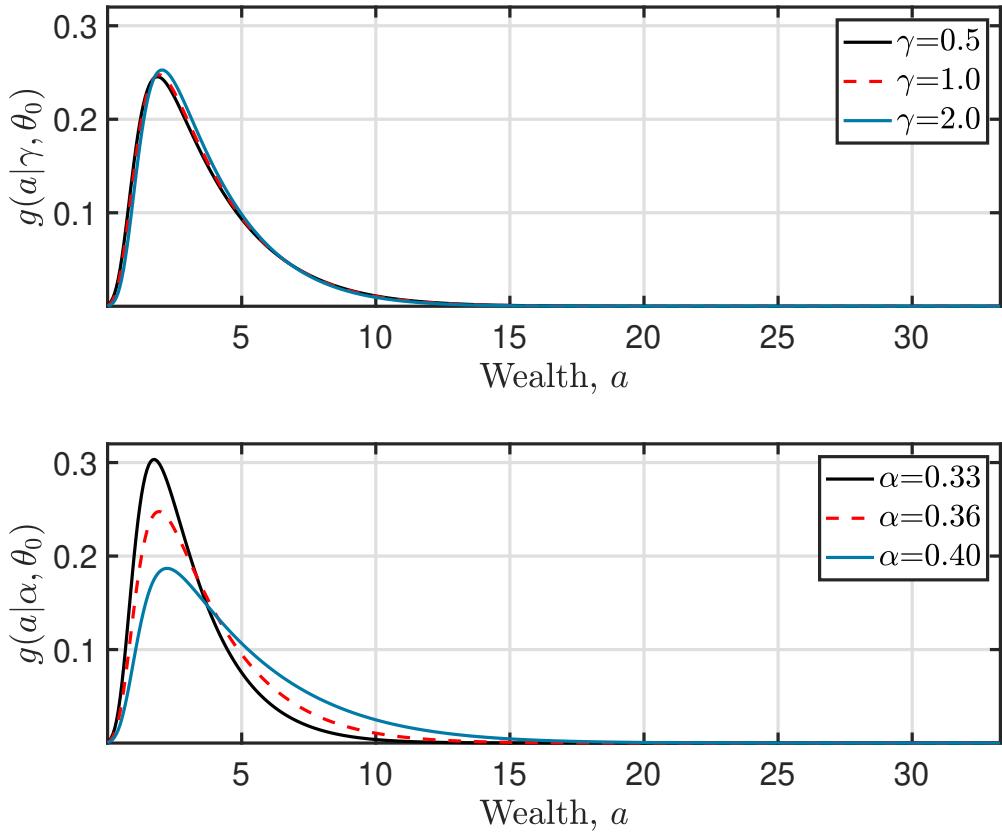
- BENHABIB, J., A. BISIN, AND M. LUO (2019): “Wealth Distribution and Social Mobility in the US: A Quantitative Approach,” *American Economic Review*, 109, 1623–47.
- BENHABIB, J., A. BISIN, AND S. ZHU (2011): “The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents,” *Econometrica*, 79, 123–157.
- BEWLEY, T. (Undated): “Interest Bearing Money and the Equilibrium Stock of Capital,” Manuscript.
- CAGETTI, M. AND M. 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. (2018): “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.
- IMROHOROĞLU, A. (1989): “Cost of Business Cycles with Indivisibilities and Liquidity Constraints,” *Journal of Political Economy*, 97, 1364–1383.
- ISKREV, N. (2010): “Local Identification in DSGE Models,” *Journal of Monetary Economics*, 57, 189–202.
- KAPLAN, G., B. MOLL, AND G. L. VIOLANTE (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.
- 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.

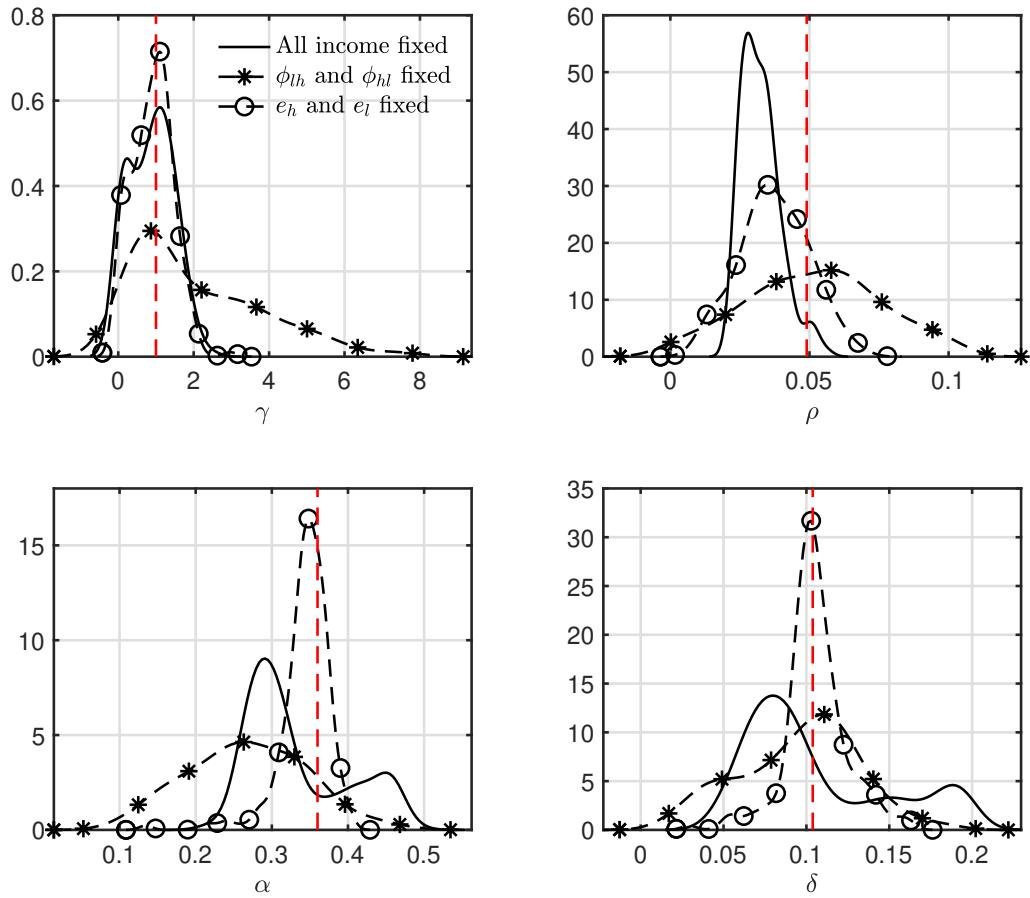
- NEWHEY, W. K. AND D. MCFADDEN (1986): “Large Sample Estimation and Hypothesis Testing,” in *Handbook of Econometrics*, ed. by R. F. Engle and D. McFadden, Elsevier, vol. 4 of *Handbook of Econometrics*, chap. 36, 2111–2245.
- OZKAN, S., K. MITMAN, F. KARAHAN, AND A. HEDLUND (2016): “Monetary Policy, Heterogeneity and the Housing Channel,” 2016 Meeting Papers 663, Society for Economic Dynamics.
- PARRA-ALVAREZ, J. C., 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.
- PRESSCOTT, 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.
- 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. (2019): “Refinancing and The Transmission of Monetary Policy to Consumption,” Unpublished.



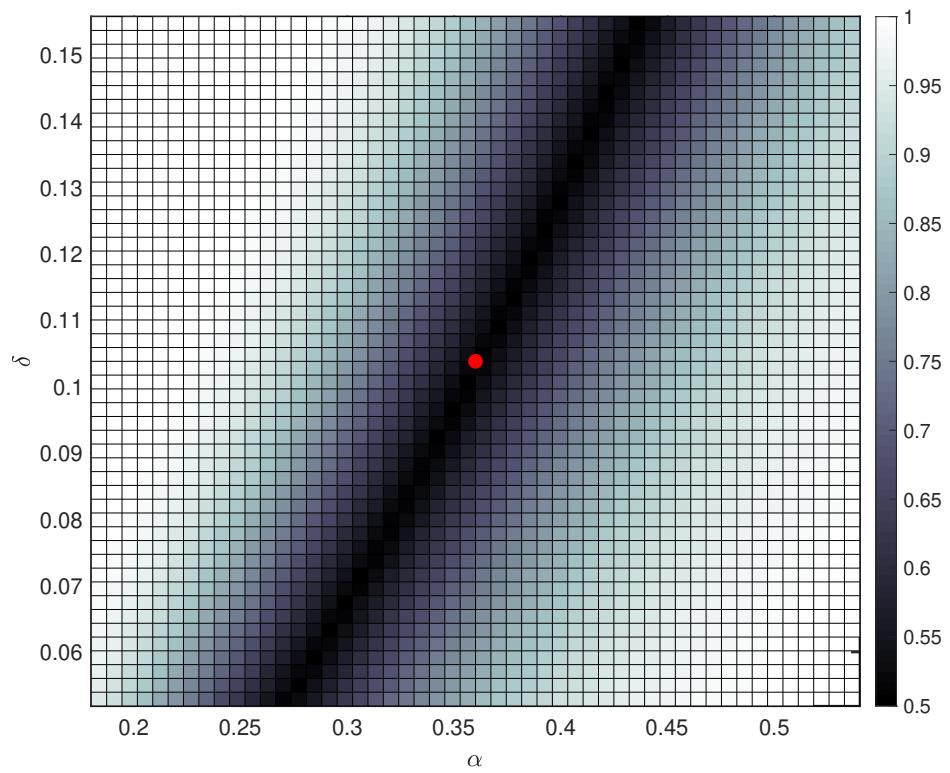
**Figure 1. Kullback-Leibler divergence.** The graph plots the implied probability  $q$  associated with the KL divergence from  $G(a | \boldsymbol{\theta})$  to  $G(a | \boldsymbol{\theta}_0)$  that results from varying each parameter at a time while keeping the remaining ones at their population value.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



**Figure 2. Sensitivity of the wealth distribution.** The graph shows the sensitivity of the distribution of wealth,  $g(a|\theta)$ , 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.



**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.



**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  $\alpha$  and  $\delta$  simultaneously while keeping the remaining parameters at their population value. The dot denotes the true parameter value.

# Research Papers

## 2020



- 2019-13: Daniel Borup and Erik Christian Montes Schütte: In search of a job: Forecasting employment growth using Google Trends
- 2019-14: Kim Christensen, Charlotte Christiansen and Anders M. Posselt: The Economic Value of VIX ETPs
- 2019-15: Vanessa Berenguer-Rico, Søren Johansen and Bent Nielsen: Models where the Least Trimmed Squares and Least Median of Squares estimators are maximum likelihood
- 2019-16: Kristoffer Pons Bertelsen: Comparing Tests for Identification of Bubbles
- 2019-17: Dakyung Seong, Jin Seo Cho and Timo Teräsvirta: Comprehensive Testing of Linearity against the Smooth Transition Autoregressive Model
- 2019-18: Changli He, Jian Kang, Timo Teräsvirta and Shuhua Zhang: Long monthly temperature series and the Vector Seasonal Shifting Mean and Covariance Autoregressive model
- 2019-19: Changli He, Jian Kang, Timo Teräsvirta and Shuhua Zhang: Comparing long monthly Chinese and selected European temperature series using the Vector Seasonal Shifting Mean and Covariance Autoregressive model
- 2019-20: Malene Kallestrup-Lamb, Søren Kjærgaard and Carsten P. T. Rosenskjold: Insight into Stagnating Life Expectancy: Analysing Cause of Death Patterns across Socio-economic Groups
- 2019-21: Mikkel Bennedsen, Eric Hillebrand and Siem Jan Koopman: Modeling, Forecasting, and Nowcasting U.S. CO<sub>2</sub> Emissions Using Many Macroeconomic Predictors
- 2019-22: Anne G. Balter, Malene Kallestrup-Lamb and Jesper Rangvid: The move towards riskier pensions: The importance of mortality
- 2019-23: Duván Humberto Cataño, Carlos Vladimir Rodríguez-Caballero and Daniel Peña: Wavelet Estimation for Dynamic Factor Models with Time-Varying Loadings
- 2020-01: Mikkel Bennedsen: Designing a sequential testing procedure for verifying global CO<sub>2</sub> emissions
- 2020-02: Juan Carlos Parra-Alvarez, Hamza Polattimur and Olaf Posch: Risk Matters: Breaking Certainty Equivalence
- 2020-03: Daniel Borup, Bent Jesper Christensen, Nicolaj N. Mühlbach and Mikkel S. Nielsen: Targeting predictors in random forest regression
- 2020-04: Nicolaj N. Mühlbach: Tree-based Synthetic Control Methods: Consequences of moving the US Embassy
- 2020-05: Juan Carlos Parra-Alvarez, Olaf Posch and Mu-Chun Wang: Estimation of heterogeneous agent models: A likelihood approach