Department Discussion Paper DDP2010

ISSN 1914-2838

HETEROGENEITY IN MACROECONOMICS AND THE MINIMAL ECONOMETRIC INTERPRETATION FOR MODEL COMPARISON

Marco Cozzi
Department of Economics, University of Victoria
Victoria, B.C., Canada V8W 2Y2

October, 2022

Abstract

I formally compare the fit of various versions of the incomplete markets model with aggregate uncertainty relying on the Minimal Econometric Interpretation, which is a computationally tractable Bayesian empirical framework. The models differ in the degree of household heterogeneity, with a focus on the role of preferences. For every specification, empirically motivated priors for the parameters are postulated to obtain the models' predictive distributions, which are interpreted as being distributions of population moments. These are in turn compared to the posterior distributions of the same moments obtained from an a-theoretical Bayesian econometric model. I show that aggregate data on consumption and income contain valuable information to determine which models are more likely to have generated the data. The two models featuring risk aversion heterogeneity have the highest marginal likelihoods, showing that this element is quantitatively important also for the study of aggregate outcomes. I also extend the framework to include the fit of the wealth Gini index, but the ranking of the models is only marginally affected.

Keywords: Heterogeneous Agents, Incomplete Markets, Unemployment Risk, Business Cycles, Bayesian Methods, Calibration, Minimal Econometric Interpretation, Model Comparison.

JEL Classifications: C63, C68, E21, E32, D52, D58.

Author Contact:
Marco Cozzi, Dept. of Economics, University of Victoria, P.O. Box 1700, STN CSC, Victoria, B.C., Canada V8W 2Y2; E-mail: mcozzi@uvic.ca; Tel: (250) 721-6535

Heterogeneity in Macroeconomics and the Minimal Econometric Interpretation for Model Comparison

Marco Cozzi, University of Victoria

October 2022

Abstract

I formally compare the fit of various versions of the incomplete markets model with aggregate uncertainty relying on the Minimal Econometric Interpretation, which is a computationally tractable Bayesian empirical framework. The models differ in the degree of household heterogeneity, with a focus on the role of preferences. For every specification, empirically motivated priors for the parameters are postulated to obtain the models' predictive distributions, which are interpreted as being distributions of population moments. These are in turn compared to the posterior distributions of the same moments obtained from an a-theoretical Bayesian econometric model. I show that aggregate data on consumption and income contain valuable information to determine which models are more likely to have generated the data. The two models featuring risk aversion heterogeneity have the highest marginal likelihoods, showing that this element is quantitatively important also for the study of aggregate outcomes. I also extend the framework to include the fit of the wealth Gini index, but the ranking of the models is only marginally affected.

JEL Classification Codes: C63, C68, E21, E32, D52, D58.

Keywords: Heterogeneous Agents, Incomplete Markets, Unemployment Risk, Business Cycles, Bayesian Methods, Calibration, Minimal Econometric Interpretation, Model Comparison.

Acknowledgments: *I am grateful to both the editor and two anonymous referees for helpful and constructive comments that greatly improved the paper, and to seminar participants at the SED Meetings in Warsaw and at the JRC in Ispra for useful suggestions. SSHRC provided financial support for this project with the IDG Grant # 4302013511.*

Contact details: *Department of Economics, University of Victoria, 3800 Finnerty Road, Victoria, BC V8P 5C2, Canada. Tel: 1-250-721-6535, E-mail:* `mcozzi@uvic.ca`

1 Introduction

Quantitative structural modeling represents a widely popular way of undertaking macroeconomic analysis. Models with incomplete markets, household-level heterogeneity, and a mix of idiosyncratic and aggregate shocks are especially valuable because they allow to perform counterfactual computational experiments together with studying the welfare and distributional implications of various policy interventions. A number of contributions assessing the consequences of business cycles with this class of models have found substantial welfare benefits of eliminating aggregate risk, which are approximately an order of magnitude larger than those originally documented by Lucas (1987). Notable studies in this literature are Krusell and Smith (1999), Storesletten, Telmer, and Yaron (2001), Mukoyama and Sahin (2006), and Krusell, Mukoyama, Sahin, and Smith (2009). Castaneda, Diaz-Gimenez, and Rios-Rull (1998), Heathcote (2005), and Chiu and Molico (2010) propose variants of the baseline framework to quantify the macroeconomic outcomes and distributional effects of cyclical variations in the income distribution, and of fiscal and monetary policies.

In order to perform a reliable welfare analysis, and provide a sound guidance for policy design, it is desirable to compare these models with the data in a systematic way. Furthermore, given that the researcher is free to consider many different dimensions of heterogeneity, among other modeling choices, the relative performance of different models should be assessed with an internally consistent empirical framework. Finally, since the issue of parameter uncertainty is important for many macroeconomic models, the empirical analysis should accommodate it. Although the use of Bayesian empirical methodologies has a long tradition in macroeconomic models with representative agents, the computational challenges of solving models with incomplete markets and aggregate shocks make using these estimation techniques an exceptional hurdle for most specifications.[1] In this paper, I compare four versions of the incomplete markets model with aggregate uncertainty by means of their Minimal Econometric Interpretation (MEI). This is a simple Bayesian empirical framework, proposed by Geweke (2010), which in turn represents a generalization of DeJong, Ingram, and Whiteman (1996). It is a computationally tractable procedure, because it requires calculating numerically the models' (marginal) likelihood only once, which can be accurately approximated with a large –but manageable– number of independent parameter draws (that are trivially parallelized across different processors). Unlike traditional Bayesian estimation, where the Markov chain Monte Carlo algorithm necessitates both an enormous number of sequential parameter draws and to evaluate the likelihood function at every iteration. From a methodological perspective, this framework interprets the economic models as purely theoretical tools that have implications for distributions of population moments. This is an appealing feature when dealing with set-ups that rely on a parsimonious specification for the shocks, an assumption made in most models with incomplete markets, because formal statistical testing based on the models' likelihood would reject many of them.[2]

[1]See, among others, An and Schorfheide (2007), Canova (2009), Fernandez-Villaverde, Guerron, and Rubio-Ramirez (2010), Rios-Rull, Schorfheide, Fuentes-Albero, Kryshko, and Santaeulalia-Llopis (2012), and Appendix F. In the recent literature, a number of contributions estimate incomplete markets models with classical methods, neglecting the role of parameter uncertainty. Also, notice that the method proposed by Ahn, Kaplan, Moll, Winberry, and Wolf (2017) is not suitable for Bayesian estimation, because in their model the reduction technique leads to 300 state variables, making the Kalman smoother intractable. Moreover, their method relies on a certainty equivalent approximation, which is known to be an inadequate tool for the computation of welfare effects.

[2]For instance, if the set of observables were to include the time series of the unemployment rate, all models considered in this

In particular, I use the MEI to analyze whether risk aversion heterogeneity helps accounting for the time series behavior of aggregate consumption, its correlation with income, and their relative volatility. These are key elements of any dynamic macroeconomic model, because they are intimately linked to consumption smoothing behavior. In models with incomplete markets, the welfare consequences of a policy intervention are affected by both the insurance properties of the policy being evaluated and by the households' response in terms of their consumption/saving decisions, which can trigger important general equilibrium effects. Ultimately, the level of trust in the quantitative assessment of the welfare effects arising from, say, a stabilization policy rests on a model's ability to adequately capture the dynamics of consumption at business cycle frequencies, and its correlation with income.

The starting point is the Krusell and Smith (1998) set-up, with or without discount factor heterogeneity. More novel variants of the model are proposed, where Panel Study of Income Dynamics (PSID) data are used to provide a parsimonious, yet data-driven, specification for heterogeneity in risk aversion. The emphasis is on treating this aspect as observed heterogeneity. The first contribution is to assess the empirical performance of these models, along the dimensions they are designed to tackle, namely the behavior of both consumption and income in a time series sense. I provide evidence on which specifications of the model are more likely to have generated the data, in an empirical framework that allows for parameter uncertainty, does not assume that one of the competing models is the true one, and interprets the model as a theoretical tool, providing information on population moments.[3] The results show that all models with preference heterogeneity possess some empirical validity, while the complete markets version of the model performs poorly. Using the Bayes factor as a formal measure of fit, the two models with risk aversion heterogeneity are found to dominate the other specifications. A second contribution is represented by a thorough Monte Carlo analysis, which shows that the MEI recovers the Data Generating Processs (DGP), by assigning higher marginal likelihoods to the true DGP. This occurs even when the model-generated samples are limited in size, and close to the dimension of actual datasets. A third contribution is to show how to incorporate microeconomic moments into the methodology, allowing to assess the models' fit also along this dimension. This is a valuable extension, because estimation methods based on the likelihood of aggregate data cannot credibly identify many microeconomic parameters, and combining likelihoods of aggregate and individual data is challenging. Consequently, I augment the set of empirical moments to include the wealth Gini index, but in my application this extension does not alter the ranking of the models in terms of their fit. A final contribution is to provide an assessment of the long-run welfare costs of business cycle fluctuations, which highlights how models with risk aversion heterogeneity can

paper would be summarily dismissed. Moreover, many DSGE models imply exact relationships between endogenous variables that are not supported by the data. Finally, a low number of shocks can be preferable, not only to reduce the computational burden, but also because a common challenge of DSGE modeling is how to credibly introduce several shocks, needed to avoid the stochastic singularity problem.

[3] Perhaps implicitly, this is the interpretation that currently applies to most applications relying on models with heterogeneous agents and aggregate uncertainty. The endogenous variables are typically obtained from arbitrarily long simulations (after an indispensable burn-in), instead of considering the actual data series length. Since these models have heterogeneous agents, the simulations are often performed with large synthetic panel data, with a cross sectional dimension that does not match the size of real longitudinal datasets. This comment is all the more true when the model's solution does not rely on simulations, and the aggregate law of motion is solved for with projection methods, as described in den Haan (2010).

have remarkably different implications.

The rest of the paper is organized as follows. Section 2 briefly presents the models and reports some of their quantitative implications. Section 3 discusses the choice of prior distributions for the parameters. Section 4 outlines the empirical methodology. Section 5 presents the results, while Section 6 concludes. Several on-line appendices discuss in more detail the (numerical and empirical) methods used, and present additional results.

2 Preference Heterogeneity in Macroeconomic Models with Aggregate Risk

I consider four versions of the incomplete markets model with heterogeneous agents and aggregate risk to address whether preference heterogeneity (in discount factors and/or risk aversions) helps accounting for the correlation between aggregate consumption and income, the autocorrelation of aggregate consumption, and the relative standard deviation of consumption and income. The four models differ in the degree and nature of household heterogeneity. The starting point is the framework proposed by Krusell and Smith (1998), and I am going to consider both their baseline model (denoted as $M1_{(IM)}$) and the extension with preference heterogeneity in the discount factors β (denoted as $M2_{(\beta)}$). The other two models introduce another layer of heterogeneity, allowing agents to differ also in their preferences for risk. One version of the model focuses only on heterogeneity in risk aversion γ (denoted as $M3_{(\gamma)}$), while another version assumes that preferences differ both in the relative risk aversion, and in the degree of patience (denoted as $M4_{(\beta,\gamma)}$).[4] For ease of comparison, the complete markets counterpart of the basic version of the model without preference heterogeneity (denoted as $M0_{(CM)}$) is also included in the analysis.

[Table 1 about here]

Table 1 lists the five specifications of the models that are going to be studied. The simplest incomplete markets framework coincides with the standard model introduced by Krusell and Smith (1998), the only difference being the formulation of the income received by the unemployed. Following den Haan, Judd, and Juillard (2010), I assume the existence of a budget-balanced Unemployment Insurance (UI) scheme, which raises contributions by taxing employed workers and distributes unemployment benefits to jobless workers.

2.1 Models Set-up

Time is discrete. The models assume a production economy with aggregate risk, such that productivity shocks hit the economy every period, inducing aggregate fluctuations. The economy is populated by a measure one of infinitely-lived agents subject to idiosyncratic risk. Agents face different employment histories (idiosyncratic

[4] In Cozzi (2014), I show that a model with risk aversion heterogeneity and endogenous sorting into risky jobs accounts for many features of the U.S. wealth distribution. That model, however, abstracts from aggregate uncertainty and time-varying preference heterogeneity.

shocks are correlated with aggregate shocks), and self-insure by accumulating a single risky asset. An exogenous borrowing constraint ($a_{\min}$) hampers the households' ability to smooth consumption.

Technology: Production is modeled as a constant returns to scale technology of the Cobb-Douglas form, which relies on aggregate capital K_t and labor L_t to generate final output $Y_t = z_t K_t^{\alpha} L_t^{1-\alpha}$. The aggregate shock takes only two values: $z_t = \{z_G, z_B\}$, where the index G (B) denotes booms (recessions), and $z_G > z_B$. The aggregate shock follows a symmetric Markov chain. Firms hire workers from a competitive labor market. Total labor services are $L_t = lN_t$, namely they are the product of the share (l) of the time endowment (normalized to 1) devoted to market activities and the employment level N_t. Firms rent capital from a competitive asset market, and this input depreciates at the exogenous rate δ. The firm's first order conditions to the profit maximization problem give the expressions for the net real return to capital r_t and the wage rate w_t:

$$r_t = \alpha z_t \left(\frac{L_t}{K_t}\right)^{1-\alpha} - \delta, \tag{1}$$

$$w_t = (1-\alpha)\, z_t \left(\frac{K_t}{L_t}\right)^{\alpha}. \tag{2}$$

Government: The government taxes the labor income of employed agents at rate τ_t to finance a budget-balanced UI scheme. Unemployed agents receive UI benefits equal to a fixed replacement rate ϕ of the going labor income. Since labor supply is fixed, and the aggregate unemployment rate can only take two values (u_G when $z_t = z_G$ and u_B when $z_t = z_B$), the equilibrium tax rate is $\tau_t = \phi(1 - N_t)/N_t$, with $N_t = 1 - u_t$.

Households: Preferences are represented by a time-separable utility function $U(.)$. Every household $i \in [0,1]$ chooses consumption ($c_{i,t}$) and future asset holdings ($a_{i,t+1}$) to maximize their objective function:

$$\max_{\{c_{i,t}, a_{i,t+1}\}_{t=0}^{\infty}} \mathbb{E}_0 \sum_{t=0}^{\infty} \beta_{i,t}^t \frac{c_{i,t}^{1-\gamma_{i,t}} - 1}{1-\gamma_{i,t}}$$

where $\mathbb{E}_0$ is the expectation operator. In the simplest set-up, preference parameters are homogeneous, and all agents share the same discount factor β and the same risk aversion γ. In general, these parameters will differ across agents and they are indexed by i to highlight this possibility. Furthermore, they are indexed by t to indicate that they can potentially evolve over time. $\beta_{i,t} \in (0,1)$ is the agents' discount factor, and in models $M2_{(\beta)}$ and $M4_{(\beta,\gamma)}$ it can take up to three different values, $\beta_{i,t} \in \{\beta_l, \beta_m, \beta_h\}$, with $\beta_l < \beta_m < \beta_h$. In these cases, each agent's discount factor can vary over time according to a three-state Markov chain. Similarly, in models $M3_{(\gamma)}$ and $M4_{(\beta,\gamma)}$, the risk aversion $\gamma_{i,t} > 0$ can take up to three different values $\gamma_{i,t} \in \{\gamma_l, \gamma_m, \gamma_h\}$, with $\gamma_l < \gamma_m < \gamma_h$. Risk aversion can also vary over time according to a three-state Markov chain.

Agents can be employed ($s = e$) or unemployed ($s = u$). The employment probabilities follow a first-order Markov process, and depend on both the idiosyncratic employment status (s) and on the aggregate state of the economy (z). I use recursive methods to solve the model, and the value function associated with this problem is denoted with $V(a, \beta, \gamma, s, z, K)$. This represents the expected lifetime utility of an agent whose current asset holdings are equal to a, whose current discount factor is β, whose current risk aversion is γ, whose current employment status is s, facing the aggregate shock z and in an economy with K units of aggregate capital. The Bellman equation is:

$$V(a,\beta,\gamma,s,z,K) = \max_{c,a'} \left\{ \frac{c^{1-\gamma}-1}{1-\gamma} + \beta E_{\beta',\gamma',s',z'|\beta,\gamma,s,z} V(a',\beta',\gamma',s',z',K') \right\}$$

$$s.t.$$

$$c + a' = (1+r)\,a + (1-\tau)\,wl, \text{ if } s = e$$

$$c + a' = (1+r)\,a + \phi wl, \text{ if } s = u$$

$$c \geq 0, \quad a' \geq a_{\min}$$

$$\ln K' = \theta_{0,j} + \theta_{1,j} \ln K, \text{ if } z = z_j, j = \{G, B\} \tag{3}$$

The appropriate Markov chains for β, γ, s and z

Agents optimally set their consumption/savings plans. They enjoy utility from consumption and face several uncertain events in the future. Notice that, according to the algorithm that I use to solve this model, the relevant state variable in the agents' problem is just aggregate capital K, rather than the whole endogenous distribution over idiosyncratic states. Agents forecast future prices relying on the (equilibrium) evolution of the aggregate capital stock, the Aggregate Law of Motion (ALM) being specified as the pair of equations (3). Moreover, every version of the model will include the laws of motion (i.e., the Markov chains) for the evolution of the exogenous stochastic state variables (β, γ, s and z) that apply to each specific case.

2.2 Long-Run Welfare Effects of Eliminating Business Cycles and Liquidity Constrained Households

A substantive issue is whether the introduction of risk aversion heterogeneity matters for the determination of macroeconomic outcomes. In this subsection, I focus on studying the welfare costs of aggregate fluctuations and the prevalence of liquidity-constrained households. Because of the intractable computational burden, it is not feasible to compute these statistics for all model parameterizations discussed in Section 3. Instead, for every model, I compute them for a benchmark case (i.e., at the average of the parameter priors postulated below).

The first macroeconomic outcome of interest is the welfare cost of aggregate fluctuations. Adapting the methods outlined in Krusell and Smith (1998), Mukoyama and Sahin (2006) and Krusell, Mukoyama, Sahin, and Smith (2009), I compute the distribution of the long-run welfare effects of eliminating business cycles, where each welfare effect depends on a specific pair of aggregate and individual states. The welfare effects are obtained by applying the integration principle, which eliminates the correlation between the aggregate and idiosyncratic shocks.[5] The welfare effects are expressed as the consumption percentage change (denoted as CEV) in all states of the world that would make the welfare in the economy with aggregate risk equal to the welfare in the economy without aggregate risk. The average welfare effects in the four models with incomplete markets are: $M1_{(IM)} = 0.12\%$, $M2_{(\beta)} = -1.05\%$, $M3_{(\gamma)} = -1.01\%$, $M4_{(\beta,\gamma)} = -62.21\%$.

[Figure 1 about here]

Figure 1 plots the welfare effects densities, omitting the distribution for model $M1_{(IM)}$, as its average is close to zero and it displays limited dispersion. The long-run welfare effects for models $M2_{(\beta)}$ and $M3_{(\gamma)}$ can be quantitatively important (a loss in excess of 3%), possess substantial dispersion, and their distributions are relatively similar. An extremely different outcome is obtained when considering model $M4_{(\beta,\gamma)}$, whose average long-run welfare effect has a stunning value. The reason behind this large cost of eliminating aggregate fluctuations stems from the role of precautionary savings. In this economy, wealth is very concentrated and it is mostly kept by the very risk-averse and very patient households. The elimination of aggregate risk leads to a large adjustment in these households' wealth holdings. This triggers a general equilibrium effect, which increases the long-run (annualized) interest rate considerably (eventually changing from 4.0% to 24.8%). The implied income effect further decreases savings, leading to a 77% drop in the aggregate capital stock. The long-run fall in aggregate income and wages explains the spectacular welfare cost of eliminating business cycles.[6] The analysis of the long-run welfare effects provides a strong reason to compare formally the fit of these models, as their welfare implications can be radically different.

Liquidity Constrained Households: Table 2 reports the percentage of households that are liquidity constrained, which is another statistic that is important for the design of macroeconomic policy. First, I consider a broad measure of "Hand-to-Mouth" (HtM) households. I then compute a measure of "Wealthy Hand-to-Mouth" (W-HtM) households that is closer in spirit to the one used by Kaplan, Violante, and Weidner (2014).

[Table 2 about here]

The share of HtM households is defined as the percentage of households that consume all of their income, or more. All models have a sizable share of households that behave in a HtM fashion, which is consistent with the empirical evidence. It is also worth reporting that this variable displays substantial volatility over the business cycle.

[5] In particular, the recursive formulation of the household problem has to be modified, by both introducing additional idiosyncratic states and amending the related Markov chain. For more details, see Appendix E.

[6] These computations focus on a comparison with the long-run equilibrium, such that the economy without aggregate fluctuations is in the steady state. The positive effects on welfare of the transitional dynamics are neglected by construction, which explains why the results for model $M2_{(\beta)}$ differ from the ones reported by Krusell, Mukoyama, Sahin, and Smith (2009). However, computing this transition is not straightforward, and a more complete welfare analysis deserves a paper in its own right.

There are several factors that affect the prevalence of HtM behavior. During booms, an important one is the employment status, as the unemployed individuals rely on the totality of their current income and some accumulated wealth to prevent their consumption profile from falling excessively.

Generally speaking, in this class of models unemployed individuals tend to display HtM behavior. This applies unless they are extremely rich, holding more than ten times the average wealth, which is a low probability event. In the model without preference heterogeneity, the saving decisions are such that during booms only the unemployed agents run down their wealth. This explains why the related number is exactly 4.0%. In the other versions of the model, some employed individuals that are (temporarily) impatient and/or with low risk aversion are also HtM. During recessions, and whenever the aggregate capital is below its long run value, a sizable share of employed individuals display HtM behavior.

The results show that preference heterogeneity has an important effect in determining this statistic. In particular, the share of HtM households in a recession increases substantially, with the models with risk aversion heterogeneity attaining the largest shares. This finding is important, as the effectiveness of fiscal policies and stimulus-style interventions depends crucially on the marginal propensity to consume.

Since the models that I consider do not have two assets, I cannot apply the same W-HtM definition as in Kaplan, Violante, and Weidner (2014). In order to approximate their concept, I apply an adjustment to the models' asset holdings. Namely, I truncate the wealth distribution, selecting the truncation point such that the truncated distribution matches the median wealth/income ratio reported in Table 2 of Kaplan, Violante, and Weidner (2014). I then consider the share of households whose assets lie in the support of the truncated distribution that consume all of their income (or more). This statistic is what I consider as the measure of W-HtM households.[7] In the two models that allow for this correction, the adjustment reduces the share of liquidity constrained individuals. In model $M2_{(\beta)}$, about 76% of HtM households are W-HtM, and the average share of liquidity constrained households is 15.5%. In model $M4_{(\beta,\gamma)}$, almost all HtM households are W-HtM. These values are roughly consistent with the estimates reported in Kaplan, Violante, and Weidner (2014).

3 Parameterizing the Models

The issue of parameter uncertainty is ubiquitous in quantitative macroeconomics. The models' level of abstraction frequently leads to mismatches between model variables and their empirical counterparts. For instance, think of capital: housing represents a large component of household wealth, and it is usually included in the value of the capital stock, which affects its depreciation rate, but to what extent can it be considered a factor of production? Parameter uncertainty can be easily accommodated by specifying prior distributions for the parameters, embracing one of the building blocks of Bayesian empirical analysis.

As for the specification of the priors, I am going to consider four different cases, listed in Table 1. The simplest case assumes independent uniform priors (Case 2), reflecting the idea that a-priori information can only provide parameter bounds. The other cases consider correlated priors (Case 1), introduce curvature using Beta priors

[7] Notice that this statistic can be computed only for two versions of the model, as in models $M1_{(IM)}$ and $M3_{(\gamma)}$ the median wealth/income ratio from the un-truncated distribution is already larger than the Kaplan, Violante, and Weidner (2014) figures.

(Case 3), and change the bounds of the heterogeneous discount factors and risk aversions, by increasing their range (Case 4).

[Table 3 about here]

3.1 Parameter Bounds common across all Models

The bounds of the prior distributions that are going to be used are reported in Table 3. The top of the Table focuses on the parameters that are common across all models (apart from $a_{\min}$, which is present only under incomplete markets).

The bounds for the capital share α are chosen on the basis of the labor share values found in the Penn World Tables 8.0. In the period $1950-2011$ the labor share has fluctuated between 62% and 68%. Since the downward trend in the labor share has been very persistent, I use an upper bound for the capital share of 40%. Moreover, since most studies in the RBC literature use a capital share of 36%, it seems appropriate to center its range at this value.[8]

The model considers only one type of capital, hence its depreciation rate δ is the weighted average of the depreciation rates of many different capital goods. A range for δ between 8% and 10% on an annual basis spans many of the estimates available in the literature.

The borrowing limit $a_{\min}$ ranges between $\underline{a_{\min}} = 0$, which is the most extreme case of market incompleteness such that borrowing is not allowed, and $\underline{a_{\min}} = -2$, which is approximately (minus) twice the average quarterly income.

As for the range related to the exogenous labor supply l, Juster and Stafford (1991) documented that the share of time devoted to market activities, averaged between males and females, is approximately 32%. The midpoint of the range is then set to this value, $l = 0.32$. Since the consensus is that answers to the time allocation survey questions are contaminated by measurement error, I consider a range of $\pm 10\%$ around this estimate.

For comparability with Krusell and Smith (1998), I work with their parameterization for the (average) values of the unemployment rate during booms (u_G) and recessions (u_B): $u_B = 10\%$ and $u_G = 4\%$. For both parameters, I consider a range which is ± 1 percentage point around these averages. In order to understand whether these bounds are plausible, I split the Bureau of Labor Statistics data on the monthly unemployment rates for the $1948M1 - 2019M12$ period into two groups, one group with the observations above the average unemployment rate of 5.7%, and the other group with the observations below it. The $3\% - 5\%$ range corresponds to the $3rd - 63rd$ percentiles of the unemployment rates below the average, while the $9\% - 11\%$ range corresponds to the top 11% of the unemployment rate distribution for the unemployment rates above the average (with 10.8%

[8]For this statistic, I do not use the more recent versions of the Penn World Tables because they feature a sizable downward revision of the U.S. labor share (larger than 2 percentage points), making it inconsistent with the value of the capital share typically used in the literature.

being the highest recorded unemployment rate).[9]

The bounds for the remaining parameters, which are also listed in Table 3, are model-specific and are discussed next.

3.2 Model-specific Parameters and their Bounds

The parameters whose bounds still need to be discussed are the discount rate and the risk aversion.[10] Because of data limitations, discount rate heterogeneity must be considered as unobserved heterogeneity. Differently, if one is willing to make a structural distributional assumption, (relative) risk aversion heterogeneity can be obtained from the data, exploiting questions on attitudes towards risk currently included in a number of large datasets, such as the PSID and the Health and Retirement Study (HRS).

In model $M1_{(IM)}$, the β and γ parameters are the same for all agents. In this case the range for γ is $[1.0, 3.0]$, which spans most of the available estimates of the inverse of the elasticity of intertemporal substitution, obtained with a variety of methods and surveyed by Attanasio and Weber (2010), among others. Choosing bounds for the subjective discount factor is more complicated, because there is not much direct information regarding this parameter. Exploiting its relationship with the interest rate, I work with a range for β equal to $[0.985, 0.995]$. The rationale being that the equilibrium annual interest rate in the steady-state of the corresponding complete markets economy is between 0.5% and 6%, and it is well-known that in the benchmark Krusell and Smith (1998) model precautionary savings have a limited quantitative importance.

In model $M2_{(\beta)}$, the range for γ is still $[1.0, 3.0]$, while there are three time-varying discount factor types. In this case, the range for β_l is $[0.9848, 0.9868]$, the range for β_m is $[0.9884, 0.9904]$, and the range for β_h is $[0.992, 0.994]$. The three ranges are chosen for the discount factors not to overlap, an assumption that will be relaxed in Case 4. Furthermore, since in Cozzi (2015) I found that this economy solved at the priors' averages matches a wealth Gini index of 0.8 and attains an average annualized interest rate of 4%, I consider fairly tight bounds around that calibration.

Models $M3_{(\gamma)}$ and $M4_{(\beta,\gamma)}$ introduce heterogeneity in the agents' preferences for risk.

Parameterizing Risk Aversion Heterogeneity: Thanks to hypothetical lottery questions included in the PSID in 1996, this can be dealt with as observed heterogeneity. In particular, I apply the procedure proposed by Kimball, Sahm, and Shapiro (2008). With appropriate statistical methods, and postulating that γ is log-

[9] This statistic, hence the choice of the range for the unemployment rate in recessions, might seem extreme. However, it should be interpreted as including the marginally attached workers, an enlarged concept of the unemployment rate for which consistent measurements are available only since 1994 (labeled U5RATE in the FRED database). In the above average unemployment rate sample, this alternative unemployment rate has been 1.3 percentage points above the standard unemployment rate. Considering this adjustment, the $7.7\% - 9.7\%$ unemployment rate range corresponds to the $73rd - 95th$ percentiles, which is somewhat less extreme.

[10] The probabilities in the Markov chains also vary across replications. In order to guarantee that they all stay non-negative, and add up to one, I consider a tight range of approximately $\pm 1\%$ around the calibrations in Krusell, Mukoyama, Sahin, and Smith (2009) and Cozzi (2015). The probabilities in the risk aversion Markov chain are taken from the bootstrapped estimates discussed below.

normally distributed in the population, I find that the estimated distribution is $LN\left(\mu_{\gamma}=1.07, \sigma_{\gamma}^{2}=0.76\right)$.[11] Although the true underlying distribution is log-normal, it is feasible to consider only three risk aversion types.[12] Therefore, I partition the PSID data in three groups with mass 20%, 30% and 50%. The conditional averages of γ for each of these subgroups are then obtained exploiting the CDF of the true distribution. This step delivers the three desired values for the risk aversion types: $\gamma_l = 0.92, \gamma_m = 2.12$, and $\gamma_h = 7.55$. There are some additional facts about the risk aversion distribution that are going to guide the next modeling choices. Since Kimball, Sahm, and Shapiro (2008) document that measurement error in the HRS is large, I consider fairly wide ranges for each parameter, which are going to be further increased in Case 4. The range for γ_l is $[0.62, 1.22]$, the range for γ_m is $[1.82, 2.42]$, and the range for γ_h is $[7.25, 7.85]$. Another fact is that the distribution of answers to the lottery questions in the HRS is virtually stationary over time. It is therefore appropriate to assume that the evolution of the risk aversion is governed by a time-invariant Markov chain.

Estimating the Risk Aversion Markov Chain: I obtain an estimate for the Markov chain probabilities by using a moment-matching procedure. I start by imposing appropriate restrictions on the structure of the Markov chain, so that it is parameterized by only three probabilities: $p_{\gamma_l}, p_{\gamma_m}$, and p_{γ_h}.[13] I can then use the observed cross sectional distribution of risk preferences to identify two of these probabilities, while the third one is backed out to replicate the intergenerational correlation between the risk aversion of parents and their children documented by Kimball, Sahm, and Shapiro (2009). The Markov chain (Π) that I have to recover from the data is:

$$\Pi \equiv \pi\left(\gamma_i, \gamma_j'\right) = \begin{bmatrix} p_{\gamma_l} & 1-p_{\gamma_l} & 0 \\ p_{\gamma_m} & 1-2p_{\gamma_m} & p_{\gamma_m} \\ 0 & 1-p_{\gamma_h} & p_{\gamma_h} \end{bmatrix}, \text{ with } i, j \in \{l, m, h\}$$

The stationarity of the risk aversion distribution allows me to consider the system of linear equations $\mu_\gamma^* = \Pi' \mu_\gamma^*$, where $\mu_\gamma^* = \left[\mu_{\gamma_l}^*, \mu_{\gamma_m}^*, 1-\mu_{\gamma_l}^* - \mu_{\gamma_m}^*\right]$ denotes the stationary distribution over the vector of risk aversions $[\gamma_l, \gamma_m, \gamma_h]$. Since the entries in μ_γ^* must add up to one, I can use only two fractions from the data to assign values to the Markov chain probabilities: $\widehat{\mu}_{\gamma_l}^* = 0.201$ and $\widehat{\mu}_{\gamma_m}^* = 0.303$.

Exploiting the stationarity condition, it is possible to write the following system of equations:

[11] Fore more details, see Cozzi (2014). In that model, I work with the assumption that preference types are permanent, so it is worthwhile considering the implications of time varying risk aversions.

[12] First, there is a data limitation issue that will be discussed below. Moreover, this model is computationally costly, and I need to simplify the problem in order to solve it many times (I draw 1,000 combinations of parameters from their priors).

[13] Note that other parameterizations of the CRRA Markov chain suffer from identification issues. For example, it is easy to show that a chain that imposes the additional restriction $p_{\gamma_h} = p_{\gamma_l}$ cannot be identified. Similarly, more general chains with four or more unknown probabilities would require more pieces of information from the data and cannot be recovered.

$$\begin{cases} \mu^*_{\gamma_l} = p_{\gamma_l}\mu^*_{\gamma_l} + p_{\gamma_m}\mu^*_{\gamma_m} \\ \\ \mu^*_{\gamma_m} = \left(1-p_{\gamma_l}\right)\mu^*_{\gamma_l} + \left(1-2p_{\gamma_m}\right)\mu^*_{\gamma_m} + \left(1-p_{\gamma_h}\right)\mu^*_{\gamma_h} \\ \\ \mu^*_{\gamma_h} = p_{\gamma_m}\mu^*_{\gamma_m} + p_{\gamma_h}\mu^*_{\gamma_h} \end{cases}$$

which, for given $\widehat{p}_{\gamma_m}, \widehat{\mu}^*_{\gamma_l}$ and $\widehat{\mu}^*_{\gamma_m}$, leads to the following expressions for the two remaining probabilities:

$$\begin{cases} \widehat{p}_{\gamma_l} = 1 - \widehat{p}_{\gamma_m}\frac{\widehat{\mu}^*_{\gamma_m}}{\widehat{\mu}^*_{\gamma_l}} \\ \\ \widehat{p}_{\gamma_h} = 1 - \widehat{p}_{\gamma_m}\frac{\widehat{\mu}^*_{\gamma_m}}{\widehat{\mu}^*_{\gamma_h}} \end{cases} \tag{4}$$

The estimation procedure then relies on a simulated minimum distance estimator. The following steps are considered: a) construct a fine grid for p_{γ_m}, b) consider one grid-point at a time, whose value is denoted with $\widetilde{p}_{\gamma_m}$, c) get $\widetilde{p}_{\gamma_l}$ and $\widetilde{p}_{\gamma_h}$ from system (4), d) simulate the implied Markov chain for 200 periods (50 years $\times$ 4 quarters per year) with 100,000 agents (i.e., risk aversion types), d) compute the intergenerational correlation ρ^{Π}_{γ}, e) select $\widehat{p}_{\gamma_m}$ as $\widehat{p}_{\gamma_m} = \underset{\widetilde{p}_{\gamma_m}}{Arg\min}\left[\rho^{\Pi}_{\gamma}\left(\widetilde{p}_{\gamma_m}\right) - \rho^{PSID}_{\gamma}\right]^2$. The procedure leads to the following estimates for the Markov chain probabilities: $p_{\gamma_h} = 0.9948, p_{\gamma_m} = 0.0085$ and $p_{\gamma_l} = 0.9872$.[14]

Finally, choosing bounds for the priors of the CRRA Markov chain probabilities boils down to choosing a range for p_{γ_m}. This is obtained by relying on the empirical distribution of $1,000$ minimum distance estimates obtained by simulating the chain with $5,662$ artificial agents, which corresponds to the cross sectional dimension of the household heads in the 1996 PSID sample that provided an answer to the lottery questions.

The last case left to consider is model $M4_{(\beta,\gamma)}$, which allows for both types of preference heterogeneity. As for the heterogeneity in the discount factor, I use the same specification as in $M2_{(\beta)}$, namely three discount factor types that evolve over time through a Markov chain. In this case, the range for β_l is $[0.980, 0.982]$, the range for β_m is $[0.984, 0.986]$, and the range for β_h is $[0.992, 0.994]$. In this model, the role of precautionary savings is more pronounced, explaining the downward shift in the ranges for the first two types. Lower discount factors are imposed for the model to display average interest rates that are similar to the other models. As for the risk aversion types, the priors are kept the same as in model $M3_{(\gamma)}$. The ranges of all preference parameters will be increased in Case 4.

3.3 Further Details on the Choice of Priors

The information provided above fully characterizes the case postulating uncorrelated Uniform priors, while the other ones need some additional details. In both the case with correlated Uniform priors and independent Beta priors, I retain the same bounds listed above. For the correlated Uniform priors, the correlation matrix is obtained as follows. From the Penn World Tables, I compute the correlations between the labor share, the real rate of return, and the capital depreciation rate. These values are used to pin down the correlation matrix between α, β (or $\beta_l, \beta_m, \beta_h$), and δ. Since borrowing constraints are less likely to be binding when workers

[14] As shown in Figure 7 in Appendix G, given two of the three probabilities, the remaining one is uniquely identified.

face better labor market conditions, the $a_{\min}$ and l priors are assumed to have a negative correlation equal to -0.25. Similar considerations lead me to set the correlation between both the $a_{\min}$ and u_G priors and the $a_{\min}$ and the u_B priors to 0.25. Since hours worked are lower when the unemployment rate is higher, I set the correlation between both the l and u_G priors and the l and the u_B priors to -0.25. Finally, lacking sound evidence on how risk aversion correlates with the other parameters, the γ (or $\gamma_l, \gamma_m, \gamma_h$) priors are assumed to be independent. In the Beta priors case, I assume that the densities are symmetric around their mean, with both shape parameters set to 5. To conclude, the case with overlaps in the heterogeneous preference parameters uses the same midpoints reported above, but increases the range of these parameters only.[15]

4 The Empirical Framework

At this stage, the analysis does not have explicit empirical content, because the models provide quantitative assessments about population moments. Moreover, for a given parameterization, each model delivers a point for every moment of interest. By assuming priors for the parameters one obtains distributions for the moments, which are generated by parameter uncertainty. DeJong, Ingram, and Whiteman (1996) stress that the framework needs another element to be able to link the models' implications to the information contained in the data. This link takes the form of a reduced-form Bayesian econometric model, which allows to relate the observables to the population moments. The interaction between these two elements of the MEI equips the researcher with an empirical framework where formal model comparison can be undertaken.[16]

Since the main focus of heterogeneous-agent models is to carefully micro-found the determination of consumption/saving decisions, I am going to assess their performance in three dimensions, all pertaining to the behavior of aggregate consumption and its relationship with income.

Data: The time period is $1947Q1 - 2019Q4$, and the number of observations is 292. As customary in the RBC literature, I consider as aggregate consumption ($\tilde{C}_t$) the sum of the expenditures on Services and Non-durables.[17] Aggregate output ($\tilde{Y}_t$) is defined as the sum of Services, Non-durables, and Investment, because all model economies are closed, and there is no public expenditure in the form of government consumption. Since it is well-known that the detrending methodology can heavily affect the business cycle statistics, the $\tilde{Y}_t$

[15]The upper and lower bounds are obtained as follows. In model $M2_{(\beta)}$, the midpoint of β_l is increased (decreased) by 0.0016 (0.001), the midpoint of β_m is increased (decreased) by 0.0016 (0.0016), and the midpoint of β_h is increased (decreased) by 0.001 (0.0016). In both model $M3_{(\gamma)}$ and $M4_{(\beta,\gamma)}$, the midpoint of γ_l is increased (decreased) by 0.55 (0.3), the midpoint of γ_m is increased (decreased) by 2.5 (0.55), and the midpoint of γ_h is increased (decreased) by 0.3 (2.5). In model $M4_{(\beta,\gamma)}$, the midpoint of β_l is increased (decreased) by 0.0016 (0.001), the midpoint of β_m is increased (decreased) by 0.003 (0.0016), and the midpoint of β_h is increased (decreased) by 0.001 (0.003).

[16]An additional appealing feature of this approach is that it treats symmetrically the uncertainty in the moments of both the theoretical model and the data. In the MEI framework, prior distributions are specified for both the theoretical and the empirical models, which induce probability distributions over moments of interest. The degree of overlap between these distributions is then used to assess the models' fit.

[17]Notice that I exclude durables from the analysis. A possible alternative could be to include durables in investment. Although the empirical results are fairly similar, this procedure is inconsistent with the model's assumption that total investment contributes to the accumulation of capital.

and $\tilde{C}_t$ series are detrended with two different filters. I consider both the one-sided HP filter of Stock and Watson (1999) and the BN filter of Kamber, Morley, and Wong (2018), which is based on a Beveridge-Nelson decomposition.[18]

A Bayesian VAR(1): In this framework, there is the need to link the distributions of population moments to the observables. In order to study the empirical characteristics of the relevant moments of consumption and income, I specify a Bayesian VAR(1) process on detrended aggregate income Y_t and detrended aggregate consumption C_t. Diffuse priors are assumed for all the parameters.

$$\begin{pmatrix} Y_t \\ C_t \end{pmatrix} = \begin{bmatrix} d_{11} & d_{12} \\ d_{21} & d_{22} \end{bmatrix} \begin{pmatrix} Y_{t-1} \\ C_{t-1} \end{pmatrix} + \begin{pmatrix} \eta_t^Y \\ \eta_t^C \end{pmatrix}$$

$$\begin{pmatrix} \eta_t^Y \\ \eta_t^C \end{pmatrix} \overset{iid}{\sim} N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \Sigma = \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{bmatrix} \right)$$

With a more compact notation, the VAR can be represented as $\mathcal{Y}_t = D\mathcal{Y}_{t-1} + \mathcal{E}_t$, where $\mathcal{Y}_t$ is a $T \times 2$ matrix, whose columns are Y_t and C_t, D is the 2×2 matrix of parameters to be estimated, and $\mathcal{E}_t$ is the $T \times 2$ matrix of innovations. The error terms η_t^Y and η_t^C are assumed to be normally distributed, with a zero mean and a variance/covariance matrix denoted by Σ.

It is well-known, see for example Koop and Korobilis (2009), that in VAR models with normally distributed shocks the flat priors assumption implies that the posteriors' marginal distributions are multivariate t distributions centered around the OLS estimates. However, the estimation for all seven parameters $(d_{11}, d_{12}, d_{21}, d_{22}, \sigma_{11}, \sigma_{12}, \sigma_{22})$ is accomplished with a posterior simulator. Simulation methods are needed because the objects of interest are not the posteriors of the parameters, but functions of them (i.e., some moments), and I draw from their posteriors.[19]

The incomplete markets models are evaluated by comparing the draws from the joint posteriors of the autocorrelation of consumption (m_{ρ_C}), the correlation between consumption and income ($m_{\rho_{CY}}$), and the relative standard deviation of consumption and income ($m_{sd_{CY}}$), all obtained from the Bayesian VAR, with the draws for the same (population) moments predicted by the theoretical models.

[18] Following Kamber, Morley, and Wong (2018), I use both a low signal-to-noise ratio ($\bar{\delta}_{BN} = 0.21$) and a Bayesian AR(12) forecasting model with a "Minnesota" shrinkage prior (on the second-difference coefficients). This formulation is desirable because it provides estimates of the cyclical component that are highly persistent and large in amplitude. As a robustness check, I also detrended the time series data with a two-sided HP filter with a smoothing parameter of $1,600$, finding that the ranking of the models in terms of their fit is preserved.

[19] The three moments' posteriors are plotted in Figure 3 in Appendix G, while Appendix C provides more details on how the computations are actually implemented. DeJong and Ripoll (2007) use a similar methodology, relying on the predictive distributions to perform model comparison.

5 Results

I begin this section by discussing the model comparison based on the MEI approach. I then present a Monte Carlo study examining the MEI performance, which is followed by an extension to the MEI that allows to exploit both macro and micro-economic data.

5.1 Model Comparison

The first step consists of computing the models' Log Marginal Likelihoods (LML), which are reported in Table 4. These are then exploited to derive the Bayes factors (BF), which are used to perform the model comparison.[20]

[Table 4 about here]

The LML results show that, even when using macro-economic moments only, models with preference heterogeneity dominate the other ones in terms of fit. Irrespective of the filtering method, models $M3_{(\gamma)}$ and $M4_{(\beta,\gamma)}$ attain the highest LML. The models without preference heterogeneity display the lowest LML, suggesting that preference heterogeneity provides a quantitatively important improvement in accounting for the consumption dynamics at business cycle frequencies. In particular, the major failure of models $M0_{(CM)}$ and $M1_{(IM)}$ is with respect to the relative standard deviation of consumption and income, as their distributions are highly concentrated around their means. These two models imply excessive consumption smoothing, as they understate the role of wealth inequality and of binding borrowing constraints. As for the different cases for the priors, the values of the LML are quite stable across them.[21]

The Bayes factors can be computed for any pair of models M_i and M_j $(i, j = 0, ..., 4)$, and they are the ratio of two marginal likelihoods, $P_1(M_i|data, E_1)$ and $P_1(M_j|data, E_1)$, whose expressions are:

$$BF_1 \equiv \frac{P_1(M_i|data, E_1)}{P_1(M_j|data, E_1)} = \frac{P_1(M_i|E_1)P_1(data|M_i, E_1)}{P_1(M_j|E_1)P_1(data|M_j, E_1)}$$

$$= \frac{P_1(M_i|E_1)\int\int\int P_1(m_{\rho_C}, m_{\rho_{CY}}, m_{sd_{CY}}|M_i)P_1(data|m_{\rho_C}, m_{\rho_{CY}}, m_{sd_{CY}}, E_1)dm_{\rho_C}dm_{\rho_{CY}}dm_{sd_{CY}}}{P_1(M_j|E_1)\int\int\int P_1(m_{\rho_C}, m_{\rho_{CY}}, m_{sd_{CY}}|M_j)P_1(data|m_{\rho_C}, m_{\rho_{CY}}, m_{sd_{CY}}, E_1)dm_{\rho_C}dm_{\rho_{CY}}dm_{sd_{CY}}}$$

$$\propto \frac{\int\int\int P_1(m_{\rho_C}, m_{\rho_{CY}}, m_{sd_{CY}}|M_i)P_1(m_{\rho_C}, m_{\rho_{CY}}, m_{sd_{CY}}|data, E_1^*)dm_{\rho_C}dm_{\rho_{CY}}dm_{sd_{CY}}}{\int\int\int P_1(m_{\rho_C}, m_{\rho_{CY}}, m_{sd_{CY}}|M_j)P_1(m_{\rho_C}, m_{\rho_{CY}}, m_{sd_{CY}}|data, E_1^*)dm_{\rho_C}dm_{\rho_{CY}}dm_{sd_{CY}}}$$

$$\approx \frac{\frac{1}{N_{M_i}N_{E_1^*}}\sum_{u=1}^{N_{M_i}}\sum_{v=1}^{N_{E_1^*}} K_1\left(m_{\rho_{C,u}}^{M_i}, m_{\rho_{CY,u}}^{M_i}, m_{sd_{CY,u}}^{M_i}; m_{\rho_{C,v}}^{E_1^*}, m_{\rho_{CY,v}}^{E_1^*}, m_{sd_{CY,v}}^{E_1^*}\right)}{\frac{1}{N_{M_j}N_{E_1^*}}\sum_{u=1}^{N_{M_j}}\sum_{v=1}^{N_{E_1^*}} K_1\left(m_{\rho_{C,u}}^{M_j}, m_{\rho_{CY,u}}^{M_j}, m_{sd_{CY,u}}^{M_j}; m_{\rho_{C,v}}^{E_1^*}, m_{\rho_{CY,v}}^{E_1^*}, m_{sd_{CY,v}}^{E_1^*}\right)}$$

[20] BF_1 (BF_2) is used to denote the Bayes factors computed on the macro (micro) moments, and a similar notation is also used to denote the other variables in the formulas, such as the densities and marginal likelihoods.

[21] For Case 2, the three distributions of the model-generated moments are plotted in Figure 4 in Appendix G.

Posterior odds ratios, denoted as $P_1(M_i|data, E_1)/P_1(M_j|data, E_1)$, are a common tool in Bayesian empirical work. Notice how these depend on both the data and an incomplete econometric model denoted as E_1. This model is needed to provide the framework with empirical content. In particular, E_1 specifies a conditional distribution of observables $P_1(Y, C|m_{\rho_C}, m_{\rho_{CY}}, m_{sd_{CY}}, \theta, E_1)$ together with a prior for the parameters θ in the econometric model. This econometric model is incomplete, because it does not provide a joint prior distribution for the moments m_{ρ_C}, $m_{\rho_{CY}}$, and $m_{sd_{CY}}$. The prior is instead provided by each theoretical model M_i in turn. Under appropriate conditions, listed in Geweke (2010), the density $P_1(data|m_{\rho_C}, m_{\rho_{CY}}, m_{sd_{CY}}, E_1)$ that appears in the triple integral can be replaced by the density $P_1(m_{\rho_C}, m_{\rho_{CY}}, m_{sd_{CY}}|data, E_1^*)$. The latter density is easy to work with, because it corresponds to the posterior density for the moments of interest implied by an auxiliary econometric model E_1^*, which in this application is the BVAR model introduced above.

In the formula approximating the integrals, I rely on an independent trivariate normal kernel $K_1(.;.)$.[22] The numbers N_{M_i} and N_{E^*} stand for the number of draws for the theoretical models ($1,000$ each) and for the Bayesian VAR ($10,000$).[23]

[Table 5 about here]

The ten Bayes factors related to fitting the three macroeconomic moments are reported in Table 5. In Case 1, when the data are detrended with the HP filter, the ordering by Bayes factors is model $M4_{(\beta,\gamma)}$ over $M3_{(\gamma)}$ over $M2_{(\beta)}$ over $M1_{(IM)}$ over $M0_{(CM)}$, with the ratios being roughly $75 : 13 : 1 : 0.0006 : 0.0002$.[24]

When focusing on a detrending method, the ordering of the models is remarkably stable across cases, with the exception of models $M3_{(\gamma)}$ and $M4_{(\beta,\gamma)}$, which can swap their place at the top. A key takeaway is that the overall empirical performance of the two models with risk aversion heterogeneity is similar, and better than what the other versions achieve, indicating that this element of heterogeneity should be routinely included in quantitative work. Differently, the two simpler models do not appear to be fruitful frameworks.

5.2 Monte Carlo Analysis

A natural question is whether the MEI methodology is a reliable tool for empirical analysis. This concern is a valid one, especially when the models' fit is assessed on the basis of a limited number of observations and moments. In order to tackle this issue, I conduct a Monte Carlo analysis. The Data Generating Process is postulated to be one of the four incomplete markets models (in turn, and with all parameters set at the average

[22] For all density approximations, I use a fixed bandwidth equal to 0.003. I also experimented using a different value, such as 0.005 and 0.01, and the ranking of the models was not affected.

[23] In previous versions of the paper, I was using $2,000$ parameter draws, and I verified that the results were similar to the ones obtained using $1,000$ draws. It is also worth mentioning that in both models with risk aversion heterogeneity the ALM does not converge and shows cycles in a small number of replications, which are discarded. This also happens in the more traditional versions of the model, but in Cozzi (2015) I have shown with a perturbation approach that self-fulfilling equilibria are not likely to arise in the incomplete markets model with aggregate shocks and discount factor heterogeneity.

[24] When the data are detrended with the BN filter, whose results are in Table 9 in Appendix G, the ordering is model $M3_{(\gamma)}$ over $M4_{(\beta,\gamma)}$ over $M2_{(\beta)}$ over $M1_{(IM)}$ over $M0_{(CM)}$, with the ratios being roughly $14.3 : 14.2 : 1 : 0.000005 : 0.0000001$.

of the uncorrelated uniform priors).[25] For each assumed DGP, I compute the equilibrium ALM and decision rules, I simulate an artificial sample, I construct the aggregate variables, and I filter their series using both the (one-sided) HP filter and the BN one. The priors for the parameters of the theoretical models are assumed to be uncorrelated Uniforms, whose bounds are the DGP parameter values $\pm 10\%$.[26] Table 6 reports the Bayes factors of the Monte Carlo experiments using 300 model-generated observations. The label at the top of a column states the specific DGP that the case is representing. The experiments with 150 and 3,000 observations are included in Appendix G.

[Table 6 about here]

On the one hand, it is reassuring that the MEI recovers the DGPs, by assigning them higher marginal likelihoods even with a relatively small, and empirically relevant, sample size. However, in a handful of comparisons, the Bayes factors do not provide overwhelming evidence in favor of the DGP. In some cases, this is unavoidable. For instance, when model $M1_{(IM)}$ is assumed to be the DGP, because the distributions of its populations moments are almost identical to those of model $M0_{(CM)}$. Any moment-based empirical procedure would have a hard time rejecting one of the two models in their pairwise comparison. A valuable lesson learned from the Monte Carlo analysis is that the MEI works well when using a fixed bandwidth for the kernel density approximations. This is not the case when the bandwidth is allowed to change from one model comparison to the next, such as when using Silverman's rule of thumb. Using relatively small bandwidths is preferable. These can also vary with the moment under consideration, but they have to be set equal in all approximations across model comparisons. Failure to do so can lead to the wrong model ranking.

5.3 Micro Vs. Macro Data and Measures of Fit

The models under study have implications for both macro and micro-economic outcomes. Given the presence of household heterogeneity, it is possible to consider formally their cross-sectional properties. I therefore examine an extension to the measure of fit considered above, to include both macro and micro-economic moments.

Data limitations coupled with modeling challenges make it unfeasible to work with joint marginal likelihoods, stemming from a mix of micro and macro data.[27] Instead, I consider the marginal likelihood conditional on the macro data separately from the marginal likelihood conditional on the micro data. To combine the two

[25] Since the complete markets model $M0_{(CM)}$ implies moments that are very close to its corresponding incomplete markets version, the Monte Carlo results for the two models must be similar.

[26] For most parameters, these bounds are similar to the ones listed in Section 3. However, it is not possible to apply the $\pm 10\%$ perturbation to the discount factors, as they would exceed 1, and I perturb them by ± 0.001 (± 0.005) when they are heterogeneous (homogeneous). Finally, the $\pm 10\%$ perturbation would imply small ranges for the risk aversions, and I perturb them by ± 0.3 (± 1.0) when they are heterogeneous (homogeneous).

[27] There are virtually no panel datasets representative of the population of interest (i.e., a whole country) with a long time series dimension and collected at a frequency that can accurately capture short-run economic fluctuations. In particular, in the PSID there are no questions that would allow to measure aggregate consumption or investment. Not to mention that those data were collected annually and, more recently, every other year. Other popular alternatives, the NLSY and the HRS, consider selected samples, which are not amenable to an aggregate analysis.

dimensions, I rely on a weighted average of the two related Bayes factors, using the relative number of moments employed in each case as weights (ω).[28] The average Bayes factors BF, which are the average of the underlying macro and micro ones, are obtained as follows: $BF = \omega * BF_1 + (1-\omega) * BF_2$, where in this application $\omega = 3/4$. It is now understood that two separate a-theoretical econometric models, E_1^* and E_2^*, are needed to obtain the posterior densities of the macro and micro moments, respectively.

As all the economic models have non-degenerate wealth distributions, the microeconomic dimension of interest is the long-run wealth Gini index (g).[29] The formula for the micro Bayes factors is:

$$BF_2 \equiv \frac{P_2(M_i|micro\ data, E_2)}{P_2(M_j|micro\ data, E_2)} = \frac{P_2(M_i|E_2)P_2(micro\ data|M_i, E_2)}{P_2(M_j|E_2)P_2(micro\ data|M_j, E_2)}$$

$$= \frac{P_2(M_i|E_2)\int P_2(g|M_i)P_2(micro\ data|g, E_2)dg}{P_2(M_j|E_2)\int P_2(g|M_j)P_2(micro\ data|g, E_2)dg}$$

$$\propto \frac{\int P_2(g|M_i)P_2(g|micro\ data, E_2^*)dg}{\int P_2(g|M_j)P_2(g|micro\ data, E_2^*)dg}$$

$$\approx \frac{\frac{1}{N_{M_i}N_{E_2^*}}\sum_{u=1}^{N_{M_i}}\sum_{v=1}^{N_{E_2^*}} K_2\left(g_u^{M_i}; g_v^{E_2^*}\right)}{\frac{1}{N_{M_j}N_{E_2^*}}\sum_{u=1}^{N_{M_j}}\sum_{v=1}^{N_{E_2^*}} K_2\left(g_u^{M_j}; g_v^{E_2^*}\right)}$$

In this case, the computation of the posterior odds ratios is achieved with simple univariate kernel density approximations.

A Bayesian Econometric Model for the Wealth Gini Index: In the extended framework, there is the need to link the distribution of the population wealth inequality to the observables, which are the cross sectional wealth distributions in the PSID waves between 1984 and 2019. Therefore, for each wave, I specify a simple cross-sectional Bayesian model for log wealth $\widetilde{a}_i \equiv \log a_i$: $\widetilde{a}_i \overset{iid}{\sim} N(\mu_a, \sigma_a^2)$.[30]

Under this distributional assumption there is a closed form formula for the Gini index $g = 2\Phi(\sigma_a/\sqrt{2}) - 1$,

[28] Since in principle there are no restrictions to the number of moments that can be considered in either dimension, it is reasonable to take this element into account, to give more importance to the aspect that is assessed with more moments. Also, computing the Bayes factor by taking ratios of average marginal likelihoods is not feasible, as in its formula the terms $P_1(Y, C|E_1^*)$ and $P_2(a|E_2^*)$ would not drop out.

[29] Recall that in the PSID the consumption measure is not ideal, as it refers to food expenditure. This makes it undesirable to consider the same correlations at both the macro and micro levels. Also, data on the wealth distribution are not available at the quarterly frequency, and the SCF data are typically collected every two years. Since the wealthiest households tend to be the entrepreneurs, which are not modeled here, I rely on the PSID wealth data instead. Finally, although the U.S. wealth distribution has been fairly stable over time, I consider several cross sections to make the population moment of interest –the long-run wealth Gini index– consistent with the available data.

[30] More precisely, I consider a truncated distribution, as wealth holdings can be negative. The log-normal assumption applies above a truncation point a_{q_1} such that $\widetilde{a}_i \overset{iid}{\sim} N(\mu_a, \sigma_a^2)|a_i > a_{q_1}$. Some experimentation showed that truncating the empirical distributions at the first quartile provides an excellent fit between the actual Gini index and the theoretical one. For computational convenience, the truncation point a_{q_1} is assumed to be known.

where $\Phi(.)$ is the CDF of the standard normal. This helps tremendously in the computation of the posterior for the wealth Gini index, as it is readily obtained by sampling from the posterior distribution of σ_a^2.

I assume a diffuse prior for both μ_a and σ_a^2. Following well-known results in Zellner (1971) and Lancaster (2004), the marginal posterior for the variance σ_a^2 is proportional to a $Gamma(\nu)$ distribution, where ν denotes the degrees of freedom. The computation of the posterior distribution for the Gini index for a given cross sectional sample is then straightforward. A complication arises from the fact that the data consists of a sequence of cross sections, and that the models have implications for the long-run degree of wealth inequality. I deal with this aspect by considering the mixture of the resulting posterior distributions, each equally weighted, and with a weight corresponding to the inverse of the number of PSID waves ($N_{PSID} = 14$) used in the analysis.

$$P_2(g|micro\ data, E_2^*) = \sum_{q=1}^{N_{PSID}} \frac{1}{N_{PSID}} P_{2,q}(g|micro\ data\ in\ wave\ q, E_2^*)$$

The results related to the marginal likelihoods are included in the bottom part of Table 4, while the bottom portions of Table 5 and 9 report the average Bayes factors.[31]

It is not surprising that, in this dimension, the model without preference heterogeneity displays a poor fit. Differently, the LML for the other three models are similar and quite stable across cases. Considering the weighted average of the two Bayes factors computed on the macro and micro moments does not alter the ranking obtained with the macro moments only. The evidence in favor of the two models with risk aversion heterogeneity becomes somewhat weaker, but the results based on detrending the aggregate data with the HP filter still offer substantial evidence in favor of model $M4_{(\beta,\gamma)}$.

[Table 7 about here]

To conclude with, I perform a Monte Carlo analysis also for the microeconomic moment, whose results are reported in Table 7. For these experiments, I consider model-generated data with 30,000 cross sectional observations collected in 4 waves (notice that these simulations do not require detrending). Also when working with the wealth Gini index, the MEI recovers the DGPs, by assigning them both higher marginal likelihoods and sizable Bayes factors. Finally, the bottom portions of the Table report the average Bayes factors of the Monte Carlo experiments, combining the macro and micro moments. In most comparisons, the true DGPs display very large average Bayes factors. The results in these Monte Carlo experiments are valuable also because they help gauging how big the average Bayes factors should be in order to be informative. When comparing the top two models, values above 2 seem to be large enough to provide evidence supporting the DGP against its closest competitor.

[31] The mixture of the wealth Gini index's posterior densities is plotted in Figure 5 in Appendix G, which in Figure 4 also includes the distributions of the model-generated wealth Gini index, for Case 2.

6 Conclusions

In this paper, I have undertaken a formal comparison of models with heterogeneous agents and aggregate uncertainty, extending the computationally tractable Bayesian framework of DeJong, Ingram, and Whiteman (1996) and Geweke (2010). I have incorporated time-varying heterogeneity in risk aversion, finding that this additional layer of heterogeneity improves the model's fit. Relying on three aggregate moments that are closely linked to consumption smoothing behavior, I have found that the two models with risk aversion heterogeneity attain the largest marginal likelihoods. This result holds also when I incorporate into the empirical framework an additional microeconomic moment, the wealth Gini index. Finally, I have also assessed the MEI methodology with a Monte Carlo analysis, showing that both macroeconomic and microeconomic data help discriminating between different specifications of the model. The results have shown that risk aversion heterogeneity is quantitatively important, and it should be routinely included in macroeconomic models with aggregate shocks.

To conclude, given that fully Bayesian methods are infeasible for most specifications of incomplete markets model with aggregate risk, the MEI represents a viable and coherent approach for model comparison in the presence of parameter uncertainty. In more general models, say with flexible labor supply or overlapping-generations, the set-up can be easily extended to consider a larger number of microeconomic moments, such as the correlation between wealth and leisure, or the life-cycle dynamics of wealth and hours worked.

Label	*Model Set-up*
M0 $_{(CM)}$	*Complete Markets and No Heterogeneity*
M1 $_{(IM)}$	*Incomplete Markets and No Preference Heterogeneity*
M2 $_{(\beta)}$	*Incomplete Markets and* β *Heterogeneity (No* γ *Heterogeneity)*
M3 $_{(\gamma)}$	*Incomplete Markets and* γ *Heterogeneity (No* β *Heterogeneity)*
M4 $_{(\beta,\gamma)}$	*Incomplete Markets and* β *and* γ *Heterogeneity*
Label	*Parameter Priors Set-up*
Case 1	*Correlated Uniforms*
Case 2	*Uncorrelated Uniforms*
Case 3	*Uncorrelated Betas*
Case 4	*Uncorrelated Uniforms with overlaps in the heterogeneous preference parameters*
Monte Carlo	*Uncorrelated Uniforms with bounds around the DGP*$(\pm 10\%)$

Table 1: Models and Cases Description.

Case	*M1* $_{(IM)}$	*M2* $_{(\beta)}$	*M3* $_{(\gamma)}$	*M4* $_{(\beta,\gamma)}$
HtM Recessions	26.0%	34.1%	53.9%	37.2%
HtM Booms	4.0%	6.9%	4.1%	14.0%
W-HtM Recessions	-	26.9%	-	35.6%
W-HtM Booms	-	4.2%	-	13.5%

Table 2: Share of Hand-to-Mouth and Wealthy Hand-to-Mouth households.

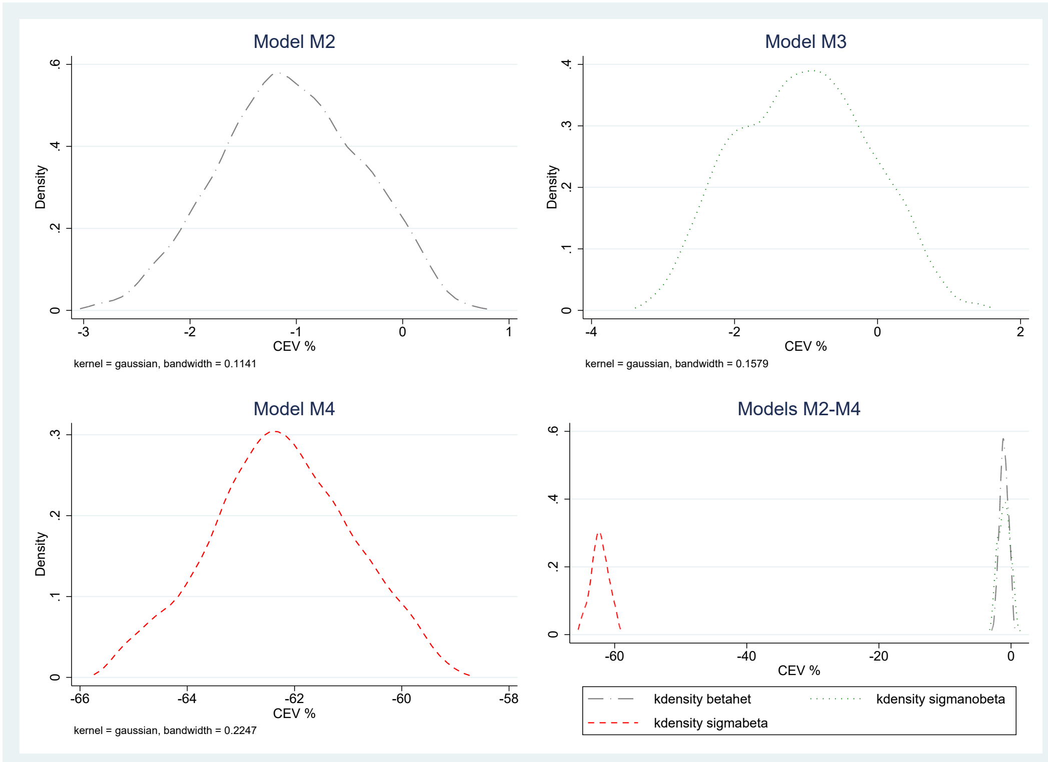

Figure 1: Kernel Densities of the Long-Run Welfare Effects (CEV%) of Eliminating Business Cycles in models $M2_{(\beta)}$, $M3_{(\gamma)}$, and $M4_{(\beta,\gamma)}$. *Notes*: The welfare effects are obtained by applying the integration principle.

Model	*Parameter*	*Description*	*Min*	*Max*
All	α	*Capital share*	$\underline{\alpha} = 0.32$	$\overline{\alpha} = 0.40$
All	δ	*Capital depreciation rate*	$\underline{\delta} = 0.020$	$\overline{\delta} = 0.025$
M1-M4	$a_{\min}$	*Borrowing limit*	$\underline{a_{\min}} = -2.0$	$\overline{a_{\min}} = 0$
All	l	*Hours worked (share of time endowment)*	$\underline{l} = 0.287$	$\overline{l} = 0.351$
All	u_G	*Unemployment rate in expansions*	$\underline{u_G} = 0.03$	$\overline{u_G} = 0.05$
All	u_B	*Unemployment rate in recessions*	$\underline{u_B} = 0.09$	$\overline{u_B} = 0.11$
M0 $_{(CM)}$	β	*Discount factor*	$\underline{\beta} = 0.985$	$\overline{\beta} = 0.995$
	γ	*Relative risk aversion*	$\underline{\gamma} = 1.0$	$\overline{\gamma} = 3.0$
M1 $_{(IM)}$	β	*Discount factor*	$\underline{\beta} = 0.9849$	$\overline{\beta} = 0.9949$
	γ	*Relative risk aversion*	$\underline{\gamma} = 1.0$	$\overline{\gamma} = 3.0$
M2 $_{(\beta)}$	β_h	*Discount factor (type-specific)*	$\underline{\beta_h} = 0.992$	$\overline{\beta_h} = 0.994$
	β_m	"	$\underline{\beta_m} = 0.9884$	$\overline{\beta_m} = 0.9904$
	β_l	"	$\underline{\beta_l} = 0.9848$	$\overline{\beta_l} = 0.9868$
	γ	*Relative risk aversion*	$\underline{\gamma} = 1.0$	$\overline{\gamma} = 3.0$
M3 $_{(\gamma)}$	β	*Discount factor*	$\underline{\beta} = 0.984$	$\overline{\beta} = 0.994$
	γ_h	*Relative risk aversion (type-specific)*	$\underline{\gamma_h} = 7.247$	$\overline{\gamma_h} = 7.847$
	γ_m	"	$\underline{\gamma_m} = 1.818$	$\overline{\gamma_m} = 2.418$
	γ_l	"	$\underline{\gamma_l} = 0.623$	$\overline{\gamma_l} = 1.223$
	p_{γ_j}	*CRRA Markov chain probabilities*	*See text*	
M4 $_{(\beta,\gamma)}$	β_h	*Discount factor (type-specific)*	$\underline{\beta_h} = 0.992$	$\overline{\beta_h} = 0.994$
	β_m	"	$\underline{\beta_m} = 0.984$	$\overline{\beta_m} = 0.986$
	β_l	"	$\underline{\beta_l} = 0.980$	$\overline{\beta_l} = 0.982$
	γ_h	*Relative risk aversion (type-specific)*	$\underline{\gamma_h} = 7.247$	$\overline{\gamma_h} = 7.847$
	γ_m	"	$\underline{\gamma_m} = 1.818$	$\overline{\gamma_m} = 2.418$
	γ_l	"	$\underline{\gamma_l} = 0.623$	$\overline{\gamma_l} = 1.223$
	p_{γ_j}	*CRRA Markov chain probabilities*	*See text*	

Table 3: Model parameters and their priors' support. *Notes*: The model period is a quarter. The parameters listed in the top of the table are common to the four incomplete markets models. The others are model-specific.

Model	*LML (Case 1)*	*LML (Case 2)*	*LML (Case 3)*	*LML (Case 4)*
Macro Moments - One-sided HP Filter				
$M0_{(CM)}$ - *CM and No Heterogeneity*	-13.953	-12.915	-14.767	-12.915
$M1_{(IM)}$ - *No Preference Heterogeneity*	-13.017	-12.167	-14.125	-12.167
$M2_{(\beta)}$ - β *Heterogeneity*	-5.539	-4.799	-8.654	-4.756
$M3_{(\gamma)}$ - γ *Heterogeneity*	-2.957	-2.955	-4.073	-3.949
$M4_{(\beta,\gamma)}$ - β *and* γ *Heterogeneity*	-1.209	-0.920	-1.863	-2.904
Macro Moments - BN Filter				
$M0_{(CM)}$ - *CM and No Heterogeneity*	-18.152	-16.766	-19.325	-16.766
$M1_{(IM)}$ - *No Preference Heterogeneity*	-14.737	-14.971	-17.662	-14.971
$M2_{(\beta)}$ - β *Heterogeneity*	-2.455	-1.177	-6.266	-1.043
$M3_{(\gamma)}$ - γ *Heterogeneity*	0.202	0.009	-0.076	0.095
$M4_{(\beta,\gamma)}$ - β *and* γ *Heterogeneity*	0.199	0.077	-1.033	-0.356
Micro Moment				
$M0_{(CM)}$ - *CM and No Heterogeneity*	NA	NA	NA	NA
$M1_{(IM)}$ - *No Preference Heterogeneity*	-0.127	-2.816	-18.112	-2.816
$M2_{(\beta)}$ - β *Heterogeneity*	3.836	3.797	3.901	3.663
$M3_{(\gamma)}$ - γ *Heterogeneity*	3.650	3.533	2.884	3.371
$M4_{(\beta,\gamma)}$ - β *and* γ *Heterogeneity*	3.972	3.950	3.820	3.786

Table 4: Log Marginal Likelihoods (LML) under the Minimal Econometric Interpretation. *Notes*: The top and middle panels (bottom panel) report the LML based on the Macro-economic (Micro-economic) moments (moment). The time series data are filtered with either the one-sided HP filter (top panel) or the BN filter (middle panel). A fixed bandwidth equal to 0.003 is used for all density approximations.

Comparison	*Macro BF (Case 1)*	*Macro BF (Case 2)*	*Macro BF (Case 3)*	*Macro BF (Case 4)*
$P(M0_{(CM)}\|.)/P(M1_{(0)}\|.)$	0.392	0.474	0.527	0.474
$P(M0_{(CM)}\|.)/P(M2_{(\beta)}\|.)$	0.000	0.000	0.002	0.000
$P(M0_{(CM)}\|.)/P(M3_{(\gamma)}\|.)$	0.000	0.000	0.000	0.000
$P(M0_{(CM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	0.000	0.000	0.000	0.000
$P(M1_{(0)}\|.)/P(M2_{(\beta)}\|.)$	0.001	0.001	0.004	0.001
$P(M1_{(0)}\|.)/P(M3_{(\gamma)}\|.)$	0.000	0.000	0.000	0.000
$P(M1_{(0)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	0.000	0.000	0.000	0.000
$P(M2_{(\beta)}\|.)/P(M3_{(\gamma)}\|.)$	0.076	0.158	0.010	0.446
$P(M2_{(\beta)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	0.013	0.021	0.001	0.157
$P(M3_{(\gamma)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	0.174	0.131	0.110	0.352
Comparison	*Average BF (Case 1)*	*Average BF (Case 2)*	*Average BF (Case 3)*	*Average BF (Case 4)*
$P(M1_{(0)}\|.)/P(M2_{(\beta)}\|.)$	0.005	0.001	0.003	0.001
$P(M1_{(0)}\|.)/P(M3_{(\gamma)}\|.)$	0.006	0.001	0.000	0.001
$P(M1_{(0)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	0.004	0.000	0.000	0.000
$P(M2_{(\beta)}\|.)/P(M3_{(\gamma)}\|.)$	0.358	0.444	0.699	0.670
$P(M2_{(\beta)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	0.228	0.230	0.272	0.339
$P(M3_{(\gamma)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	0.312	0.263	0.180	0.429

Table 5: Model Comparison: Bayes Factors, One-sided HP Filter.

Comparison \ DGP	$M1_{(IM)}$	$M2_{(\beta)}$	$M3_{(\gamma)}$	$M4_{(\beta,\gamma)}$
One-Sided HP Filter				
$P(M0_{(CM)}\|.)/P(M1_{(IM)}\|.)$	0.935	0.753	0.845	0.366
$P(M0_{(CM)}\|.)/P(M2_{(\beta)}\|.)$	1.624	0.239	2.130	0.0001
$P(M0_{(CM)}\|.)/P(M3_{(\gamma)}\|.)$	4.060	1.069	0.612	0.0001
$P(M0_{(CM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	506.233	17.971	98.652	0.0001
$P(M1_{(IM)}\|.)/P(M2_{(\beta)}\|.)$	1.737	0.317	2.520	0.0001
$P(M1_{(IM)}\|.)/P(M3_{(\gamma)}\|.)$	4.345	1.419	0.724	0.0001
$P(M1_{(IM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	541.691	23.857	116.682	0.0001
$P(M2_{(\beta)}\|.)/P(M3_{(\gamma)}\|.)$	2.501	4.479	0.287	0.004
$P(M2_{(\beta)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	311.776	75.325	46.307	0.003
$P(M3_{(\gamma)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	124.674	16.818	161.239	0.774
BN Filter				
$P(M0_{(CM)}\|.)/P(M1_{(IM)}\|.)$	0.839	0.474	0.611	1.000
$P(M0_{(CM)}\|.)/P(M2_{(\beta)}\|.)$	1.798	0.077	1.254	0.0001
$P(M0_{(CM)}\|.)/P(M3_{(\gamma)}\|.)$	3.970	0.383	0.355	0.0001
$P(M0_{(CM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	989.205	8.389	108.171	0.0001
$P(M1_{(IM)}\|.)/P(M2_{(\beta)}\|.)$	2.144	0.162	2.053	0.0001
$P(M1_{(IM)}\|.)/P(M3_{(\gamma)}\|.)$	4.734	0.808	0.581	0.0001
$P(M1_{(IM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	1179.529	17.687	177.119	0.0001
$P(M2_{(\beta)}\|.)/P(M3_{(\gamma)}\|.)$	2.208	4.979	0.283	0.0001
$P(M2_{(\beta)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	550.141	109.010	86.287	0.0001
$P(M3_{(\gamma)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	249.182	21.893	304.668	0.586

Table 6: Model Comparison for the Monte Carlo analysis, $T = 300$: Macro Bayes Factors. *Notes*: The true DGP is indicated by the column labels.

Model or Comparison \ DGP	$M1_{(IM)}$	$M2_{(\beta)}$	$M3_{(\gamma)}$	$M4_{(\beta,\gamma)}$
Micro BF				
$P(M1_{(IM)}\|.)/P(M2_{(\beta)}\|.)$	1567.7	0.013	∞	0.328
$P(M1_{(IM)}\|.)/P(M3_{(\gamma)}\|.)$	12.368	0.056	0.046	0.916
$P(M1_{(IM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	127.60	0.029	552.293	0.115
$P(M2_{(\beta)}\|.)/P(M3_{(\gamma)}\|.)$	0.008	4.429	0.000	2.788
$P(M2_{(\beta)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	0.081	2.329	0.000	0.350
$P(M3_{(\gamma)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	10.317	0.526	12091.9	0.126
Average BF (HP Filter)				
$P(M1_{(IM)}\|.)/P(M2_{(\beta)}\|.)$	393.23	0.241	∞	0.082
$P(M1_{(IM)}\|.)/P(M3_{(\gamma)}\|.)$	6.351	1.078	0.554	0.229
$P(M1_{(IM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	438.17	17.900	225.585	0.029
$P(M2_{(\beta)}\|.)/P(M3_{(\gamma)}\|.)$	1.878	4.466	0.215	0.700
$P(M2_{(\beta)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	233.85	57.076	34.730	0.090
$P(M3_{(\gamma)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	96.085	12.745	3143.9	0.612
Average BF (BN Filter)				
$P(M1_{(IM)}\|.)/P(M2_{(\beta)}\|.)$	393.54	0.125	∞	0.082
$P(M1_{(IM)}\|.)/P(M3_{(\gamma)}\|.)$	6.642	0.620	0.447	0.229
$P(M1_{(IM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	916.55	13.273	270.912	0.029
$P(M2_{(\beta)}\|.)/P(M3_{(\gamma)}\|.)$	1.658	4.842	0.212	0.697
$P(M2_{(\beta)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	412.63	82.340	64.715	0.087
$P(M3_{(\gamma)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	189.46	16.552	3251.4	0.471

Table 7: Monte Carlo Analysis for the Wealth Gini Index: Bayes Factors. *Notes*: The true DGP is indicated by the column labels.

References

Ahn, S., Kaplan, G., Moll, B., Winberry, T., and Wolf, C. (2017). "When Inequality Matters for Macro and Macro Matters for Inequality," *University of Chicago Press*, NBER Macroeconomics Annual, Vol. 32.

An, S., and Schorfheide, F. (2007). "Bayesian Analysis of DSGE Models," *Econometric Reviews*, Vol. 26, 113-172.

Attanasio, O., and Weber, G. (2010). "Consumption and Saving: Models of Intertemporal Allocation and Their Implications for Public Policy," *Journal of Economic Literature* Vol. 48, 693-751

Canova, F. (2009). "How Much Structure in Empirical Models?" in *Palgrave Handbook of Applied Econometrics*, edited by T. Mills and K. Patterson, Vol. 2, 68-97.

Canova, F. (1994). "Statistical Inference in Calibrated Models," *Journal of Applied Econometrics*, Vol. 9, S123-S144.

Castaneda, A., Diaz-Gimenez, J., and Rios-Rull, J.V. (1998). "Exploring the income distribution business cycle dynamics," *Journal of Monetary Economics*, Vol. 42, 93-130.

Cozzi, M. (2015). "The Krusell-Smith Algorithm: Are Self-fulfilling Equilibria Likely?" *Computational Economics*, Vol. 46, 653-670.

Cozzi, M. (2014). "Risk Aversion Heterogeneity, Risky Jobs and Wealth Inequality," QED Working Paper n. 1286.

Chiu, J., and Molico, M. (2010). "Liquidity, Redistribution, and the Welfare Cost of Inflation," *Journal of Monetary Economics*, Vol. 57, 428-38.

DeJong, D., Ingram, B., and Whiteman, C. (1996). "A Bayesian Approach to Calibration," *Journal of Business and Economic Statistics*, Vol. 14, 1-9.

DeJong, D., and Ripoll, M. (2007). "Do self-control preferences help explain the puzzling behavior of asset prices?" *Journal of Monetary Economics*, Vol. 54, 1035-1050.

Eichenbaum, M. (1991). "Real business-cycle theory. Wisdom or whimsy?" *Journal of Economic Dynamics and Control*, Vol. 15, 607-626.

Fernandez-Villaverde, J., Guerron, P., and Rubio-Ramirez, J. (2010). "The New Macroeconometrics: a Bayesian Approach," *Handbook of Applied Bayesian Analysis*, edited by O'Hagan and West, Oxford University Press.

Geweke, J. (2010). *Complete and Incomplete Econometric Models*. Princeton University Press, Princeton (NJ).

den Haan, W. (2010). "Comparison of solutions to the incomplete markets model with aggregate uncertainty," *Journal of Economic Dynamics & Control*, Vol. 34, 4-27.

den Haan, W., Judd, K., and Juillard, M. (2010). “Computational suite of models with heterogeneous agents: Incomplete markets and aggregate uncertainty,” *Journal of Economic Dynamics & Control*, Vol. 34, 1-3.

Heathcote, J. (2005). “Fiscal Policy with Heterogeneous Agents and Incomplete Markets.” *Review of Economic Studies*, Vol. 72, 161-188.

Juster, T., and Stafford, F. (1991). “The Allocation of Time: Empirical Findings, Behavioral Models, and Problems of Measurement,” *Journal of Economic Literature*, Vol. 29, 471-522.

Kamber, G., Morley, J., and Wong, B. (2018). “Intuitive and reliable estimates of the output gap from the Beveridge-Nelson filter,” *The Review of Economics and Statistics*, Vol. 100, 550-566.

Kaplan, G., Violante, G., and Weidner, J. (2009). “The Wealthy Hand-to-Mouth,” *Brookings Papers on Economic Activity*, Spring, 77-153.

Kimball, M., Sahm, C., and Shapiro, M. (2009). “Risk Preferences in the PSID: Individual Imputations and Family Covariation,” *American Economic Review: Papers & Proceedings*, Vol. 99, 363-368.

Kimball, M., Sahm, C., and Shapiro, M. (2008). “Imputing Risk Tolerance From Survey Responses,” *Journal of the American Statistical Association*, Vol. 103, 1028-1038.

Koop, G., and Korobilis, D. (2009). “Bayesian Multivariate Time Series Methods for Empirical Macroeconomics,” *Foundations and Trends in Econometrics*, Vol. 3, 267-358.

Krueger, D., Perri, F., Pistaferri, L., and Violante, G. (2010). “Cross Sectional Facts for Macroeconomists,” *Review of Economic Dynamics*, Vol. 13, 1-14.

Krusell, P., Mukoyama, T., Sahin, A., and Smith, A. (2009). “Revisiting the welfare effects of eliminating business cycles,” *Review of Economic Dynamics*, Vol. 12, 393-404.

Krusell, P., and Smith, A. (1999). “On the welfare effects of eliminating business cycles,” *Review of Economic Dynamics*, Vol. 100, 550-566.

Krusell, P., and Smith, A. (1998). “Income and Wealth Heterogeneity in the Macroeconomy,” *Journal of Political Economy*, Vol. 106, 867-896.

Lancaster, T. (2004). *An Introduction to Bayesian Modern Econometrics*. Blackwell Publishing, MA, USA.

Lucas, R. (1987). *Models of Business Cycles*. Basil Blackwell, Oxford.

Mukoyama, T., and Sahin, A. (2005). “Costs of Business Cycles for Unskilled Workers,” CIREQ, Cahier 15-2005.

Mukoyama, T., and Sahin, A. (2006). “Costs of business cycles for unskilled workers,” *Journal of Monetary Economics*, Vol. 53, 2179-2193.

Rios-Rull, J.V., Schorfheide, F., Fuentes-Albero, C., Kryshko, M., and Santaeulalia-Llopis, R. (2012). “Methods versus substance: Measuring the effects of technology shocks,” *Journal of Monetary Economics*, Vol. 59, 826-846.

Stock, J., and Watson, M. (1999). "Forecasting Inflation," *Journal of Monetary Economics*, Vol. 44, 293-335.

Storesletten, K., Telmer, C., and Yaron, A. (2001). "The welfare cost of business cycles revisited: Finite lives and cyclical variation in idiosyncratic risk," *European Economic Review*, Vol. 45, 1311-1339.

Zellner, A. (1971). "Bayesian and Non-Bayesian Analysis of the Log-Normal Distribution and Log-Normal Regression." *Journal of the American Statistical Association*, Vol. 66, 327-330.

Appendix A - Computation

- All codes solving the economies were written in the FORTRAN 95 language, relying on the Intel Fortran Compiler, build 18.0.02 (with the IMSL library). They were compiled selecting the O3 option (maximize speed), and without automatic parallelization. They were run on a 64-bit PC platform, running Windows 10 Professional Edition, with either an Intel Xeon $E5-2687Wv2$ Octo-core processor clocked at 3.4 Ghz, or an Intel Core $i7-9900k$ Octo-core processor clocked at 4.8 Ghz.

- The $1,000$ replications are run in parallel across cores and for model *M4* take up to 4 days to complete. Notice that typically from 15 to 25 iterations on the ALM are needed to find each equilibrium.

- In the actual solution of the model I need to discretize the continuous state variables a and K (the discount factor heterogeneity β, the risk aversion heterogeneity γ, the employment status s, and the aggregate productivity shock z are already discrete). For the household assets a I rely on an unevenly spaced grid, with the distance between two consecutive points increasing geometrically. This is done to allow for a high precision of the policy rules at low values of a, where the change in curvature is more pronounced. In order to keep the computational burden manageable, I use 101 grid points on the household assets space (75 grid points for model *M4*), the lowest value being the borrowing constraint b and the highest one being a value $a_{\max}$ high enough not to be binding in the simulations ($a_{\max} = 500$). For the aggregate capital K I use a fairly dense grid. Over the $[3, 27]$ interval I use 25 points, which are far more than the typical 4-6. Given the large number of parameterizations, in the iterative process on the ALM parameters the simulations do visit regions for aggregate capital that are very far from the support of the ergodic equilibrium distribution, causing convergence issues when using a coarse grid.

- As for the solution method for the household problem, I use the Endogenous Grid Method (EGM) with linear interpolation in the (a, K) dimensions. This method is much faster and more stable than the relatively popular value function iteration scheme with cubic spline interpolation.

- The aggregate dynamics are computed by simulating a large panel of individuals for $6,000$ periods, with the first $2,000$ periods being discarded as a burn-in. The panel size is $50,000$ agents for the economies with preference heterogeneity, and $30,000$ agents for the other one. As for the approximation method, I rely on a bi-linear approximation scheme for the saving functions, for values of a and K falling outside the grid.

- The welfare effects are computed by simulating a large panel of individuals for $6,000$ periods, with the first $2,000$ periods being discarded as a burn-in. The panel size is $100,000$ agents.

- If the numerical procedure fails to converge in some of its steps, the related results are discarded.

Appendix B - Algorithm for the Models Solution

The computational procedure used to solve the incomplete markets model economies can be represented by the following algorithm:

1. Solve each version of the model at the average of the uniform prior distributions and store the equilibrium ALM parameters Θ^*.
2. Consider model M_i, $i = 1, ..., 4$.
3. Draw $1,000$ combinations of parameters from their prior distributions and store them.
4. Begin the replications and set $j = 1$.
5. Parametrize the model by reading the vector j of parameter draws and begin the model solution.
6. Generate a discrete grid over the aggregate capital space $[K_{\min}, ..., K_{\max}]$.
7. Generate a discrete grid over the individual asset space $[-b, ..., a_{\max}]$.
8. Guess a vector of parameters Θ^g representing the ALM, using in the first iteration Θ^*, the converged parameters for the model solved at the average of the priors.
9. Get the saving functions $a'(a, \beta, \gamma, s, z, K)$.
10. Simulate the model under the guessed ALM, and compute an update $\Theta^{g\prime}$ as the parameter estimates of OLS regressions on the simulated data.
11. Repeat steps $8 - 10$ until the four parameters in Θ converge.
12. Compute the BN or HP filtered series for consumption and income, and get $m_{\rho_{CY}}$, m_{ρ_C} and $m_{sd_{CY}}$.
13. Save the output, set $j = j + 1$ and repeat steps $5 - 12$ N_{M_i} times.
14. Move to the next model and repeat all the steps above.

To ensure similar numerical accuracy, also the complete markets model *M0* is solved with global methods, following a procedure similar to the one outlined above. Since the social planner problem breaks the computational efficiency of the EGM, it is solved using a collocation method instead.

Appendix C - Algorithm for the Bayesian VAR and the Moments' Posteriors

The computational procedure used to obtain the posterior distributions for the moments of interest can be represented by the following algorithm:

1. Get the OLS estimates $\widehat{D}_{OLS}$ for the VAR parameters and stack the estimates in the vector $\widehat{d}_{OLS} = vec(\widehat{D}_{OLS})$.

2. Compute the residuals $\mathcal{Y}_t - \widehat{D}_{OLS}\mathcal{Y}_{t-1}$ and the associated matrix $S = \left(\mathcal{Y}_t - \widehat{D}_{OLS}\mathcal{Y}_{t-1}\right)' \left(\mathcal{Y}_t - \widehat{D}_{OLS}\mathcal{Y}_{t-1}\right)$.

3. Get the inverse S^{-1} and $V = \left(\mathcal{Y}'_{t-1}\mathcal{Y}_{t-1}\right)^{-1}$.

4. Begin the replications and set $j = 1$.

5. Since the posterior for the matrix Σ is a Wishart distribution $\Sigma_{POST}|data \sim W(S, v)$, where v denotes the degrees of freedom, draw from the inverse Wishart distribution $\Sigma^{-1}_{POST} \sim W(S^{-1}, v)$.

6. To obtain Wishart draws, get the Cholesky decomposition of S^{-1}, draw v times from a correlated bivariate normal (there are only two variables in the VAR), collect the draws in the matrix H, and set $\Sigma^{-1}_{POST} = H'H$.

7. Get the inverse Σ_{POST} and use it to draw the VAR parameters from their posterior, which is a correlated normal $\widehat{d}_{POST}|\Sigma, data \sim N\left(\widehat{d}_{OLS}, \Sigma \otimes V\right)$.

8. Set up a system of three linear equations, whose unknowns are the population second order moments of Y and C. The entries in the vector of constants are the three distinct elements of the current draw Σ_{POST}, while the entries in the matrix of coefficients are simple functions of the current draws of the VAR parameters $\widehat{d}_{POST}$.

9. Consider the current posterior draws, solve the system, compute the moments of interest and store them.

10. Set $j = j + 1$ and repeat steps $5 - 9$ N_{E^*} times.

Appendix D - Data

The time series were obtained from the Federal Reserve Bank of St. Louis FRED II data base. The data are quarterly and the range is $1947Q1 - 2019Q4$ (Source: http://research.stlouisfed.org/fred2/).

Aggregate output is defined as the sum of Services, Non Durables and Investment, and the corresponding series are:

- Real personal consumption expenditures per capita: Nondurable goods (series ID: A796RX0Q048SBEA).
- Real personal consumption expenditures per capita: Services (series ID: A797RX0Q048SBEA).
- Real Gross Private Domestic Investment (series ID: GPDIC1).

Since the Investment series represents total investment in the economy, I divide it by the Civilian Non-institutional Population (series ID: B230RC0Q173SBEA).

In the empirical analysis, all time series are filtered, and so are the simulated series obtained from each model.

The cross sectional wealth data were obtained from the Survey Research Center at the University of Michigan. The wealth data were collected irregularly, first every five years, then every other year (Source: https://simba.isr.umich.edu/data/data.aspx). I use the following waves: 1984, 1989, 1994, 1999, 2001, 2003, 2005, 2007, 2009, 2011, 2013, 2015, 2017, 2019. Each cross section is truncated below the first quartile.

Appendix E - Long-Run Welfare Effects of Eliminating Business Cycles

In order to compute the welfare effects arising from the elimination of aggregate risk I follow the procedure proposed by Krusell and Smith (1999), Mukoyama and Sahin (2006) and Krusell, Mukoyama, Sahin, and Smith (2009). There are two issues worthy of being discussed. First, aggregate risk and idiosyncratic risk are correlated via the state dependent employment process. Eliminating aggregate risk involves adjusting the employment process, a procedure which does not correspond to considering the unconditional average of the Markov chains. Krusell, Mukoyama, Sahin, and Smith (2009) refer to this step as the integration principle, which requires formulating the models without aggregate risk in a recursive way by including two new values for the employment state variable. I follow the same formulation used by Mukoyama and Sahin (2006) and carefully presented in the Appendix of Mukoyama and Sahin (2005). In the models with risk aversion heterogeneity there is an additional complication, as this household characteristic is time varying. Because of this element, the CRRA coefficient is changing over time and, unlike in Krusell, Mukoyama, Sahin, and Smith (2009), the welfare effects cannot be obtained in closed form. The long-run welfare effects need to be computed by solving a non-linear equation, which is computationally demanding when evaluating the whole distribution of welfare effects (there is one potential welfare effect per initial condition, namely a specific combination of both aggregate and individual states).

Appendix F - The Relation to some Recent Methodological Papers

In this appendix, I outline how my approach differs from some recent papers in the literature, and what are the relative pros and cons.

Ahn, Kaplan, Moll, Winberry, and Wolf (2017) is an important contribution, but its limitations should not be underplayed. Their method allows to solve models with both heterogeneity and frictions that are impossible to tackle with the traditional recursive methods and solution algorithms. I believe that this is where the method should be applied, while using it for estimation is problematic. The appeal of Ahn, Kaplan, Moll, Winberry, and Wolf (2017) is the impressive speed of their solution method. An important caveat is that the authors are not solving the Krusell and Smith (1998) (KS) economy. In their formulation, there is one simplification that might seem innocuous, but has important implications. The stochastic process for the employment status must be time-invariant, and cannot depend on the aggregate state, which is a key assumption for their procedure to work. Granted, their aggregate shock has a continuous support, rather than the two-point process that I am considering. However, a simple simulation would reveal an important shortcoming: their labor market does not display booms and recessions, as the unemployment rate is constant over time. I should stress that this is not just a way to simplify the exposition in their paper. Rather, it is an essential feature of their method, as the model first needs to be solved by eliminating the effect of the aggregate shock. They are considering idiosyncratic uncertainty, but this cannot depend on the aggregate state, which is one of the defining features of KS's framework. Therefore, the accuracy results they report are not very informative, because they do not incorporate the errors introduced by neglecting the aggregate fluctuations in the labor market. Their DM statistic is somewhat better than what is reported in den Haan (2010). However, they are solving a different model, so the actual errors are likely to be large, because the decision rules are the ones for the average unemployment rate. Moreover, in the discussion section of the paper, Per Krusell points out that applying certainty equivalence (as they are doing) precludes studying important outcomes, such as the welfare costs of business cycles. Finally, likelihood based estimation is not feasible, because in their two-asset model their reduction technique leads to 300 state variables, which makes using the Kalman smoother intractable. It goes without saying that the MEI can be easily applied also to that model, and it is feasible. Moreover, I am providing a relatively easy way to incorporate cross sectional information into the empirical implementation of the theoretical models.

Winberry (2018) proposed to solve models with heterogeneity using perturbation methods. However, it is not clear how accurate these solutions are. In the firm heterogeneity set-up, Terry (2017) has shown that the method is not very accurate. In the household heterogeneity set-up, I used the codes included in Winberry (2018), and the results are inaccurate, as shown in the plot below for the unemployed agents. In terms of the long-run average of aggregate capital, comparing these results to what is delivered by an accurate solution method shows that there is a 3% discrepancy. This gap is extremely high (as a reference, when assessing different solution methods, den Haan already considers a 1% gap as being excessive). Clearly, a 3% gap can bias in severe and hard to predict ways the welfare effects computations of different policies, among many other outcomes. And the procedure would fare worse in models with fat tails, which is a feature of my models with preference heterogeneity. This problem is not present in the solution method I am using, as I checked that, when solving the steady state version of the model, simulating a large panel (as done in KS) or approximating numerically the CDF give very similar values for aggregate capital (and the other endogenous variables). In terms of estimation, notice that Winberry (2018) uses only two observables, and the feasibility of his method is due to a combination of model characteristics, and parameters that are estimated, because they do not affect the steady state of the model. Finally, in Table 8 I compare the moments obtained solving Winberry's household model with aggregate shocks (using the Krusell and Smith (1998) calibration), to those that I get with the KS algorithm. Some statistics are remarkably different. In terms of computing time, using the same PC, Winberry's method solves the model in approximately 155 seconds, while my solution implemented in Fortran 95 takes 125 seconds.

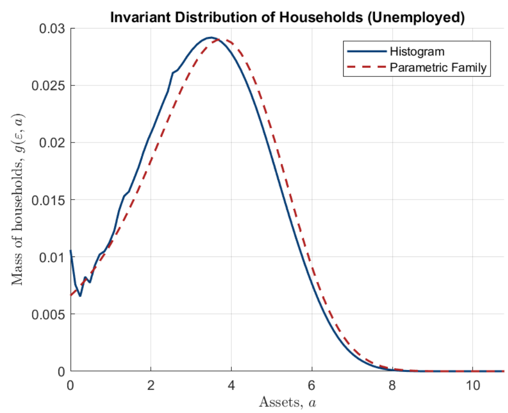

Figure 2: Cross Sectional Distributions.

Method	*KS*	*Winberry*
Sd(Y)	0.0237	0.0116
Sd(C)	0.0035	0.0041
Sd(I)	0.0882	0.0281
$\rho_{C,Y}$	0.7369	0.8372
$\rho_{C,I}$	0.6790	0.7401
$\rho_{Y,I}$	0.9966	0.9980
ρ_Y	0.5929	0.5985
ρ_C	0.7913	0.8131
ρ_I	0.5866	0.5600

Table 8: Comparison of Krusell and Smith (1998) and Winberry (2018).

References

Terry, S. (2017). "Alternative Methods for Solving Heterogeneous Firm Models," *Journal of Money, Credit, and Banking*, Vol. 49, 1081-1111.

Winberry, T. (2018). "A method for solving and estimating heterogeneous agent macro models," *Quantitative Economics*, Vol. 9, 1123-1151.

Appendix G - Additional Plots, Tables and Results

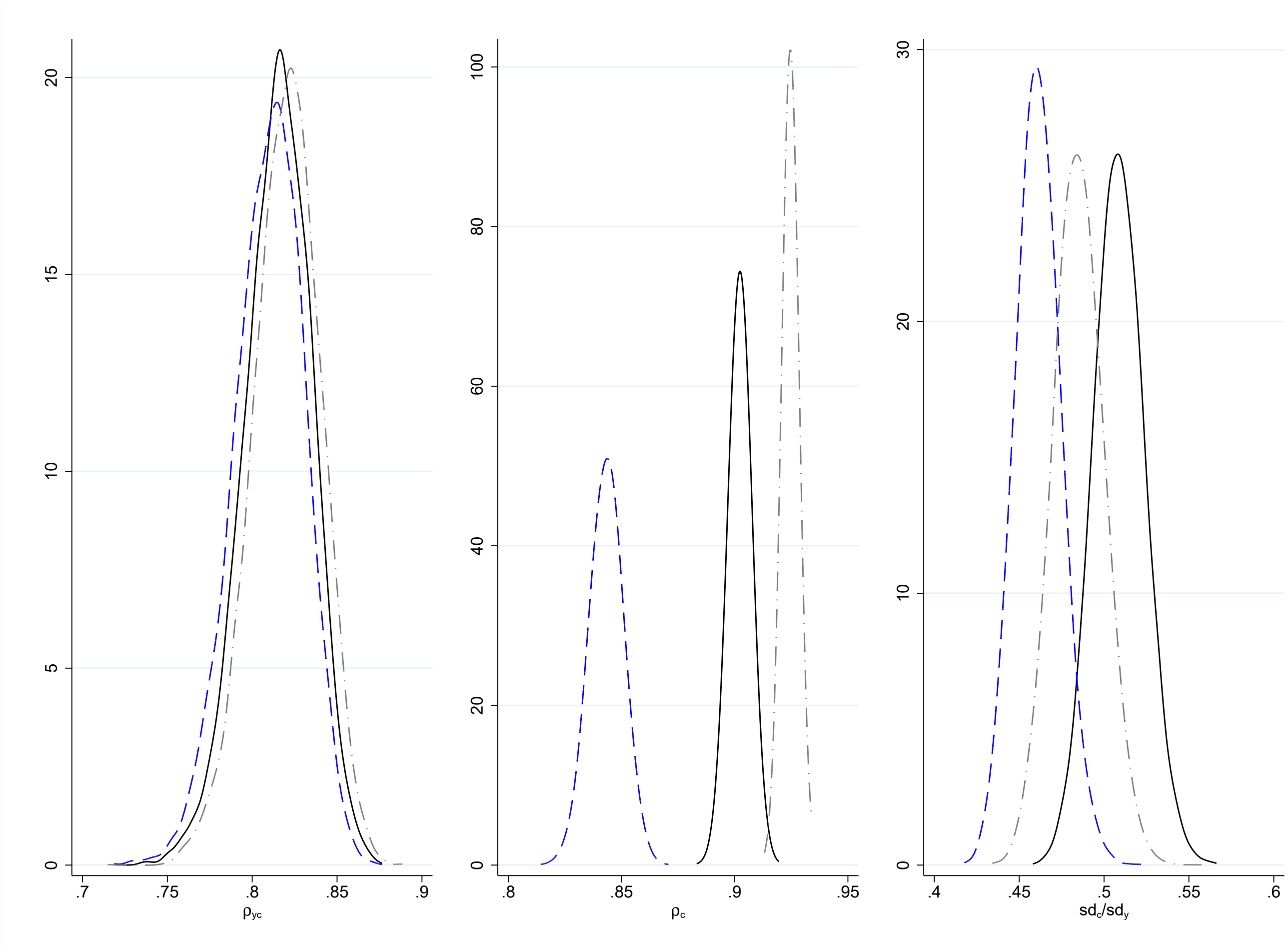

Figure 3: Kernel Approximation of the Moments' Posterior Densities, BVAR(1) on US Data. Solid line: one-sided HP Filter; Dashed line: two-sided HP Filter ($\lambda_{HP} = 1600$); Dash-dotted line: BN Filter (AR(12) with $\bar{\delta}_{BN} = 0.21$). Gaussian Kernels with fixed bandwidth ($bw = 0.003$).

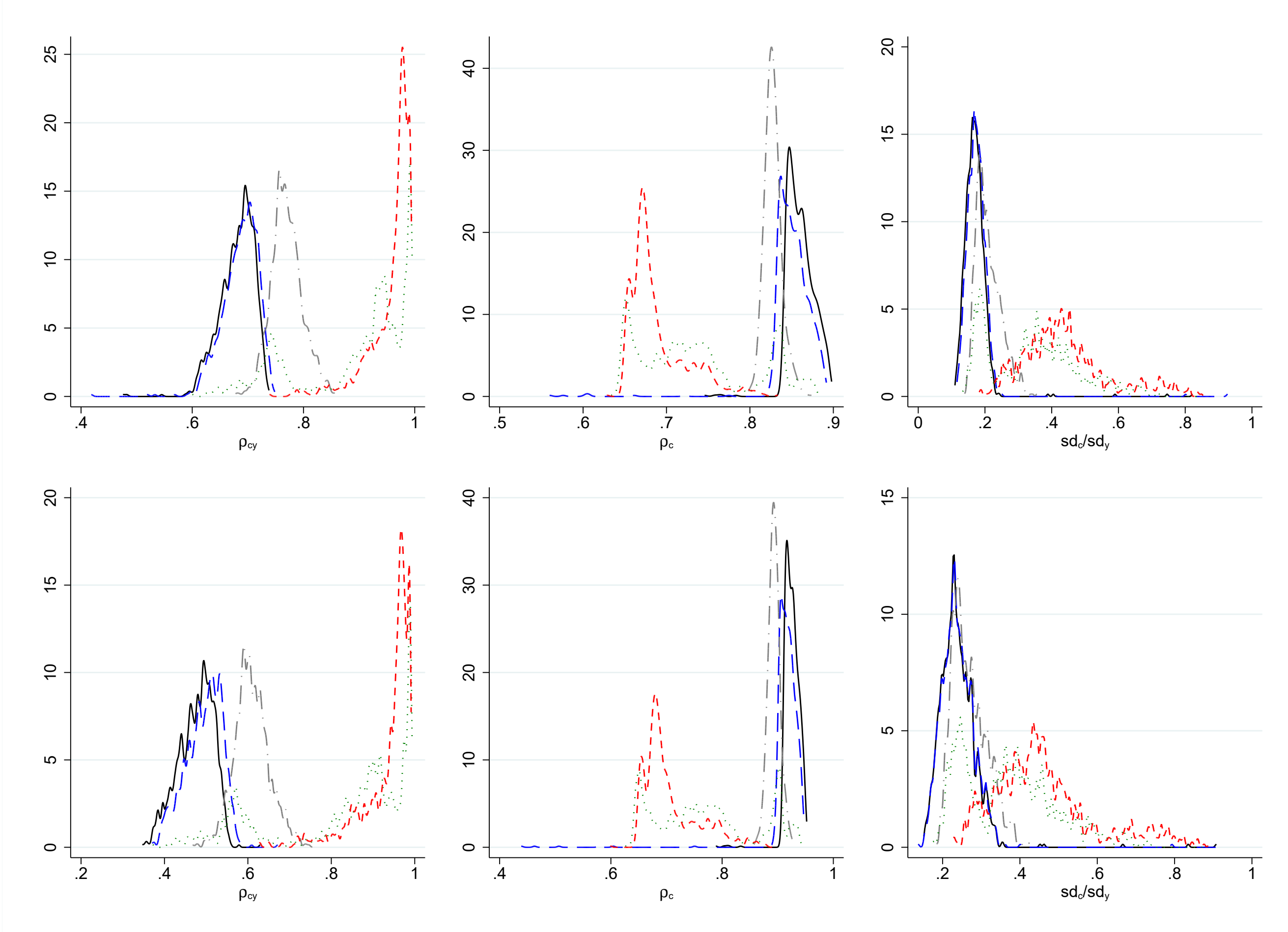

Figure 4: Kernel Approximation of the Densities of the Model-Generated Moments, Uncorrelated Uniform Priors. Top (Bottom) row: one-sided HP Filter (BN) Filter. Black line: $M0_{(CM)}$; Blue line: $M1_{(IM)}$; Grey line: $M2_{(\beta)}$; Green line: $M3_{(\sigma)}$; Red line: $M4_{(\beta,\sigma)}$.

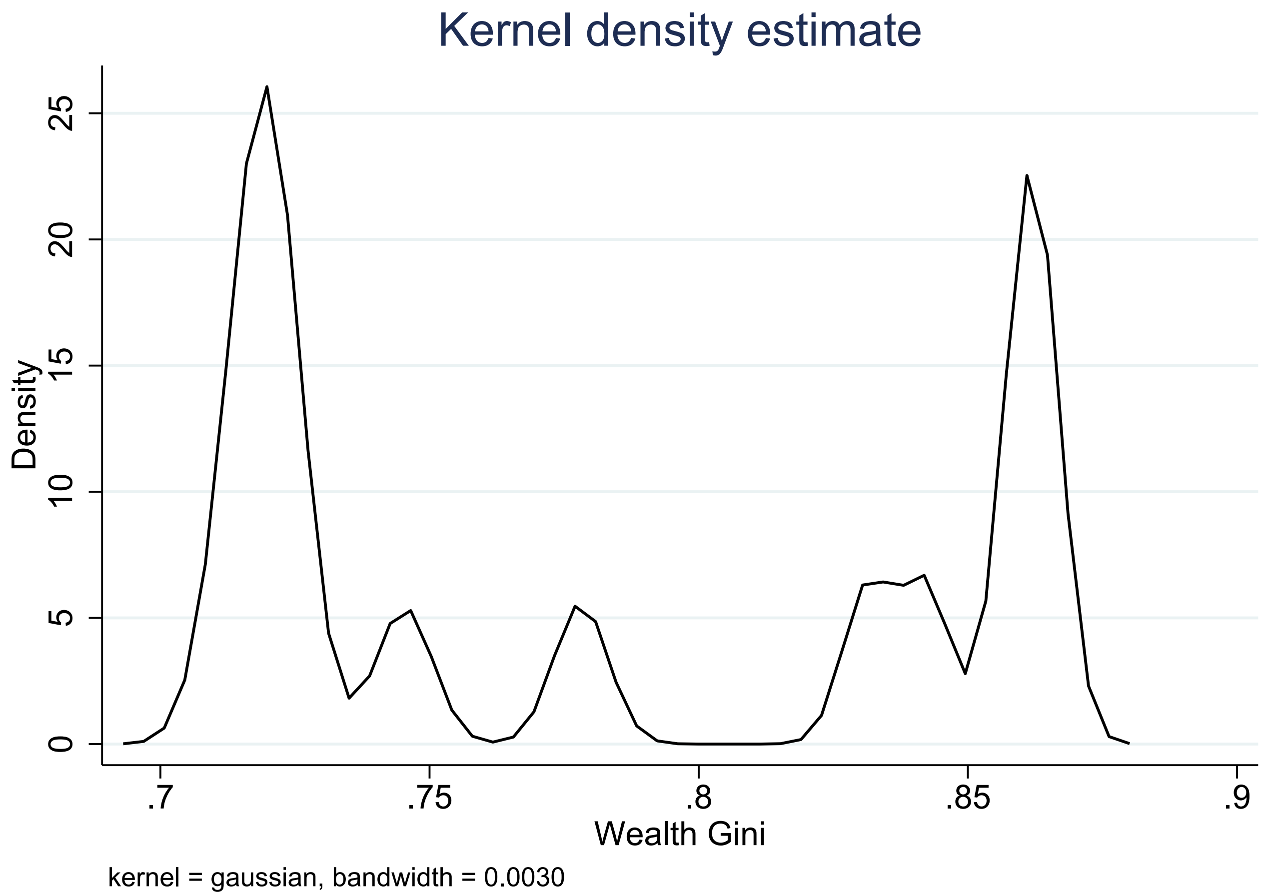

Figure 5: Kernel Approximation of the Mixture of Wealth Gini Index's Posterior Densities, PSID 1984-2019 Data. Each Cross Section is Truncated Below the First Quartile. Gaussian Kernel with fixed bandwidth ($bw = 0.003$).

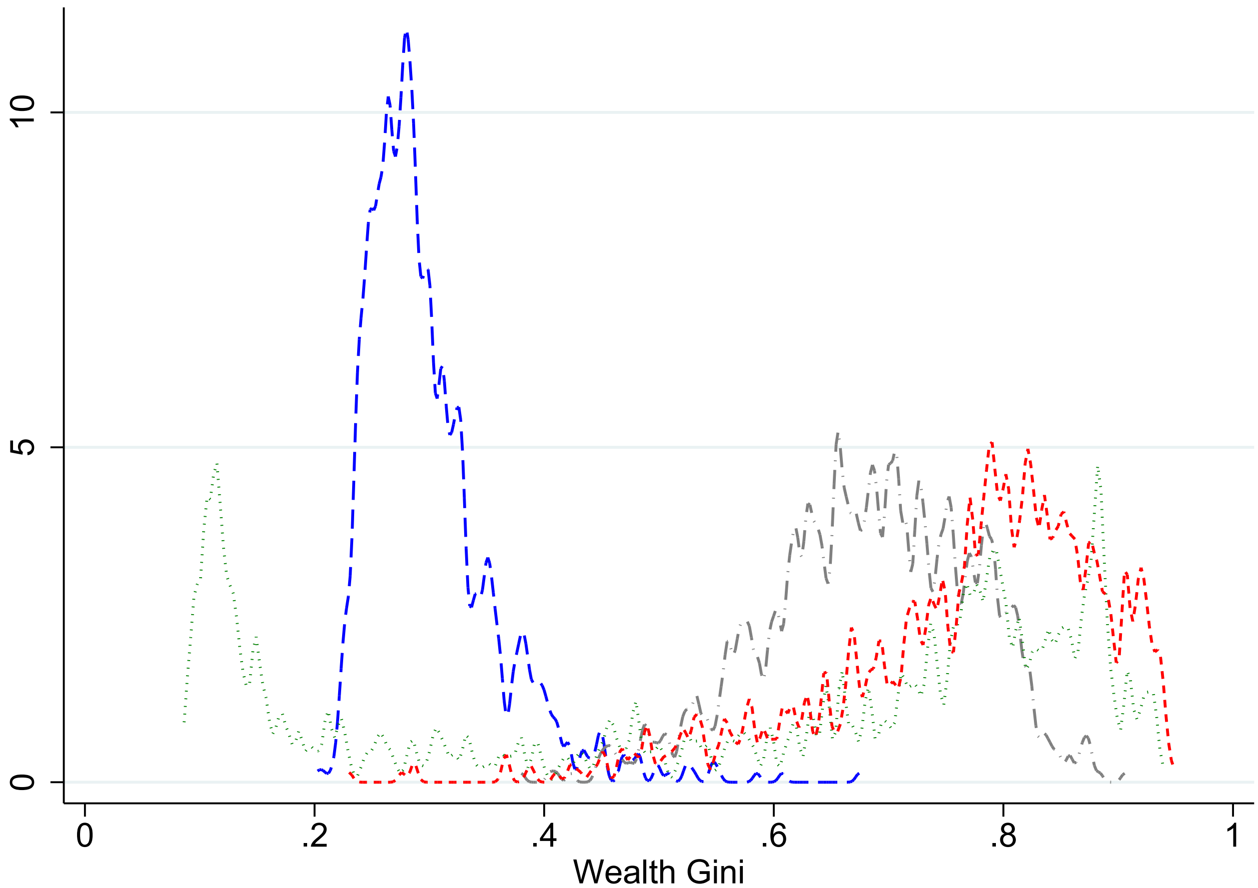

Figure 6: Kernel Approximation of the Densities of the Model-Generated Wealth Gini Index, Uncorrelated Uniform Priors. Blue line: $M1_{(IM)}$; Grey line: $M2_{(\beta)}$; Green line: $M3_{(\sigma)}$; Red line: $M4_{(\beta,\sigma)}$ Gaussian Kernels with fixed bandwidth ($bw = 0.003$).

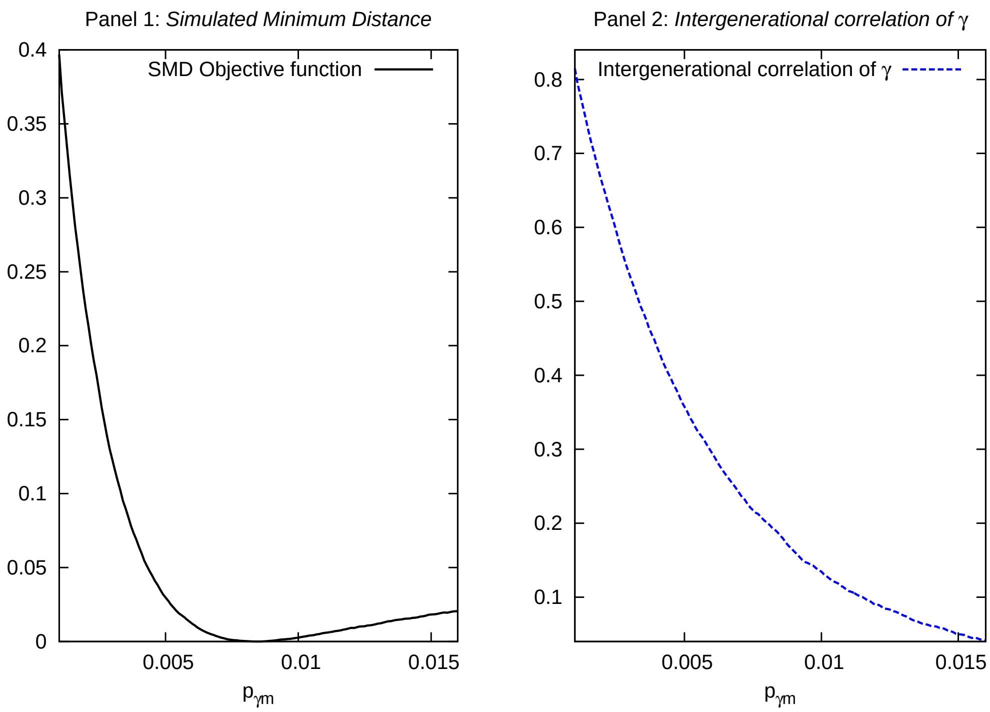

Figure 7: Identification of the CRRA Markov chain probability p_{γ_m}.

Comparison	*Macro BF (Case 1)*	*Macro BF (Case 2)*	*Macro BF (Case 3)*	*Macro BF (Case 4)*
$P(M0_{(CM)}\|.)/P(M1_{(0)}\|.)$	0.033	0.166	0.190	0.166
$P(M0_{(CM)}\|.)/P(M2_{(\beta)}\|.)$	0.000	0.000	0.000	0.000
$P(M0_{(CM)}\|.)/P(M3_{(\gamma)}\|.)$	0.000	0.000	0.000	0.000
$P(M0_{(CM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	0.000	0.000	0.000	0.000
$P(M1_{(0)}\|.)/P(M2_{(\beta)}\|.)$	0.000	0.000	0.000	0.000
$P(M1_{(0)}\|.)/P(M3_{(\gamma)}\|.)$	0.000	0.000	0.000	0.000
$P(M1_{(0)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	0.000	0.000	0.000	0.000
$P(M2_{(\beta)}\|.)/P(M3_{(\gamma)}\|.)$	0.070	0.305	0.002	0.320
$P(M2_{(\beta)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	0.070	0.285	0.005	0.503
$P(M3_{(\gamma)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	1.003	0.934	2.602	1.569
Comparison	*Average BF (Case 1)*	*Average BF (Case 2)*	*Average BF (Case 3)*	*Average BF (Case 4)*
$P(M1_{(0)}\|.)/P(M2_{(\beta)}\|.)$	0.005	0.000	0.000	0.000
$P(M1_{(0)}\|.)/P(M3_{(\gamma)}\|.)$	0.006	0.000	0.000	0.001
$P(M1_{(0)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	0.004	0.000	0.000	0.000
$P(M2_{(\beta)}\|.)/P(M3_{(\gamma)}\|.)$	0.354	0.555	0.693	0.575
$P(M2_{(\beta)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	0.271	0.429	0.275	0.598
$P(M3_{(\gamma)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	0.934	0.865	2.050	1.342

Table 9: Model Comparison: Bayes Factors, BN Filter.

Comparison \ *DGP*	$M1_{(IM)}$	$M2_{(\beta)}$	$M3_{(\gamma)}$	$M4_{(\beta,\gamma)}$
One-Sided HP Filter				
$P(M0_{(CM)}\|.)/P(M1_{(IM)}\|.)$	0.989	0.798	0.912	0.277
$P(M0_{(CM)}\|.)/P(M2_{(\beta)}\|.)$	2.344	0.333	3.099	0.0001
$P(M0_{(CM)}\|.)/P(M3_{(\gamma)}\|.)$	5.039	1.506	1.028	0.0001
$P(M0_{(CM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	903.175	43.077	267.004	0.0001
$P(M1_{(IM)}\|.)/P(M2_{(\beta)}\|.)$	2.371	0.417	3.397	0.0001
$P(M1_{(IM)}\|.)/P(M3_{(\gamma)}\|.)$	5.096	1.888	1.127	0.0001
$P(M1_{(IM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	913.485	53.989	292.651	0.0001
$P(M2_{(\beta)}\|.)/P(M3_{(\gamma)}\|.)$	2.149	4.526	0.332	0.001
$P(M2_{(\beta)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	385.275	129.414	86.155	0.001
$P(M3_{(\gamma)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	179.253	28.595	259.652	0.774
BN Filter				
$P(M0_{(CM)}\|.)/P(M1_{(IM)}\|.)$	0.976	0.585	0.776	1.000
$P(M0_{(CM)}\|.)/P(M2_{(\beta)}\|.)$	3.434	0.160	2.481	0.0001
$P(M0_{(CM)}\|.)/P(M3_{(\gamma)}\|.)$	5.664	0.790	0.888	0.0001
$P(M0_{(CM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	1950.789	42.291	397.940	0.0001
$P(M1_{(IM)}\|.)/P(M2_{(\beta)}\|.)$	3.517	0.273	3.198	0.0001
$P(M1_{(IM)}\|.)/P(M3_{(\gamma)}\|.)$	5.802	1.349	1.144	0.0001
$P(M1_{(IM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	1998.164	72.278	512.808	0.0001
$P(M2_{(\beta)}\|.)/P(M3_{(\gamma)}\|.)$	1.650	4.948	0.358	0.0001
$P(M2_{(\beta)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	568.151	265.059	160.371	0.0001
$P(M3_{(\gamma)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	344.414	53.564	448.286	0.614

Table 10: Model Comparison for the Monte Carlo analysis, $T = 3000$: Macro Bayes Factors. *Notes*: The true DGP is indicated by the column labels.

Comparison \ DGP	$M1_{(IM)}$	$M2_{(\beta)}$	$M3_{(\gamma)}$	$M4_{(\beta,\gamma)}$
One-Sided HP Filter				
$P(M0_{(CM)}\|.)/P(M1_{(IM)}\|.)$	0.991	0.794	0.900	0.594
$P(M0_{(CM)}\|.)/P(M2_{(\beta)}\|.)$	2.896	0.417	3.301	0.015
$P(M0_{(CM)}\|.)/P(M3_{(\gamma)}\|.)$	5.668	1.734	1.133	0.001
$P(M0_{(CM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	1159.762	52.129	282.778	0.0001
$P(M1_{(IM)}\|.)/P(M2_{(\beta)}\|.)$	2.922	0.525	3.668	0.025
$P(M1_{(IM)}\|.)/P(M3_{(\gamma)}\|.)$	5.718	2.185	1.259	0.001
$P(M1_{(IM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	1170.094	65.681	314.299	0.001
$P(M2_{(\beta)}\|.)/P(M3_{(\gamma)}\|.)$	1.957	4.162	0.343	0.036
$P(M2_{(\beta)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	400.449	125.108	85.676	0.029
$P(M3_{(\gamma)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	204.619	30.056	249.635	0.795
BN Filter				
$P(M0_{(CM)}\|.)/P(M1_{(IM)}\|.)$	0.897	0.531	0.686	0.145
$P(M0_{(CM)}\|.)/P(M2_{(\beta)}\|.)$	2.204	0.114	1.713	0.0001
$P(M0_{(CM)}\|.)/P(M3_{(\gamma)}\|.)$	4.467	0.572	0.527	0.0001
$P(M0_{(CM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	1194.104	19.528	199.551	0.0001
$P(M1_{(IM)}\|.)/P(M2_{(\beta)}\|.)$	2.457	0.214	2.496	0.0001
$P(M1_{(IM)}\|.)/P(M3_{(\gamma)}\|.)$	4.980	1.076	0.769	0.0001
$P(M1_{(IM)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	1331.308	36.761	290.886	0.0001
$P(M2_{(\beta)}\|.)/P(M3_{(\gamma)}\|.)$	2.027	5.017	0.308	0.0001
$P(M2_{(\beta)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	541.795	171.420	116.520	0.0001
$P(M3_{(\gamma)}\|.)/P(M4_{(\beta,\gamma)}\|.)$	267.313	34.168	378.321	0.622

Table 11: Model Comparison for the Monte Carlo analysis, $T = 150$: Macro Bayes Factors. *Notes*: The true DGP is indicated by the column labels.

The Models' Performance in other Dimensions: Tables 12 and 13 report the overlap between the 5-th and 95-th percentiles of the prior (model-generated) and posterior (BVAR-generated) densities of the three macro moments. Table 14 reports the overlap between the 5-th and 95-th percentiles of the prior (model-generated) and posterior densities of the wealth Gini index. The results are similar to what is found on the basis of the overall posteriors (joint, for the macro moments).

Moment	*Area (Case 1)*	*Area (Case 2)*	*Area (Case 3)*	*Area (Case 4)*
ρ_{CY}				
M0$_{(CM)}$ - *CM and No Heterogeneity*	0.267	0.271	0.273	0.271
M1$_{(IM)}$ - *No Preference Heterogeneity*	0.306	0.306	0.309	0.306
M2$_{(\beta)}$ - β *Heterogeneity*	0.805	0.783	0.789	0.745
M3$_{(\gamma)}$ - γ *Heterogeneity*	0.455	0.441	0.496	0.496
M4$_{(\beta,\gamma)}$ - β *and* γ *Heterogeneity*	0.282	0.300	0.275	0.354
ρ_{C}				
M0$_{(CM)}$ - *CM and No Heterogeneity*	0.798	0.795	0.796	0.795
M1$_{(IM)}$ - *No Preference Heterogeneity*	0.756	0.752	0.755	0.752
M2$_{(\beta)}$ - β *Heterogeneity*	0.608	0.613	0.625	0.630
M3$_{(\gamma)}$ - γ *Heterogeneity*	0.190	0.209	0.371	0.282
M4$_{(\beta,\gamma)}$ - β *and* γ *Heterogeneity*	0.032	0.037	0.022	0.059
sd_C/sd_Y				
M0$_{(CM)}$ - *CM and No Heterogeneity*	0.000	0.001	0.000	0.001
M1$_{(IM)}$ - *No Preference Heterogeneity*	0.000	0.000	0.000	0.000
M2$_{(\beta)}$ - β *Heterogeneity*	0.001	0.001	0.000	0.002
M3$_{(\gamma)}$ - γ *Heterogeneity*	0.244	0.248	0.111	0.193
M4$_{(\beta,\gamma)}$ - β *and* γ *Heterogeneity*	0.455	0.394	0.229	0.313

Table 12: Overlap between the 5-th and 95-th percentiles of the prior (model-generated) and posterior (BVAR-generated) densities of the three macro moments. One-sided HP Filter. A fixed bandwidth equal to 0.003 is used for all density approximations.

Moment	*Area (Case 1)*	*Area (Case 2)*	*Area (Case 3)*	*Area (Case 4)*
ρ_{CY}				
M0 $_{(CM)}$ - *CM and No Heterogeneity*	0.000	0.000	0.000	0.000
M1 $_{(IM)}$ - *No Preference Heterogeneity*	0.000	0.000	0.000	0.000
M2 $_{(\beta)}$ - β *Heterogeneity*	0.039	0.049	0.017	0.043
M3 $_{(\gamma)}$ - γ *Heterogeneity*	0.447	0.409	0.347	0.373
M4 $_{(\beta,\gamma)}$ - β *and* γ *Heterogeneity*	0.410	0.416	0.428	0.442
ρ_C				
M0 $_{(CM)}$ - *CM and No Heterogeneity*	0.897	0.894	0.899	0.894
M1 $_{(IM)}$ - *No Preference Heterogeneity*	0.888	0.882	0.893	0.882
M2 $_{(\beta)}$ - β *Heterogeneity*	0.838	0.838	0.849	0.845
M3 $_{(\gamma)}$ - γ *Heterogeneity*	0.268	0.283	0.479	0.384
M4 $_{(\beta,\gamma)}$ - β *and* γ *Heterogeneity*	0.051	0.060	0.029	0.095
sd_C/sd_Y				
M0 $_{(CM)}$ - *CM and No Heterogeneity*	0.009	0.014	0.003	0.014
M1 $_{(IM)}$ - *No Preference Heterogeneity*	0.009	0.013	0.003	0.013
M2 $_{(\beta)}$ - β *Heterogeneity*	0.039	0.045	0.013	0.043
M3 $_{(\gamma)}$ - γ *Heterogeneity*	0.364	0.346	0.174	0.309
M4 $_{(\beta,\gamma)}$ - β *and* γ *Heterogeneity*	0.545	0.496	0.385	0.428

Table 13: Overlap between the 5-th and 95-th percentiles of the prior (model-generated) and posterior (BVAR-generated) densities of the three macro moments. BN Filter. A fixed bandwidth equal to 0.003 is used for all density approximations.

Moment	*Area (Case 1)*	*Area (Case 2)*	*Area (Case 3)*	*Area (Case 4)*
Wealth Gini				
M1$_{(IM)}$ *- No Preference Heterogeneity*	0.014	0.001	0.001	0.001
M2$_{(\beta)}$ *-* β *Heterogeneity*	0.648	0.624	0.679	0.549
M3$_{(\gamma)}$ *-* γ *Heterogeneity*	0.548	0.492	0.265	0.420
M4$_{(\beta,\gamma)}$ *-* β *and* γ *Heterogeneity*	0.758	0.740	0.644	0.625

Table 14: Overlap between the 5-th and 95-th percentiles of the prior (model-generated) and posterior densities of the wealth Gini index. A fixed bandwidth equal to 0.003 is used for all density approximations.