Queen's Economics Department Working Paper No. 1277

Equilibrium Heterogeneous-Agent Models as Measurement Tools: some Monte Carlo Evidence

Marco Cozzi
Queen

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

8-2013

Equilibrium Heterogeneous-Agent Models as Measurement Tools: some Monte Carlo Evidence

Marco Cozzi, Queen's University

August 2013

Abstract

This paper discusses a series of Monte Carlo experiments designed to evaluate the empirical properties of heterogeneous-agent macroeconomic models in the presence of sampling variability. The calibration procedure leads to the welfare analysis being conducted with the wrong parameters. The ability of the calibrated model to correctly predict the long-run welfare changes induced by a set of policy experiments is assessed. The results show that, for the policy reforms with sizable welfare effects (i.e., more than 0.2%), the model always predict the right sign of the welfare effects. However, the welfare effects can be evaluated with the wrong sign, when they are small and when the sample size is fairly limited. Quantitatively, the maximum errors made in evaluating a policy change are very small for some reforms (in the order of 0.02 percentage points), but bigger for others (in the order of 0.6 p.p.). Finally, having access to better data, in terms of larger samples, does lead to substantial increases in the precision of the welfare effects estimates, though the rate of convergence can be slow.

JEL Classification Codes: C15, C54, C68, D52.

Keywords: Monte Carlo, Heterogeneous Agents, Incomplete Markets, Ex-ante Policy Evaluation, Welfare.

Acknowledgements: *I am grateful to two anonymous referees, Thor Koeppl, James MacKinnon, Lee Ohanian, Gregor Smith, and seminar participants at Queen's University, the SED meetings in Ghent, and the ESEM meetings in Oslo for useful comments and suggestions.* *Contact details: Department of Economics, Queen's University, 94 University Avenue, Kingston, ON K7L 3N6, Canada. Tel: 1-613-533-2264, Fax: 1-613-533-6668, Email: mcozzi@econ.queensu.ca.*

1 Introduction

Fact: quantitative macroeconomics is a burgeoning field. In this field, Heterogenous-Agent (HA) models have become a fruitful approach to conduct modern research in macroeconomics, as discussed in Krusell and Smith (2006) and Heathcote, Storesletten and Violante (2009). The need to go beyond the assumption of the representative agent for several important questions, Carroll (2000), the availability of cheap computational power, the increased access to both cross-sectional and panel data, together with a wider exposure of the profession to both numerical and recursive methods are among the reasons why macroeconomics research is performed more often with the aid of a computer, and with household level heterogeneity.

The typical project starts by posing a well defined economic question, develops a model firmly grounded on microfoundations (specifically tailored to address the question at hand, particularly for the role of heterogeneity), and numerically solves the model by implementing a calibration strategy. Often the aim is to use the theory to quantify a variable that cannot be directly measured for, say, lack of data (e.g. the extent of frictions in a market, as in Krusell, Mukoyama, Rogerson, and Sahin (2010), or the nature of unobservable shocks, as in Storesletten, Telmer and Yaron (2004) and Guvenen (2007)). Another common objective is to perform counterfactual analysis, by computing the welfare effects of a policy change together with its distributional impacts in an ex-ante policy evaluation, as in Alvarez and Veracierto (2001) and Fuster, Imrohoroglu and Imrohoroglu (2007), or its optimal scheme, as in Conesa, Kitao and Krueger (2009).

Ultimately, quantitative stochastic macroeconomic models are measurement devices. Important and influential contributions relying on HA quantitative models have found, for example, that the observed changes in the US wage structure led on average to welfare gains, Heathcote, Storesletten and Violante (2010), that a theory of uninsured income risk accounts for the US earnings and wealth inequality almost exactly, Castaneda, Diaz-Gimenez, and Rios-Rull (2003), that skill-biased technological change accounts for the evolution of wage inequality in the US, Heckman, Lochner and Taber (1998), that the size of precautionary savings is modest, Aiyagari (1994), just like the welfare costs of Business Cycles, Imrohoroglu (1989).[1]

These models are designed to quantify some variables of interest, being functions of a set of empirical facts: the features of the data they are asked to replicate in the calibration stage. However, little is known about their empirical properties, in particular their reliability as tools to perform welfare analysis. As noted in the past in the Real Business Cycle literature, when the objective of the computational experiment is empirical in nature, the calibration methodology is not free of potential pitfalls.[2] On the one hand, relying on external estimates to pin down some of the parameters leads to identification issues, whenever the assumptions of the empirical methodology estimating those parameters are not consistent with the ones of the macroeconomic model. On the other hand, sampling variability can be an obstacle that could cast doubts on some of the findings. More precisely, in face of sampling variability, it is not known how much the object of measurement is going to depend

[1] This list is by no means exhaustive: it is just a subset of the several important contributions in the HA macroeconomic literature. In particular, since the analysis will be in a steady-state, models with aggregate uncertainty are omitted. For a comprehensive survey see Heathcote, Storesletten and Violante (2009).

[2] See Gregory and Smith (1990), Canova (1995), Hoover (1995), and Hansen and Heckman (1996), among others.

on the tight requirement of asking the model to match exactly selected features of the data.[3]

This paper is going to deal with the latter aspect in two classes of equilibrium HA economies.[4] In order to shed some light on the empirical properties of these models I am going to propose and perform a series of Monte Carlo experiments.

First, I am going to specify a simple HA economy with Overlapping Generations (OLG). This is a version of the Huggett (1996) economy, appropriately modified to allow for an interesting public policy (the budget-neutral elimination of a capital income tax), and for several sources of heterogeneity in labor income (following the more recent literature, such as Storesletten, Telmer and Yaron (2004) and Heathcote, Storesletten and Violante (2010)).

In order to consider the empirical properties of another popular class of HA models, I am also going to specify a simple economy with infinitely lived agents. This is a version of the Huggett (1993) and Aiyagari (1994) economies, appropriately modified to allow for a different public policy (a reduction in the unemployment benefits) and for a clear-cut calibration of key parameters.

More in detail, the two economies are going to be production economies with an endogenous asset distribution, where a government collects taxes to finance either some public expenditure or an unemployment benefit scheme. Agents are ex-ante identical, while they differ ex-post, due to different realizations of a sequence of idiosyncratic shocks.

In the OLG economy, agents are going to face uninsurable income shocks. Two common parameterizations of the stochastic income process, estimated on PSID data, will be considered as the Data Generating Process (DGP). I will simulate the economy and I will draw 500 samples of different sizes, namely 200, 800, $3,200$, and $12,800$ individuals. For each simulated sample, I am going to estimate the same stochastic income process with a GMM estimator, obtaining sampling distributions for the persistence and variance of the income shocks. With these parameters, I am going to calibrate the model economy, solve it, compute the relevant endogenous variables, together with welfare measures under both the current policy regime and under a policy change. The procedure is going to deliver sampling distributions of the calibrated parameters and, more importantly, of the endogenous variables.

In the infinitely lived agents economy, the individuals are going to face an exogenous employment opportunity shock. This is a one-shock economy, where the sequence of shocks is independently distributed across the agents, and follows a two-state Markov chain. This is an important aspect, because the probabilities of the exogenous stochastic process do not depend on other parameters, and are uniquely identified: data on the unemployment rate and the unemployment duration are sufficient to pin them down. A baseline parameterization of the stochastic process, consistent with US labor market outcomes, will be considered as the DGP. I will simulate the economy and I will draw 500 samples of different sizes, namely $1,000$, $4,000$, $16,000$, and $64,000$ individuals.[5]

[3] These are among the reasons why carefully designed calibration exercises tend to provide some robustness checks, by perturbing the benchmark parameters.

[4] The general idea is not new: two decades ago some authors, Gregory and Smith (1991), Watson (1993), and Canova (1994), developed a set of methodologies to assess the empirical properties of RBC models.

[5] These artificial datasets are meant to mimic the size of the US datasets used more often by macroeconomists to estimate/calibrate the stochastic processes driving the uncertainty in the model economy, such as the CPS, and the NLSY/PSID.

For each replication, a sample with different realizations of the employment opportunity shocks is drawn. This is going to lead to non-degenerate distributions of two key moments: the unemployment rate and the unemployment duration. With these data in hand, I am going to follow Huggett (1993) strategy to calibrate the model, and I will follow the same steps outlined for the OLG model.

The contributions of this paper are several. First, to the best of my knowledge, this is the first Monte Carlo experiment dealing with equilibrium HA models. With this methodology it is possible to assess the small sample properties of these models as measurement devices. This allows to gauge some of their features as tools to conduct empirical analysis. Second, by varying the sample size it is possible to see how the behavior of the relevant statistics changes, namely if there is a quick convergence to the true values. Finally, ranges of estimates for the proposed policy changes are going to be provided.

For the policy reforms under analysis, I will be able to answer questions such as: How big is the largest mistake due to sampling variability that a quantitative macroeconomist can incur when evaluating a policy change with these models? Is there a tendency for the distribution of welfare effects to become substantially more concentrated when the researcher has access to more precise information (i.e., larger datasets, which lead to better calibrations)?

The most important finding is that, irrespective of the HA model, the postulated DGP, the specific policy change being examined, and the sample size in the Monte Carlo simulation, the calibration exercises hardly fail to assess the sign of the welfare change. Unless the true welfare effect is particularly small (i.e., less than 0.2%), the answer from the calibrated models has proven to be of the right sign. This is true both for the welfare enhancing cases and for the welfare reducing ones. However, the welfare effects can be evaluated with the wrong sign, when they are small and when the sample size is fairly limited.

As for the magnitude of the welfare effects, the results are less straightforward. Quantitatively, when calibrating from the largest sample, the maximum errors made in evaluating a policy change are always modest when considered in levels (between 0.03 and 0.1 percentage points). However, relative to true welfare effect, the maximum errors are small for some experiments (in the order of $\pm 2\%$ of the true value), but bigger for others (in the order of $\pm 30\%$ of the true value). Unfortunately, there does not seem to be a powerful rule-of-thumb suggesting when we are going to be in presence of small mistakes and when of large ones. These errors obviously get larger when considering smaller samples (in the worst cases considered here they get up to 0.6 p.p., and up to $\pm 100\%$ of the true value).

Moving from 200 to $12,800$ sampled individuals (and from $1,000$ to $64,000$) does lead to a decrease in the dispersion of the welfare effects. The model and the implementation of the welfare comparisons are consistent by construction: with knowledge of the fixed parameters, a law of large numbers applies, there is a unique stationary distribution, and a unique equilibrium. Hence, if we could sample an infinite number of agents for a long enough period of time, the calibration would deliver unique values for the free parameters, and the results would collapse to the population ones, recovering the DGP. The Monte Carlo experiments show that the distributions of welfare effects become more and more concentrated around the true values. However, the speed

For the OLG model, I use smaller samples to be consistent with the fewer observations that are typically available to estimate the income processes in panel datasets.

of convergence is faster in some experiments than in others. Working with the wrong parameters triggers an endogenous response of the model, due to its GE nature, leading to changing aggregates, prices, distributional features and welfare measures. In some experiments, this prevents the welfare effects from converging quickly to their true values.

The rest of the paper is organized as follows. Section 2 briefly presents the OLG model and describes the Monte Carlo experiments. Section 3 discusses the main results. Section 4 briefly presents the infinitely lived agents model and the related results. Section 5 concludes. Some appendices are also included: they discuss in more detail the models and the numerical methods used. They also present additional results and a set of robustness exercises.

2 The OLG HA Economy

Time is discrete. The economy is populated by a measure one of finitely lived agents facing an age-dependent death probability π_t^d. Age is denoted with t and there are T overlapping generations, each consisting of a continuum of agents. At age T_R all alive agents retire. The population grows at rate g_n.[6]

Preferences: Agents' preferences are assumed to be represented by a time separable utility function $U(.)$. Agents' utility is defined over stochastic consumption sequences $\{c_t\}_{t=1}^T$: their aim is to choose how much to consume (c_t), and how much to save in an interest bearing asset (a_{t+1}) in each period of their lives, in order to maximize their objective function. The agents' problem can be defined as:

$$\max_{\{c_t,a_{t+1}\}_{t=1}^T} E_0 U(c_0, c_1, ...) = \max_{\{c_t,a_{t+1}\}_{t=1}^T} E_0 \sum_{t=1}^T \beta^{t-1} \left[\prod_{j=1}^t \left(1 - \pi_j^d\right) \right] u(c_t)$$

where E_0 represents the expectation operator over the idiosyncratic sequences of shocks, and $\beta > 0$ is the subjective discount factor. I assume that $u(c_t) = \frac{c_t^{1-\theta}-1}{1-\theta}$, that is the per-period utility function is strictly increasing, strictly concave, satisfies the Inada conditions, and has a CRRA$=\theta$. Notice that there is no direct disutility from work, hence labor supply is fixed.

Endowments: Agents differ in their labor endowments $\epsilon_{t,\varepsilon,f}$. There are three channels that contribute to the determination of the total efficiency units that the workers supply in the labor market. First, there is a deterministic age component e_t, which is the same for all agents. Second, there is a stochastic component ε, whose log follows a stationary $AR(1)$ process: $\log \varepsilon_t = \rho_y \log \varepsilon_{t-1} + \xi_t$, with $\xi_t \sim N(0, \sigma_y^2)$. Third, there is a fixed effect component f, with half of the agents being born with the highest realization, and the other half with the lowest. These two values are chosen to match the variance of the fixed effect σ_f^2. The total efficiency units a worker is endowed with are the product of these three components, and labor earnings are $w\epsilon_{t,\varepsilon,f} = w \times e_t \times \varepsilon \times f$.

[6] In the interest of space, I am going to present just a sketch of the model. For more details see Appendix A, Huggett (1996), Storesletten, Telmer and Yaron (2004) and Rios-Rull (1999). The numerical methods and the computational algorithms are reported in Appendices C and D.

After the common retirement age T_R, the labor endowment drops to zero, and the agents receive a pension $\overline{y}_R$ paid for with the contributions of the economically active agents. The pension is a fixed replacement rate ϕ_R of the average labor earnings, and agents pay proportional taxes (τ_R) to contribute to the budget-balanced pension scheme. They also finance the public expenditure G and a public transfer TR with proportional capital (τ_a) and labor/pension earnings (τ_w). Agents can insure against their mortality risk in a perfectly competitive market for annuities. As a consequence, on average agents die with zero wealth and there are no accidental bequests. All agents enter the economy with a zero asset endowment and with the average realization of the stochastic component of labor earnings, which is normalized to 1.

Technology: The production side of the economy is modeled as a constant returns to scale technology of the Cobb-Douglas form, which relies on aggregate capital K and labor L to produce the final output Y.

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

Capital depreciates at the exogenous rate δ and firms hire capital and labor every period from competitive markets. The first order conditions of the firm give the expressions for the net real return to capital r and the wage rate w:

$$r = \alpha \left(\frac{L}{K}\right)^{1-\alpha} - \delta, \tag{1}$$

$$w = (1-\alpha)\left(\frac{K}{L}\right)^{\alpha}. \tag{2}$$

Notice that the marginal productivity of labor is always positive, hence the equilibrium labor input L will be always equal to the (exogenous) total supply of efficiency units. It follows that in the steady-state:

$$L = \sum_{t=1}^{T_R-1} \Phi_t \int_{\mathcal{A}\times\mathcal{E}\times\mathcal{F}} \epsilon_{t,\varepsilon,f} d\mu_t\left(a,\varepsilon,f\right)$$

where $\mu_t\left(a,\varepsilon,f\right)$ denotes the stationary distribution of working-age agents over their state variables, and Φ_t denotes the share of each cohort in the total population.

Government: The government carries out some public expenditure G and rebates some lump-sum transfers TR to every household. In order to finance the cost of these purchases and subsidies, both capital (τ_a) and labor (τ_w) taxes are levied. G, TR and τ_a are going to be policy parameters set exogenously, while labor taxes are set residually to guarantee a balanced budget.

Other market arrangements: The final good market is competitive, and firms hire capital and labor every period from competitive markets. Capital is supplied by rental firms that borrow from the agents at the risk-free rate r and invest in physical capital. There are no state-contingent markets to insure against the income risk, but workers can self-insure by saving into the risk-free asset. The agents also face an exogenous borrowing limit, denoted as b, which is set to $b = 0$ to avoid capital income taxes turning into subsidies for agents in debt.

2.1 Discussion

Why the focus on this specific OLG HA economy?

1. At its heart it represents the core of many OLG HA economies, sharing with them similar intertemporal trade-offs, insurance motives, and distortions.
2. As shown in the literature, e.g. by Storesletten, Telmer and Yaron (2004) and Conesa, Kitao and Krueger (2009), it accounts for several life-cycle facts, such as the evolution of earnings inequality and of wealth holdings.
3. Although it considers many layers of heterogeneity, it is simple enough to be efficiently solved on a computer with an extremely high numerical precision in a matter of minutes, allowing for a Monte Carlo experiment to be feasible.
4. It allows for a neat calibration procedure of key parameters, in particular the ones related to the income process.
5. The simple stochastic structure, together with the finite lifetimes assumption, allows to easily solve for the endogenous distributions numerically without resorting to simulations of samples of agents, which would lead to confounding effects in the welfare effects computation.[7]

2.2 Design of the Monte Carlo Experiment

The Monte Carlo experiment consists of two steps.

In the first step, I postulate the stochastic process for the labor income dynamics, parameterizing it according to two popular estimates on US data. As in the model, the stochastic component of labor earnings follows an $AR(1)$ specification. For the two parameters representing the persistence (ρ_y) and the standard deviation (σ_y) of the income shocks I consider, in turn, the estimates obtained by Guvenen (2009) and French (2005). I then solve the model economy and compute the welfare effects arising from a policy experiment. Coming from the DGP, these represent the true welfare effects.

In the second step, relying on the true stochastic income process, I simulate 500 samples of a given cross sectional dimension N from it. More in detail, I simulate the labor earnings dynamics of N individuals for $P = 20$ periods. This procedure generates a sequence of artificial panel datasets, and I estimate the two parameters ρ_y and σ_y with the Blundell and Bond GMM estimator.[8] I then re-calibrate the model's parameters according

[7] As shown in Michaelides and Ng (2000) in a Monte Carlo comparison of simulation estimators, exploiting the numerical approximation of the invariant distribution rather than simulating samples of agents provides substantial efficiency gains.

[8] I also considered some alternative dynamic panel data estimators, such as POLS and the Anderson-Hsiao ones, and the results were very similar. As customary in this literature, the estimation of the process for the persistent income shocks deals with a second stage regression, namely with the residual income uncertainty. Notice that with the specified DGP there is no need to use more sophisticated estimators to jointly estimate the distribution of the fixed effects and the age profile. A simple fixed effect estimator would recover its values virtually exactly, just like an appropriately specified first stage regression on age would recover the exact age effects. This is why I do not consider these variables in the simulations: across replications, they would not vary. In order to

to the estimates obtained in the simulated samples. With the new calibration, I implement the same policy change considered in the first step and compute the welfare effects induced by the new policy regime. These will (virtually always) differ from the true ones, because of the underlying parameterization being different from the DGP. A crucial aspect, related to the reliability of calibrated HA equilibrium models as empirical devices, will be how disperse the welfare effects estimates are going to be, and by how much they will differ from the true ones.

I repeat the second step four times (with samples of different sizes, to consider the effect of having more data, namely of size $N = 200$, 800, 3,200, and 12,800), and perform several robustness checks (by recalibrating in equilibrium another key parameter, β, and with different fixed parameters, to check if these affect the results).[9]

2.2.1 Parameterization

DGP: The baseline parameterization of the DGP relies on the estimates in Guvenen (2009): in the PSID, he finds $\rho_y = 0.988$ and $\sigma_y^2 = 0.015$. A second parameterization will follow French (2005) estimates: still in the PSID, he finds $\rho_y = 0.977$ and $\sigma_y^2 = 0.014$.

Monte Carlo Simulations and Calibration: The HA economy under study has several independent parameters: $\rho_y, \sigma_y, \beta, \sigma_f^2, \theta, \alpha, \delta, T_R, T, g_n, \pi_t^d, e_t$. These parameters are divided into three categories. In the first category there are two parameters, ρ_y and σ_y, which are "simulated". This means that they are going to be recalibrated at each replication of the experiment, on the basis of the specific estimates obtained in the sample simulation of the DGP. In the second category there are five parameters: $\beta, \sigma_f^2, \theta, \alpha, \delta$. In the baseline experiment, the discount factor β is going to be kept fixed, while in other experiments it is going to be recalibrated at each replication to match an equilibrium interest rate of 4%. The four remaining parameters are going to be kept fixed. However, a set of robustness checks with respect to their values are performed in order to understand if changing them affects the findings. Finally, the third category consists of all the demographics/life-cycle parameters: $T_R, T, g_n, \pi_t^d, e_t$. These are borrowed from other studies, and are always kept fixed. Agents become economically active at age 20, retire at age $T_R = 66$, and they can live up to $T = 101$ years. These are fairly conventional and secondary assumptions. The population growth rate g_n, the survival probabilities π_t^d, and the age profile for the efficiency units e_t are all taken from widely accepted studies, which are heavily used in the calibration of similar models.

A few words are in order, to explain why I limit the attention to calibrating only three parameters, while fixing the other ones. Recalibrating β is of primary importance. First and foremost, dealing with a capital income tax reform, the economy in its pre-reform equilibrium needs to display an empirically plausible rate of return of savings, 4% being the average between the rate of return of stocks and that of safe assets. Furthermore, with a Cobb-Douglas specification, recalibrating β allows to match almost exactly also the investment/output

meaningfully use minimum distance panel data estimators, I should either complicate the model, or add some measurement errors. However, choosing reasonable values for their variances is a non-trivial task.

[9] I set the number of periods in the panel dataset to $P = 20$, as this is one of the common restrictions imposed when estimating income processes from the Panel Study of Income Dynamics (see Heathcote, Perri and Violante (2010) for a thorough discussion). Decreasing this number to $P = 15$ doesn't have an important effect on the results.

ratio of 25% in all replications, without the need to adjust δ as well.[10] To conclude with, as argued before, appropriate regressions on the simulated data would retrieve the exact e_t profile, and the exact values of f and σ_f^2. Hence, there is no need to change these parameters in the replications.[11]

"Simulated" Parameters: The two parameters driving the labor income uncertainty in the economy are going to be assigned many different values, one for each replication of the simulated sample. The income shocks persistence and standard deviation (ρ_y and σ_y) are going to be recalibrated for every sample drawn at the iteration m of the Monte Carlo, according to the formulas of the Blundell and Bond GMM estimator (with the lagged labor income $y_{n,p-1}$ specified as the only covariate):

$$\widehat{\rho}_{y,m} = \frac{\sum_{n=1}^{N}\sum_{p=3}^{P} \triangle \log y_{n,p-1}^{m} \log y_{n,p}^{m}}{\sum_{n=1}^{N}\sum_{p=3}^{P} \triangle \log y_{n,p-1}^{m} \log y_{n,p-1}^{m}} \tag{3}$$

$$\widehat{\sigma}_{y,m}^{2} = \frac{1}{N\,(P-3)} \sum_{n=1}^{N}\sum_{p=3}^{P} \left(\eta_{n,p}^{m} - \overline{\eta_{n}^{m}}\right)^{2} \tag{4}$$

where $\triangle$ stands for the first difference operator, $\log y_{n,p}^{m}$ stands for the log of the labor income of individual n at period p computed in the simulated sample m, $\eta_{n,p}^{m}$ for the corresponding residuals, and $\overline{\eta_{n}^{m}}$ for their individual-specific time average.

[Figures 1 and 2 about here]

Figures (1) and (2) show the distributions of the two parameters resulting from the estimation procedure. Figure (1) presents the non-parametric kernel estimate of four densities (one for each sample size in the simulation stage) of the income shocks persistence ρ_y. These are computed relying on equation (3), once for every artificial dataset. Figure (2) presents a similar graph for the standard deviation of the income shocks σ_y, computed relying on equation (4). As expected, with small samples the distributions are quite dispersed and sampling more individuals from the true stochastic process leads to more precise estimates for both parameters.

[10] See Figure (12) in Appendix E for the plot of the investment/ouput densities with the calibration in equilibrium of β, and Figure (13) for the ones without it. Notice how in the former case all the densities collapse to a spike at the 25% value. This happens because a 4% interest rate is obtained with a capital/output ratio of approximately 3, which is the ratio needed to get the desired investment rate along the balanced growth path, as $I/Y = (\delta + g_n)\,K/Y$.

[11] Although conceptually feasible, it is not straighforward that matching the labor shares and investment rates obtained in the simulations is desirable. First, these variables are genuinely macro aggregates, namely they are computed from data sources (e.g., NIPA and CPS) that are different from the panel datasets used to estimate the income processes. Furthermore, this would require to assume the availability of reliable data on the change of asset holdings (to compute the households' aggregate savings, hence the economy's total investment), a type of information which is usually hard to get in real household-level panel data. A similar comment applies to the CRRA, which could be identified by the simulated consumption growth rates. Finally, this alternative procedure would lead to several parameters changing at once, making the interpretation of the results less transparent.

[Table 1 about here]

Table (1) provides a set of descriptive statistics for the two income dynamics parameters. It is worth stressing that the parameters' range used to solve the model and compute the related welfare effects is quite large. As for ρ_y, with only 200 individuals at our disposal, this parameter ranges from 0.9379 to 0.999. This is less so for the other parameter, the standard deviation of the income shocks σ_y, which appears to be always fairly concentrated. In the 200 individuals case this parameter ranges from 0.1172 to 0.1299. Given the parameters' dispersion, the amount of idiosyncratic uncertainty in the economy varies substantially, getting less dispersed with larger samples. With $12,800$ artificial individuals ρ_y ranges from 0.9794 to 0.9952, while σ_y ranges from 0.1217 to 0.1231. Notice how the median values are always very close to the DGP, irrespective of the sample size. As for ρ_y, there is one fact to point out. In the top two panels of Figure (1) the densities are bimodal, with one of the modes being close to the value of 1.0. As the income process must be stationary, all estimates $\widehat{\rho}_{y,m} > 1.0$ have to be capped at 0.999 instead. This happens often in the two smaller samples, where with 200 (800) individuals 28.6% (11.4%) of the estimates in the replications are constrained ex-post to be inside the unit circle. Differently, in the two larger samples, this happens in 1% (0%) of the replications. Another consequence of this fact concerns the rate of convergence of the parameters. For ρ_y this is close to $\sqrt{N}$ only in the unconstrained cases, which visually can be seen by moving from one panel to the next, with the estimated densities doubling their maximum value. Differently, σ_y shows a $\sqrt{N}$ convergence behavior virtually in all cases, even though the distributions are not always symmetric around the true value. These findings will be useful in interpreting some of the welfare effects results.

Fixed Parameters: The concavity of the utility function is pinned down by the CRRA coefficient θ, which is set to 1.5, a common value in the literature. For the reason outlined above, the borrowing limit b is set to zero. The capital depreciation rate is set to replicate an investment/output ratio of approximately 25.0%. This is achieved with $\delta = 0.07$. With a Cobb-Douglas production function the capital share is captured by the parameter $\alpha = 0.34$, which matches the capital share of income. The rate of time preference β is calibrated to get an equilibrium interest rate of approximately 4% on an annual basis, obtained when $\beta = 0.97244$, with a capital/output ratio of 3.1. The income process studied by Guvenen (2009) also provides an estimate for the variance of the fixed effect, which is set accordingly to $\sigma_f^2 = 0.058$. The complete parameterization of the model is reported in Table (2).

[Table 2 about here]

Demographics/Life-cycle "Consensus" Parameters: The US long-run average of the yearly population growth rate is 1.1%, hence $g_n = 0.011$. For the life-cycle aspects of the model, I consider standard profiles for both the survival probabilities, from Bell and Miller (2002), and for the efficiency units e_t, from Hansen (1993).

Policy Parameters: Differently from the parameters above, τ_a, G, TR and ϕ_R are policy parameters. They are pinned down by the institutional features of the economy they are meant to represent (its fiscal policies and social security system, in particular). Their values are reported in Table (3), and will be explained in detail in the policy experiments description. Notice that I consider a lump-sum transfer to allow for some progressivity in the income tax code, and that I don't consider experiments with respect to ϕ_R, because of the stylized pension scheme I rely on. This parameter's calibrated value corresponds to the pension replacement rate of the average individual earnings, computed for the US economy in OECD (2011): $\phi_R = 0.394$.

[**Table 3 about here**]

2.2.2 Model Solution and Policy Changes

For each set of calibrated parameters the equilibrium of the model is computed twice. The first time under the current policy regime, i.e. for specific values of the triplet $\{\tau_a, G, TR\}$, the second time under the counterfactual economy, i.e. after a policy change.[12] The policy reform is going to deal with the elimination of the capital income tax. This is meant to highlight the distortionary effects of capital taxation, the endogenous response of savings, and the concurrent shift in taxation from capital to labor income.

Moreover, notice that in the numerical solution of the model there is no sampling variability (due for example to Monte Carlo integration). The only sources of error are the numerical errors induced by the discretization of the state space and by the convergence criteria, which are kept as small as the computational burden makes possible. It follows that, across replications, the computed equilibria vary only because of the different calibrated parameters. Differently, the change in equilibria would partially reflect the simulation error: aggregate quantities would vary randomly, leading to an induced endogenous response of the model and an additional source of error for the welfare effects.

2.2.3 Experiment 1 - Capital Taxation and Public Expenditure

This experiment deals with a change in the tax code, starting from a very distortionary capital income taxation τ_a.[13] There are exogenously given public expenditure G and transfers TR. They are both financed by collecting taxes on capital and labor income. G, TR and τ_a are set by a policy maker, while τ_w is set residually to ensure a balanced budget.

In the initial policy regime, capital income is heavily taxed. This leads agents to save little, as the after-tax rate of return is low. At the same time, labor supply is inelastic, so taxing labor does not imply any distortions. Agents cannot reduce their labor supply and do not accumulate human capital while on the job. Unexpectedly,

[12] The welfare effects are going to compare two different steady-states. I consider both the percentage change in aggregate welfare, and a consumption equivalent welfare measure, denoted with ϖ. See equations (12) and (13) in Appendix A. Considering transitional dynamics for the Monte Carlo procedure is computationally infeasible.

[13] An analogous policy reform is studied in a simpler economy by Imrohoroglu (1998), the desirability of changing the current tax code to a flat one by Ventura (1999), and the characterization of the optimal capital tax by Conesa, Kitao and Krueger (2009).

the government implements a budget-neutral tax cut: $\tau_a = 0.3$ in the benchmark economy and $\tau_a^{new} = 0$ in the counterfactual one. This simple reform implies that, for a given level of public expenditure, lower capital taxes are going to increase the incentives to save, leading to a larger capital stock, and higher wages. Overall this explains why in the baseline case there is a welfare gain from reducing the capital tax, even though a priori the policy reform could lead to a welfare loss (labor income is going to be taxed more, which is very unattractive for -income and wealth poor- young individuals).

Case 1a): For this case $\tau_a = 0.3, \tau_a^{new} = 0, \frac{G}{Y} = \overline{\frac{G}{Y}} = 0.2$ and $\frac{TR}{Y} = \overline{\frac{TR}{Y}} = 0.05$. This specification of G matches the public expenditure/output ratio in the initial steady-state, which is approximately 20%. It also keeps the government share of output $\left(\frac{G}{Y}\right)$ in the counterfactual economy constant at 20%. Likewise, the public transfers/output ratio $\left(\frac{TR}{Y}\right)$ is assumed to be 5% in both the initial and final steady-states. This term captures in a parsimonious way the progressivity of the income tax code, together with the redistribution intrinsic in many other insurance and safety net programs, such as Medicare/Medicaid and the food stamps ones, as discussed in Aiyagari and McGrattan (1998).

For constant total public outlays as a share of income $\overline{\frac{G}{Y}} + \overline{\frac{TR}{Y}}$, a decrease in capital taxes τ_a mechanically leads to an increase in labor taxes τ_w.[14] However, labor supply is rigid in this economy and we are getting closer to the principles of Ramsey taxation: higher taxes should be set for the goods with lower elasticities of substitution. Notice also that the decrease in distortionary taxation has a GE effect as well: a higher return on capital increases savings and investment, leading to a higher output and to a lower interest rate. It follows from these considerations that, if I were to assume a constant G in the counterfactual economy, the government share of output would be less than 20% in the final steady-state, leaving more resources for consumption and investment. This outcome would inflate the welfare effects, which explains the assumptions I work with.

The true welfare effect is sizable, and it is quantified in a 1.067% increase of the average steady-state utility. The corresponding figure for the consumption equivalent welfare measure is even larger, being $\varpi = 2.175\%$. The increased capital supply raises output and wages, which more than compensate the drawbacks of the increased taxation of labor earnings.

Case 1b): For this case, both the policy experiment and the DGP are the same as in case 1a, namely $\tau_a = 0.3, \tau_a^{new} = 0$. Differently from before, the parameter β is no longer kept fixed, and it is recalibrated in each replication to match a 4% interest rate in the pre-reform equilibrium. It goes without saying that, being the DGP identical, also the true welfare effects are the same as in case 1a.

Welfare Effects Sampling Distributions: Figures (3) and (4) plot the non-parametric kernel densities of the welfare effects obtained in the four Monte Carlo experiments for the cases 1a and 1b.[15]

[14] See equation (9) in Appendix A.

[15] A side remark is that, for this economy, studying the welfare effects directly in terms of the average steady state utility change vis-à-vis a consumption based welfare measure turns out to be immaterial. When considering the latter, there is just a level effect, with the shape of the distributions and the convergence behavior being identical. For example, compare Figure (3) to Figure (11) in Appendix E. Hence, in the remainder of the paper, I will focus only on the consumption equivalent measure ϖ.

[Figures (3) and (4), and (5) about here]

The first consideration arising from inspecting the supports of the sampling distributions is that, in both cases, the computed welfare effects portray the policy as leading unambiguously to a welfare gain. Irrespective of the sample size and the richness of the calibration, ϖ is bounded below by a 2% increase. Furthermore, the plots show that the economy-wide welfare effects of the capital tax reforms are converging at a fast rate to their true value, especially in case 1a. Notice how the additional calibration in equilibrium of β leads to more imprecise welfare effects, for any sample size N. This is due to the additional dispersion of the discount factor, whose distribution can have a long tail when the sample size is fairly small, as shown in Figure (5). In particular, when $N = 200$ the range of β is $[0.97197, 0.97421]$, but the median is very close to the true value, being $\beta^{med} = 0.97246$.

The bimodal distributions of the parameter ρ_y are reflected in the non-symmetric shape of the welfare effects' densities. As the sample size grows, there are fewer constrained persistence parameters, and the distribution of the welfare effects eventually becomes bell-shaped. To conclude with, the top left panels of the two figures (referring to $N = 200$) show that a small sample can lead to relatively large mistakes in the assessment of the aggregate welfare, because of the more likely extreme calibrations. Differently, the bottom right panels (referring to $N = 12,800$) show that the welfare effect distributions are very concentrated around their true value.

Cases 1c-d): These two additional cases consider the same policy reform, but rely on a different DGP. Once more, $\tau_a = 0.3, \tau_a^{new} = 0, \frac{G}{Y} = \overline{\frac{G}{Y}} = 0.2$ and $\frac{TR}{Y} = \overline{\frac{TR}{Y}} = 0.05$. However, following French (2005), now $\rho_y = 0.977$ and $\sigma_y^2 = 0.014$. As expected, under the new parameterization, the true welfare effect changes, now being $\varpi = 2.112\%$, with a recalibrated $\beta = 0.9729$. The consequences of considering this alternative DGP are minimal. Case 1c deals with a fixed β across replications, while case 1d recalibrates it in equilibrium. As for the welfare effects sampling distributions, qualitatively there are no substantial differences from cases 1a-b, and all the related Figures and Tables are reported in Appendix E. The only outcome worthy of being mentioned is the "nicer" shape of the distributions, driven by the lower number of replications displaying a capped ρ_y.

2.3 Robustness Checks

A thorough robustness analysis is performed. In particular, other plausible calibrations are considered, setting alternative values for every fixed parameter. Table (4) lists all the cases, together with the recalibrated discount factor β and the true value of the policy change ϖ that applies to the new calibration.[16]

[Table 4 about here]

[16] Table (5) reports a set of statistics of the welfare effects for the baseline cases, while Table (6) reports the same statistics for the robustness checks, when β is recalibrated in each replication. These will be commented upon in the next section. All the plots of the related sampling distributions can be found in Appendix E.

Capital depreciation rate (δ): The concept of investment used for the baseline calibration is quite broad, because it includes non-residential and residential fixed investment and purchases of consumer durables. Narrower definitions lead to lower investment/output ratios: for example, in the PENN World Tables the average investment share over the period 1990-2010 (of PPP converted GDP per-capita at current prices) has a value of 22.3%. Hence, it is worthwhile investigating a lower share of 22.5%. This is achieved with $\delta = 0.0457$. The implications for the true welfare effect are quantitatively non-negligible, now being $\varpi = 2.792\%$. A lower δ implies a lower user cost of capital, hence the economy accumulates more capital and the benefits of eliminating the capital income tax are more pronounced. However, although the equilibrium interest rate is still 4%, the capital/output ratio is now 3.97, which is on the high end of the available estimates. In terms of the Monte Carlo experiment, the resulting sampling distributions appear to be similar to the previous ones, but they are shifted to the right and more dispersed. The overall patterns are analogous to the baseline calibration, even though the rate of convergence is somewhat slower.

Capital share (α): The long-run average of the labor share leads to a value for α which is around 1/3, corresponding to the typical assumption for this parameter. However, this share varies over the cycle and during the last dozen years a very persistent increase in its value has been observed. In 2012 $\alpha = 0.38$, up steadily from $\alpha = 0.34$ in 2001. This is why I also consider a value of $\alpha = 0.4$. The implications for the true welfare effect are less drastic, now being $\varpi = 2.420\%$. As for the Monte Carlo experiment, the sampling distributions of ϖ are comparable to the ones obtained with the lower δ, but they are even more dispersed and the rate of convergence appears to be slightly slower.

Variance of the fixed effect (σ_f^2): In the baseline calibration, I followed the results in Guvenen (2009) and set $\sigma_f^2 = 0.058$. In the literature, however, there are other estimates for this parameter. In particular, in their influential work Storesletten, Telmer and Yaron (2004) find $\sigma_f^2 = 0.211$, which is almost four times larger. I also use this calibration, which implies that the two values of f are moving farther apart. The implications for the true welfare effect are really minor, now being $\varpi = 2.064\%$. The lower ϖ is due to the high fixed effect types enjoying higher earnings, consumption and welfare, while the low types have the opposite outcomes, are worse off and more than compensate the former group's welfare improvement. Both their lower labor income and their increased precautionary savings lead them to consume less. This increases the marginal utility of consumption of young individuals with the low fixed effect, that suffer a utility loss when the labor tax increases, explaining why there is a lower aggregate welfare effect. Interestingly, in this case the rate of convergence is pretty high.

CRRA (θ): The robustness analysis shows that the welfare effects are very sensitive with respect to the choice of θ, a result already found by both Imrohoroglu (1998) and Conesa, Kitao and Krueger (2009). With $\theta = 3.0$, a calibration which is well inside the range of empirically relevant values, the true welfare effect is now only $\varpi = 0.284\%$, representing an almost eightfold decrease. Nevertheless, the shape and the rate of convergence of the welfare effects sampling distributions are not extensively affected.

The lower the EIS (=1/CRRA) parameter, the greater the negative consequences in terms of welfare for the younger agents. Effectively, the elimination of the capital income tax partially shifts the burden of the public expenditure from older agents, that hold a lot of wealth, to younger ones, that are wealth poor. The increase in labor taxes decreases their disposable income, reducing their ability to achieve consumption smoothing, with a higher θ making it more costly in utility terms. It turns out that it is possible to find a θ high enough for the steady state welfare to change sign: with $\theta > 3.2$ the economy-wide welfare effect is now negative. Although this result raises concerns on the ex-ante evaluation of such a policy, it allows to conduct further robustness checks, when the true welfare effect is quantitatively small. I parameterize the economy to imply, by construction, a true welfare effect bounded away from zero, but positive and modest. As many calibration studies routinely find welfare effects that are substantially less than 1%, this seems a worthwhile investigation. Starting from the baseline parameterization, I set $\theta = 3.15$, in order for the true welfare effect to be a fraction of a percentage point.[17] More precisely, the value of the consumption equivalent welfare measure in this case is $\varpi = 0.091\%$, such that the Monte Carlo experiment is faced with a challenging task. As (plausibly) expected, in this case the calibrations sometimes deliver estimates of the welfare effect that have the wrong sign. However, also in this extreme case, the results seem to provide evidence supporting the use of HA models as quite reliable tools to estimate welfare effects. In this experiment, it is sufficient to have $3,200$ individuals for this to happen in only 2.2% of the replications. Furthermore, the sign reversal never happens with the largest sample of $12,800$ individuals, while with 800 (200) individuals this happens in 19.4% (30.2%) of the cases.[18]

3 Estimating Welfare Effects with a Calibrated OLG HA Model

What happens to the estimated welfare effects obtained from the Monte Carlo experiments? In general, working with a calibrated model with parameters that differ from the true ones leads to errors that many would consider tolerable. Tables (5) and (6) report several statistics of the economy-wide welfare effects computed in each experiment.

[Tables 5 and 6 about here]

When the true welfare effect was sufficiently far from zero, even as small as $\varpi = 0.2\%$, the calibrated HA model always got the right sign of the welfare effects. Irrespective of the data abundance and the alternative calibrations of the fixed parameters, the quantified welfare effects were always of the same sign as the true ones. Unfortunately, this was no longer true in the absence of a very large panel dataset, when the true ϖ was deliberately close to zero.

Another positive, but perhaps not surprising, aspect is that the range of the welfare effects decreases monotonically with the sample size. For example, for case 1a, it moves from $[2.060, 2.239]$ in the least pre-

[17] I always recalibrate the discount factor to match a 4% interest rate, now attained with $\beta = 0.9868$ when $\theta = 3.0$, and with $\beta = 0.9878$ when $\theta = 3.15$.

[18] It is relatively rare that researchers estimate the stochastic income processes relying on less than $1,000$ individuals, even though for some applications the need to consider some specific subgroups leaves the researcher with smaller samples. For example, the selection criteria in Guvenen (2009) leave $1,270$ individuals, which drop to 335 and 882 when further splitted into two education categories.

cise case to $[2.128, 2.219]$ in the most precise one. Both the upper bound and the lower bound of the welfare effects support are shrinking considerably with the sample size.

An implication of this behavior is that the maximum error that we could incur when evaluating the welfare effects of a policy change decreases with the quantity of the information at our disposal. In the experiment 1a (1b) the maximum error never exceeds 0.115 (0.154) percentage points, and its absolute value ranges between 2.04% and 5.28% (2.30% and 7.08%) of the true value. Also the alternative DGP shows similar outcomes. In experiment 1c (1d), the maximum error never exceeds 0.129 (0.134) percentage points, and it ranges between 0.51% and 6.11% (1.55% and 6.33%) of the true value.

Setting aside the CRRA robustness check for a moment, the experiments related to the other robustness checks showed errors similar in magnitude to the cases 1a-d. The maximum error never exceeds 0.303 percentage points (obtained with $\alpha = 0.4$) and the range of the maximum error is between 1.45% and 12.52% of the true value. In all these cases, the model proved to be a sound measurement tool.

The means and medians of each Monte Carlo experiment are remarkably close to the corresponding true values. If we consider the model as an estimator for the welfare effects, it displays a small sample bias that does not seem to have a regular pattern with the sample size. Hence, the average of the welfare effects sampling distribution does not coincide with the true effect. The bias, though, is quantitatively negligible for all possible experiments, being at most 1.40% of the true value. This result suggests that a well designed sensitivity analysis can go a long way in quantifying a reliable estimate of an ex-ante policy evaluation.

A different set of considerations apply to the CRRA robustness check. Also in this case, the maximum error is limited in size, being at most 0.28 percentage points, but it ranges between 27.09% and a whopping 98.31% of the true value, due to the very small true ϖ. More importantly, it was found that a calibration study can estimate the wrong sign of a policy reform. Although not excessively severe, this issue calls for the implementation of a systematic robustness analysis of HA models aimed at quantifying the welfare effects of a policy change.

The rate at which the welfare effects distributions converge to their true values differ in the various experiments under analysis. In case 1d the sampling distributions show a slow rate of convergence, unlike in cases 1a and 1c, which seems to be collapsing to their true values at quite a fast rate. The welfare effects sampling distributions do get more concentrated around the true value, but the speed of convergence is often slower than what is observed for the underlying parameters.

4 The Infinitely Lived HA Economy

Another popular class of HA models considers infinitely lived agents subject to uninsurable income shocks. In this section, I specify a very simple model in this class, and perform a different policy experiment. Time is discrete. The economy is populated by a measure one of infinitely lived ex-ante identical agents.[19]

[19] For a more detailed description of the model, see Appendix B, Huggett (1993), Aiyagari (1994), and Rios-Rull (1999).

Preferences: The agents' problem can be defined as:

$$\max_{\{c_t,a_{t+1}\}_{t=0}^{\infty}} E_0 U(c_0, c_1, ...) = \max_{\{c_t,a_{t+1}\}_{t=0}^{\infty}} E_0 \sum_{t=0}^{\infty} \beta^t u(c_t)$$

where E_0 represents the expectation operator over the employment opportunity shocks $s \in S = \{e, u\}$. $\beta \in (0,1)$ is the subjective discount factor, and I assume that $u(c_t) = \frac{c_t^{1-\theta}-1}{1-\theta}$.

Endowments: Agents can be employed (e), or unemployed (u). If employed, they earn a wage w and pay proportional taxes (τ_u) to finance the unemployment benefit scheme. If unemployed, they collect unemployment benefits, which are specified as a constant replacement rate ϕ_u of the going wage. The stochastic employment opportunities follow a two-state first-order Markov process. The transition function of the employment opportunities is represented by the matrix $\Pi(s', s) = [\pi(i,j)]$, where each element $\pi(i,j)$ is defined as $\pi(i,j) = \Pr\{s_{t+1} = i | s_t = j\}$, $i, j = \{e, u\}$. Finally, every agent is endowed with exogenous efficiency units normalized to 1.

Technology: The production side is similar to the one of the OLG model, and the equilibrium prices are still given by equations (1) and (2). Notice that all the people with a favorable employment shock are going to be employed, and in the steady-state:

$$L = \frac{\pi(e', u)}{1 - \pi(e', e) + \pi(e', u)}.$$

Government: The government runs an Unemployment Insurance (UI) benefits scheme, by taxing the labor income of the employed workers at rate τ_u and subsidizing the unemployed workers at the replacement rate ϕ_u. ϕ_u is a policy parameter exogenously given, while τ_u is set residually to ensure a self-financing scheme.

4.1 The Monte Carlo Experiment

Also in this case the Monte Carlo experiment shares many features with its OLG counterpart. In the first step, I postulate the stochastic process for the economy and parameterize it to match US data on long-run labor market outcomes. A different aspect of this application is that I will simulate individuals that can only be in two states, employed or unemployed. Although the true stochastic process is a two-state Markov chain, the simulations will deliver cross sectional data, reflecting how National statistical offices actually collect labor market data. This procedure generates a sequence of moments obtained from the simulated samples by simply computing cross sectional averages (i.e., the average unemployment rate and duration). I then re-calibrate the model's parameters according to the simulated moments and I repeat these experiments four times (with samples of size $N = 1,000$, $4,000$, $16,000$, and $64,000$).[20]

[20] Also in this model I performed a set of robustness checks with different fixed parameters, and the main results were not affected. Notice that due to the very onerous computational burden of this model, the calibration in equilibrium of β proved infeasible. Finally, in a previous version of the paper, I solved the model with both a calibration relying on more recent US labor market figures, and I performed the capital tax experiment also in this environment. The results very qualitatively similar, and in the interest of space I do not report them.

4.1.1 Parameterization

DGP: The baseline DGP parameterization targets long-run averages for the US economy, namely an unemployment rate of 6.22% and an average unemployment duration of 18.72 weeks. These are the average US labor market figures for the last 30 years, from July 1983 to June 2013. In order to properly capture the labor market dynamics, I need to work with a short time period: in the benchmark case, one model period corresponds to a quarter. It follows that $\pi(e', e)_{DGP} = 0.9539$, and $\pi(e', u)_{DGP} = 0.6946$.[21]

Monte Carlo Simulations and Calibration: The HA economy under study has seven independent parameters: $\pi(e', u), \pi(e', e), b, \theta, \beta, \alpha$, and δ. These parameters are divided into two categories. The first two parameters, $\pi(e', u)$ and $\pi(e', e)$, are "simulated", meaning that they are going to be recalibrated at each replication of the experiment, on the basis of the specific results obtained in the sample simulation of the DGP. With the information from the simulated samples the parameters I focus on are exactly and uniquely identified and do not need the model solution to be assigned a value. The five remaining parameters are going to be kept fixed.

"Simulated" Parameters: The two independent probabilities representing the Markov chain for the employment opportunities, $\pi(e', e)$ and $\pi(e', u)$, are going to be re-calibrated for every sample drawn in iteration m of the Monte Carlo, according to the following formulas:

$$\pi(e', u)_m = \frac{1}{U\ duration_m} \tag{5}$$

$$\pi(e', e)_m = 1 - \pi(e', u)_m \times \left(\frac{U\ rate_m}{1 - U\ rate_m}\right) \tag{6}$$

where "$U\ duration_m$" stands for the average unemployment duration computed in the simulated sample m, while "$U\ rate_m$" stands for the corresponding simulated unemployment rate.

[Figures 6 and 7 about here]

Figures (6) and (7) show the densities of the two parameters resulting from the calibration procedure. Figure (6) refers to the job finding probability $\pi(e', u)$, computed relying on equation (5), while Figure (7) presents a similar graph for the job retention probability $\pi(e', e)$, computed relying on equation (6). As before, sampling more individuals from the true stochastic process leads to more precise estimates of both the unemployment rate and the unemployment duration. These, in turn, lead to more concentrated distributions of the job finding and job retention probabilities, with the rate of convergence being $\sqrt{N}$.

[21] Although labor market dynamics happen at very high frequencies, considering a quarterly period strikes a balance between the computational burden and the labor market flows. Solving the model with a monthly period more than doubles the computational time, which in this case is already close to ten days. The main drawback of a quarterly model could be to constrain the two Markov chain parameters in some of the calibrations. With this parameterization, this happened only in the smallest sample, just in 1% of the replications.

[Table 7 about here]

Table (7) provides a set of descriptive statistics for the two Markov chain probabilities. Also in this application the parameters' range used to solve the model and compute the related welfare effects is quite large. As for $\pi(e', e)$, with only $1,000$ observations at our disposal this parameter ranges from 0.9047 to 0.9802. This is even more so for the other parameter in the Markov chain, the job finding probability $\pi(e', u)$. In the $1,000$ observations case this parameter ranges from 0.4206 to 1.0. As a consequence, the amount of idiosyncratic uncertainty in the economy varies substantially, getting less disperse with larger samples. With $64,000$ artificial observations $\pi(e', e)$ ranges from 0.9492 to 0.9588, while $\pi(e', u)$ ranges from 0.6548 to 0.7330. Notice how the median values are always very close to the DGP, irrespective of the sample size.

Fixed Parameters: The concavity of the utility function is pinned down by the CRRA coefficient θ, which is set to 1.5. The borrowing limit b is set to zero, as using less strict values had minor implications. The capital depreciation rate is set to replicate an investment/output ratio of approximately 22.5%. This is achieved with $\delta = 0.02$. With a Cobb-Douglas production function the capital share is captured by the parameter $\alpha = 0.34$, which matches the capital share of income. The rate of time preference β is calibrated to get an equilibrium interest rate of approximately 4% on an annual basis, obtained when $\beta = 0.9898$. The complete parameterization of the model is reported in Table (8).

[Table 8 about here]

Policy Parameters: In this model, ϕ_u is the only policy parameter, which represents the institutional features of the UI scheme in the US. It follows that $\phi_u = 0.5$.

4.1.2 Model Solution and Policy Changes

The policy experiment is going to deal with a reduction in the generosity of the UI scheme. This is meant to highlight the insurance properties of this policy and the endogenous response of precautionary savings.[22] It is important to stress that this paper does not deal primarily with the optimality of the UI program, as in Hansen and Imrohoroglu (1992). It is mainly targeted at assessing the empirical properties of the underlying quantitative macroeconomic HA model.[23]

[22] In the experiment, I consider a large change in the UI benefits. An alternative, and perhaps more informative, exercise could have been to consider their elimination. However, in the model I am preventing agents to react to such changes through relevant margins. Agents cannot change their labor supply (taxes on labor income vary as a consequence of the decreased benefits), and they can't put more effort while searching for a job. Although admittedly important, these extensions would increase substantially both the computational burden and the likelihood of the algorithm failing to converge for some calibrations. These complications could make the Monte Carlo experiment quite intractable.

[23] Once more, notice that in the numerical solution of the model there is no sampling variability (due to Monte Carlo integration), but in this model I don't consider a stochastic process for wages, as in Aiyagari (1994). Approximating numerically the stationary distributions with several gridpoints for the income shock proved computationally too expensive. For more details on the numerical solution, see Appendices C and D.

4.1.3 Experiment 2 - UI

For this experiment I assume that the only public policy in place in the economy is the UI. Unexpectedly, the government implements a cut in the replacement rate ϕ_u : $\phi_u = 0.5$ in the benchmark economy and $\phi_u^{new} = 0.25$ in the counterfactual one. Taxes react to the new regime, and their new equilibrium value decreases to $\tau_u^{new} = \frac{\phi_u^{new}(1-L)}{L}$.[24]

The economics of this policy change are very simple. With a lower ϕ_u, agents are less insured against their idiosyncratic shocks. A priori we cannot say if the policy reform is going to lead to a welfare gain or to a welfare loss. Income during an unemployment spell decreases, making it more difficult for agents to achieve consumption smoothing. However, quantitatively, this reform leads to an aggregate welfare increase. This is due to the agents' ability to self-insure by accumulating more assets: capital supply increases, because of the increased precautionary savings. At the same time employed agents pay less taxes, and are going to consume part of these resources. Overall, this leads unemployed agents to be worse off in the new steady-state, while employed ones are better off.

The unemployment rate in the economy is kept constant, hence the capital demand schedule does not change and we are just moving along the curve. A higher capital supply is going to trigger a decrease in the interest rate clearing the asset market, leading to: a) a larger capital stock, b) higher wages and a lower decrease in the UI benefits, c) a lower return on savings. The last effect implies that some agents could start substituting away from saving and consuming more. However, in the aggregate, these shifts are dominated by the increased precautionary saving motive.

Overall, together with the larger relative size of employed agents, this explains why there is a welfare gain when the unemployment benefits are reduced.

The true welfare effect for this case is rather small, and quantified in a 0.1444% increase of the average steady-state utility. Employed agents enjoy a 0.1671% welfare increase, while unemployed ones suffer a 0.2011% welfare loss. The corresponding figures for ϖ, the consumption equivalent welfare measure, are a 0.2881% increase for the overall economy, a 0.3334% increase for employed agents, and a 0.4033% decrease for the unemployed ones.

Welfare Effects Sampling Distributions Figures (8), (9) and (10) plot the non-parametric kernel density estimates of the (consumption equivalent) welfare effects obtained in the four Monte Carlo experiments (with different sample sizes).

[Figures (8), (9) and (10) about here]

Figure (8) shows the economy-wide welfare measures. A few features of the plots are worth stressing. The distributions are almost centered around the true values. Moreover, when moving from the top left panel ($N = 1,000$) to the bottom right one ($N = 64,000$), the dispersion of the welfare effect estimates gets smaller. The more people are sampled, the tighter the calibration, and the less imprecise the welfare effects estimates.

[24] See equation (16) in appendix B.

However, the rate of convergence is not close to the rate of convergence of the parameters, and seems to be happening quite slowly. The endogenous response of the model, when working with parameters different from the true ones, seems to be preventing a rapid convergence of the welfare effects. When the true welfare effects are quantitatively small, as in this case, even small differences in the computed equilibria involve a non-negligible impact on the welfare effects.

The welfare analysis can be broken down for two subgroups, the employed and the unemployed agents. The related plots are shown in Figures (9) and (10). Compared to the economy-wide welfare effects it is possible to appreciate some differences. The welfare effects of the unemployed can be quite imprecise when working with small samples, but they show a rate of convergence which is very fast and closer in magnitude to the parameters' one.[25] As for the employed agents, the same observations seem to hold, but less strongly so. The errors made with very limited sample sizes are less extreme, but they tend to improve at a lower rate.

Also in this alternative set-up, working with a calibrated model with parameters that differ from the true ones leads to errors that are rather small in magnitude. Table (9) reports several statistics of the welfare effects computed in each experiment.

[Table 9 about here]

In this application, the calibrated HA model always got the right sign of the welfare effects. Irrespective of the data abundance, the quantified welfare effects were always of the same sign as the true ones. Once more, the range of the welfare effects decreases monotonically with the sample size. For example, for the economy-wide welfare measure, it moves from $[0.176, 0.365]$ in the least precise case to $[0.239, 0.335]$ in the most precise one. Both the upper bound and the lower bound of the welfare effects support are shrinking considerably with the sample size.

Although the maximum error that one incurs into when evaluating the welfare effects of a policy change decreases with the quantity of the available information, in the UI experiment the maximum error can be sizable. Both for the economy-wide welfare and for the welfare of the employed this never exceeds 0.13 percentage points, and (in absolute value) it ranges between 15.09% and 38.78% of the true value. Differently, the welfare of the unemployed can be quantified in an imprecise way. With the smallest sample, the maximum error reaches 0.65 percentage points, representing a 160.47% mistake relative to the true value. With the largest sample, the severity of the mistakes improve, as now the maximum error is only 0.06 p.p., representing a 16.67% relative error. As for the means and medians of each Monte Carlo experiment, these are again remarkably close to the corresponding true values, providing further evidence that a thorough sensitivity analysis can lead to reliable estimates of an ex-ante policy evaluation.

The main difference between the capital tax experiment in the OLG model and the UI experiment in the infinitely lived agents one is the much slower rate of convergence of the welfare effects distributions towards their true values obtained in the latter model.

[25] This finding provides evidence that the slow convergence of the welfare effects for the employed agents and for the overall economy is a genuine result. The numerical error could have been behind those behaviors.

5 Discussion and Conclusions

More than fifteen years ago, in their survey on computational experiments in macroeconomics, Hansen and Heckman (1996) argued that equilibrium HA models represented a potential solution to some of the problems inherent in the calibration methodology. In this paper I performed a series of Monte Carlo experiments for two classes of HA economies. The general idea was to evaluate some of their properties as tools to perform empirical research. More precisely, I focused on the misspecification of the models due to sampling variability in some statistics that the calibration methodology aims at matching.

This paper does not claim to have provided the final word on the matter. However, as a first assessment, the findings seem to be encouraging: 1) in the simulations, for the vast majority of parameter configurations, calibration procedures, and policy changes, the quantified measure of the welfare change was of the right sign; 2) for several parameter configurations and policy changes the distributions of welfare effects were quite concentrated around the true value, leading to what many would consider tolerable errors when evaluating the effects of a policy change.

On a less bright note, in some experiments the welfare effects were less concentrated, leading to a maximum error as big as 160.47% of the true (albeit rather small) value. In one case, purposedly designed to assess the model's performance when the true welfare effect is modest, the quantified welfare effects were of the wrong sign in a non trivial number of replications. However, mistakes as severe as these seemed to be rare, with the $[2\%, 12\%]$ interval being the most common range, and without a reversal of the sign, even when estimating income processes with only 200 individuals.

An insight provided by these experiments refers to the need for an extensive robustness analysis whenever the welfare effects estimated with an HA model are found to be quantitatively small. In such circumstances, which appear to be the norm rather than a mere intellectual curiosity, sampling variability does bite, and can have serious repercussions. The welfare effects can well be of the wrong sign, when the samples have a limited size (or when the data are very noisy). A natural complementary step consists of performing a more thorough robustness analysis, by assuming empirically motivated prior distributions for the parameters, in the spirit of Canova (1994). With today's computational power, this methodology has become feasible also for rich HA models, as I have shown in Cozzi (2012), a related contribution.

For some experiments, the distribution of welfare effects shows a fast convergence to the true value when increasing the sample size (hence the quality of the calibration). While for other experiments the rate of convergence appeared to be rather slow. This was particularly true for the economy-wide ϖ in the UI experiment.

There are two additional comments related to the interpretation of this Monte Carlo study.

The first comment concerns how to interpret it as a robustness analysis, and if there are simpler alternatives. Rather than considering many replications, one could use the simulation results to obtain ranges for the exogenous parameters, discretize them and perform detailed robustness checks. On the one hand, this approach would miss a key aspect: from the simulations, it is apparent that there are very strong correlations among the calibrated parameters. In particular, the correlation between ρ_y and σ_y is $0.5 < Corr(\rho_y, \sigma_y) < 0.6$, between ρ_y and β is $-0.9 < Corr(\rho_y, \beta) < -0.99$, and that between σ_y and β is $-0.5 < Corr(\sigma_y, \beta) < -0.7$. A simple discretization of the parameters' space would not give a lower importance to the combinations of parameters that are empirically less relevant. On the other hand, a coarse discretization with, say, 10 points per parameter would imply computational costs similar to the ones of the Monte Carlo procedure.

A second, and possibly more interesting, comment is related to how one should interpret the distributions of welfare effects, and how seriously they should be taken in terms of the ex-ante evaluation of the two policy reforms studied here. Following Geweke (2010), this Monte Carlo procedure can be considered as a simple Bayesian prior predictive analysis. More precisely, the income process simulations provide distributions for two of the model's parameters, which can be seen as their priors. From this perspective, the prior distributions are taken to be the sampling distributions of some classical estimators (which admittedly can be objected). Still, a crucial shortcoming of the analysis is that the quantitative predictions of the model are not confronted with the data in a systematic way. Unfortunately, the computation of the likelihood function and the posteriors for this class of models is still a daunting task. Moreover, when DSGE models are considered as incomplete econometric models, formally testing them with likelihood based (or other classical) procedures is not always deemed reasonable. An alternative approach is to interpret the model as a tool delivering population moments, which should then be linked to the data, for instance with a Bayesian reduced-form econometric model with implications for the same moments. The empirical advantages of this research avenue notwithstanding, labeled the Minimal Econometric Interpretation by Geweke (2010), it has not yet been explored in the literature on HA macroeconomic models.

Rather than fully specifying the DGP and creating artificial samples from it, I could have considered real data from, say, the PSID or the CPS, bootstrap the sample, calibrate and solve the model. Although the statistical foundations of bootstrapping are different, I believe that I would have obtained similar results, at least qualitatively. The advantage of the fully specified DGP approach is that it provides a natural benchmark comparison: by design, I know how far each replication is from the true model, and from the true welfare effects. With a bootstrapping experiment I would still be able to compute, for example, bounds for the welfare gains. However, I would not be able to assess how far these computations are from the true value, whenever I rely on parameters that are not the true ones. That is, I would not be able to tell how the endogenous response of the model to parameter misspecifications "confounds" the estimates of the welfare gains/losses.

It would be interesting to evaluate the empirical properties of these classes of HA models both with boostrapping methodologies, and with full-fledged structural estimation/Bayesian ones. I leave these extensions and modifications for future work.

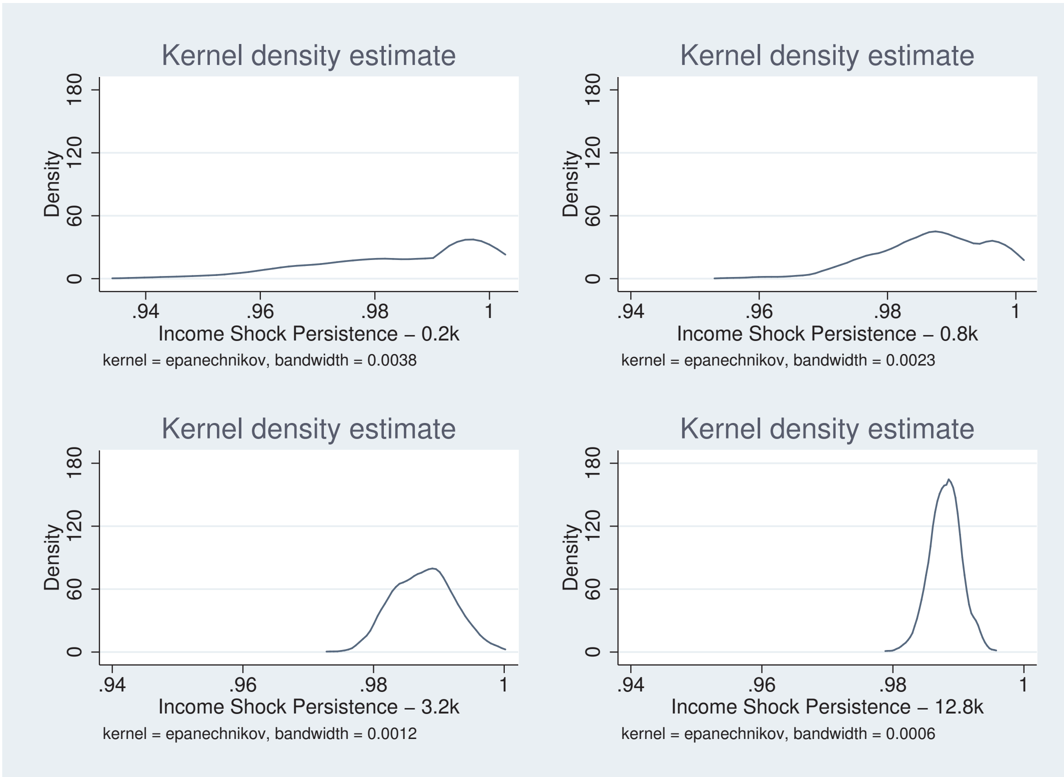

Figure 1: Calibration, OLG Model - "Simulated" Parameters: ρ_y

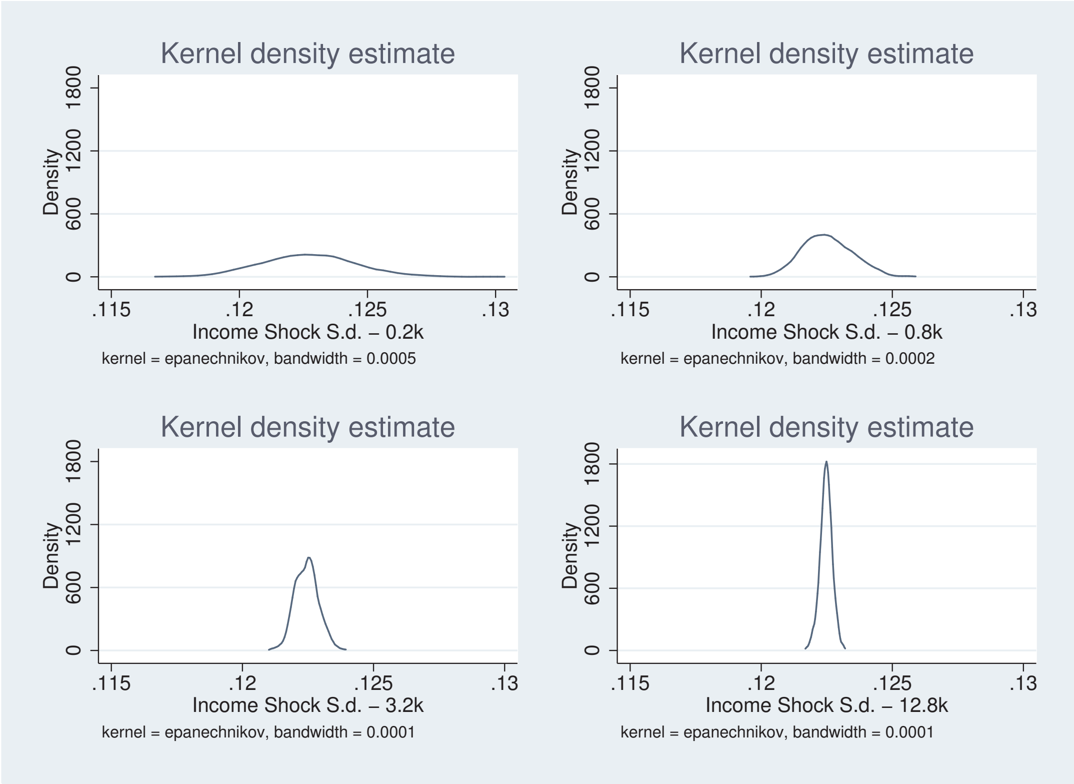

Figure 2: Calibration, OLG Model - “Simulated” Parameters: σ_y

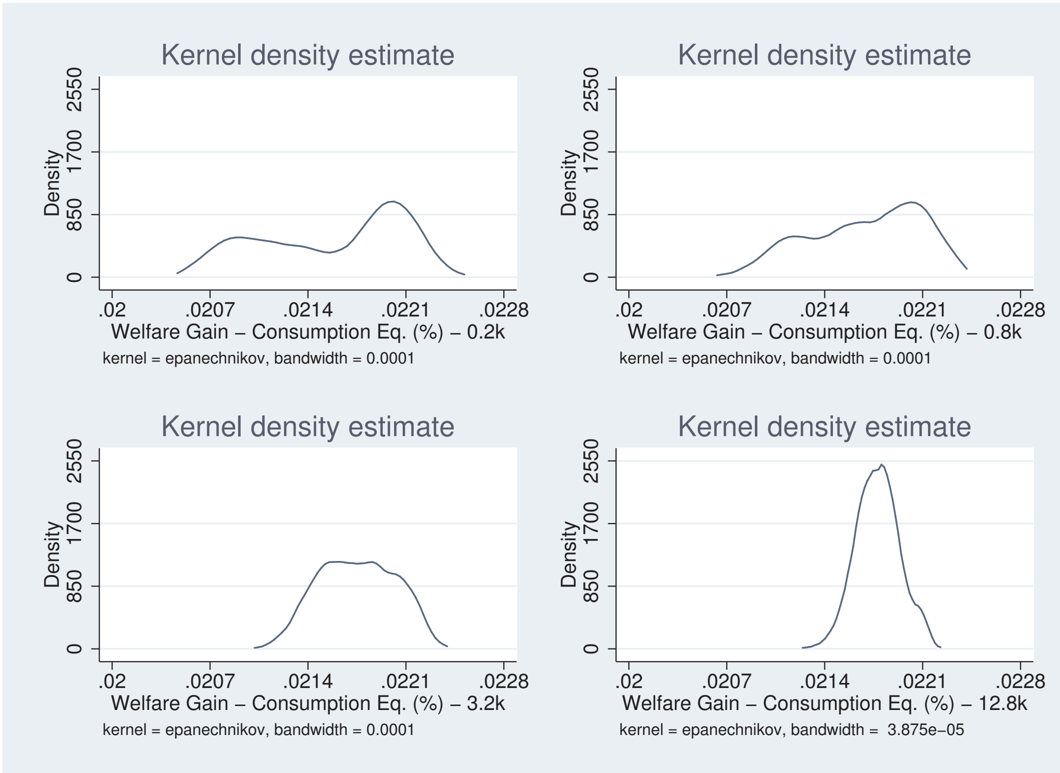

Figure 3: Overall Welfare Gains (ϖ), OLG model, Capital Tax Experiment, Case 1a (Fixed β)

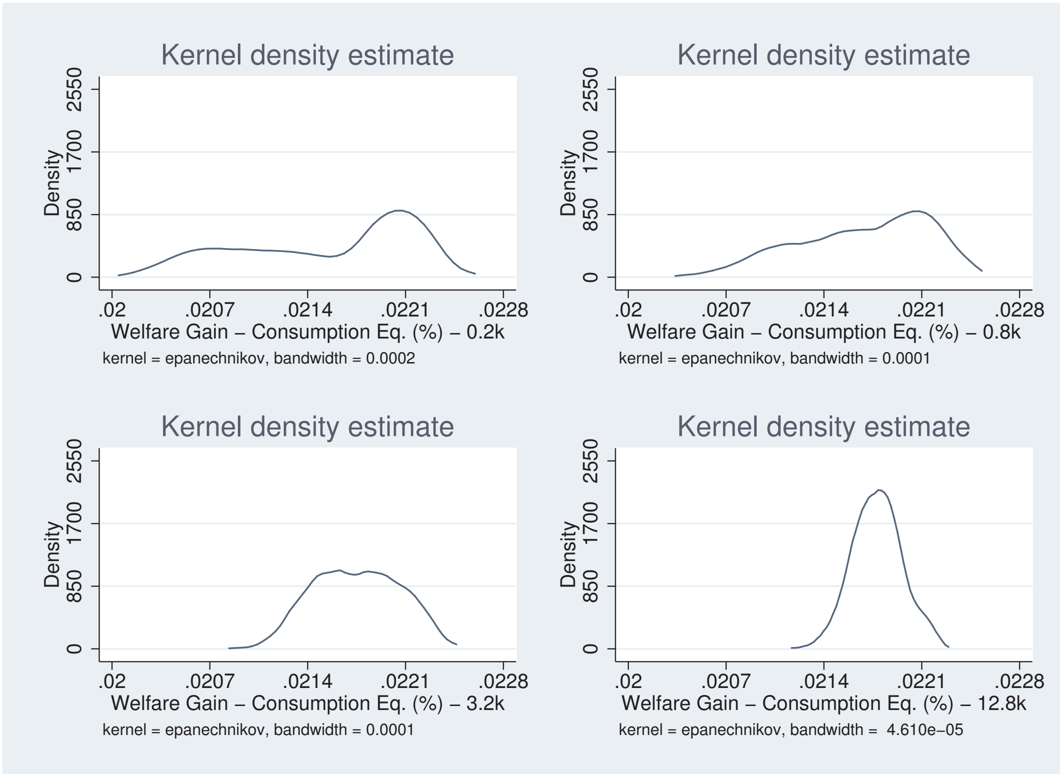

Figure 4: Overall Welfare Gains (ϖ), OLG model, Capital Tax Experiment - Case 1b (Recalibrated β)

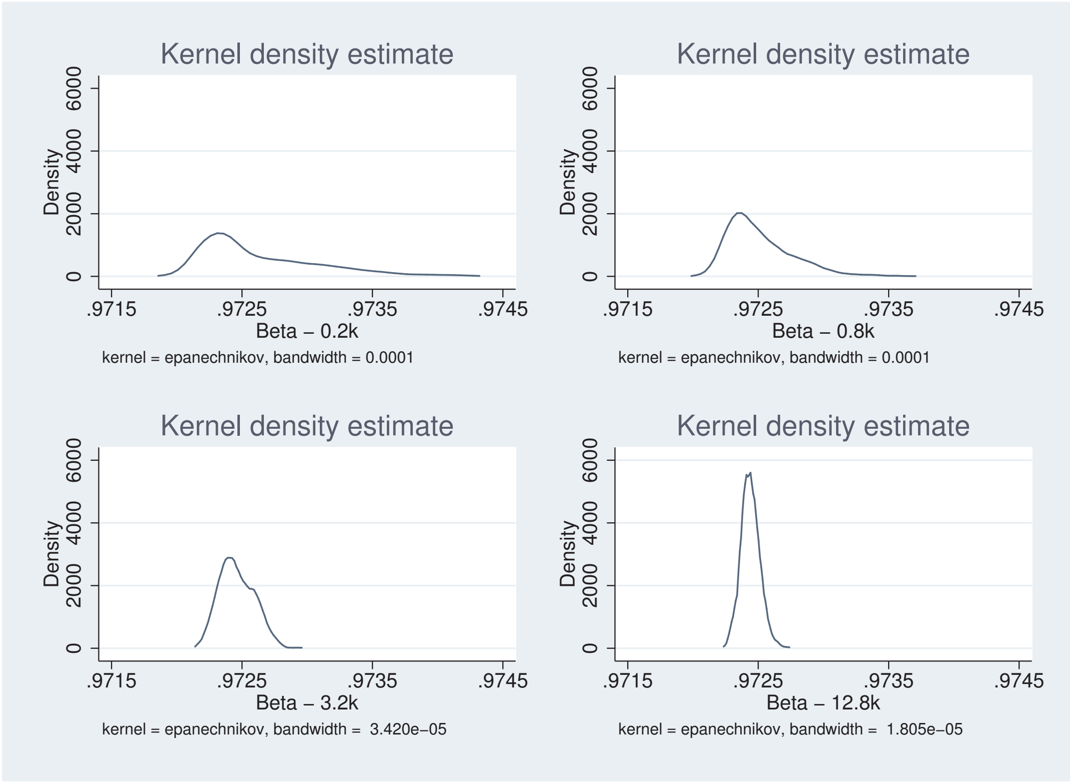

Figure 5: Calibration in Equilibrium, OLG Model - Parameter: β

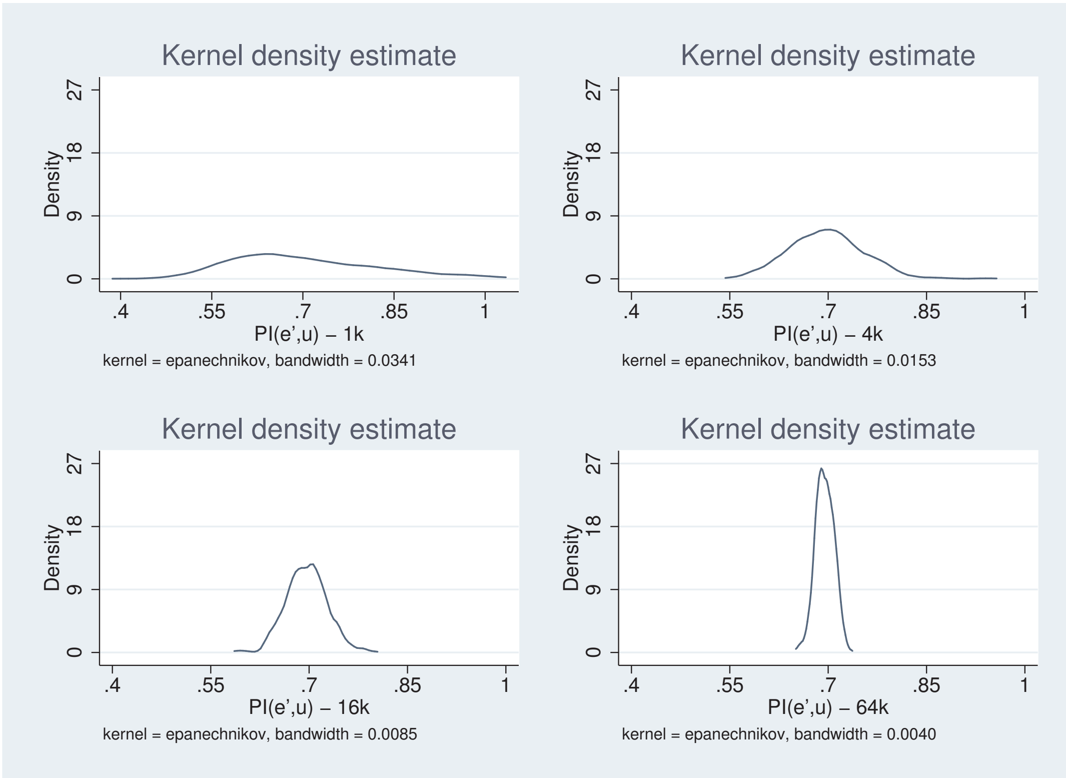

Figure 6: Calibration, ∞-lived Agents Model - "Simulated" Parameters: $\pi(e\prime, u)$

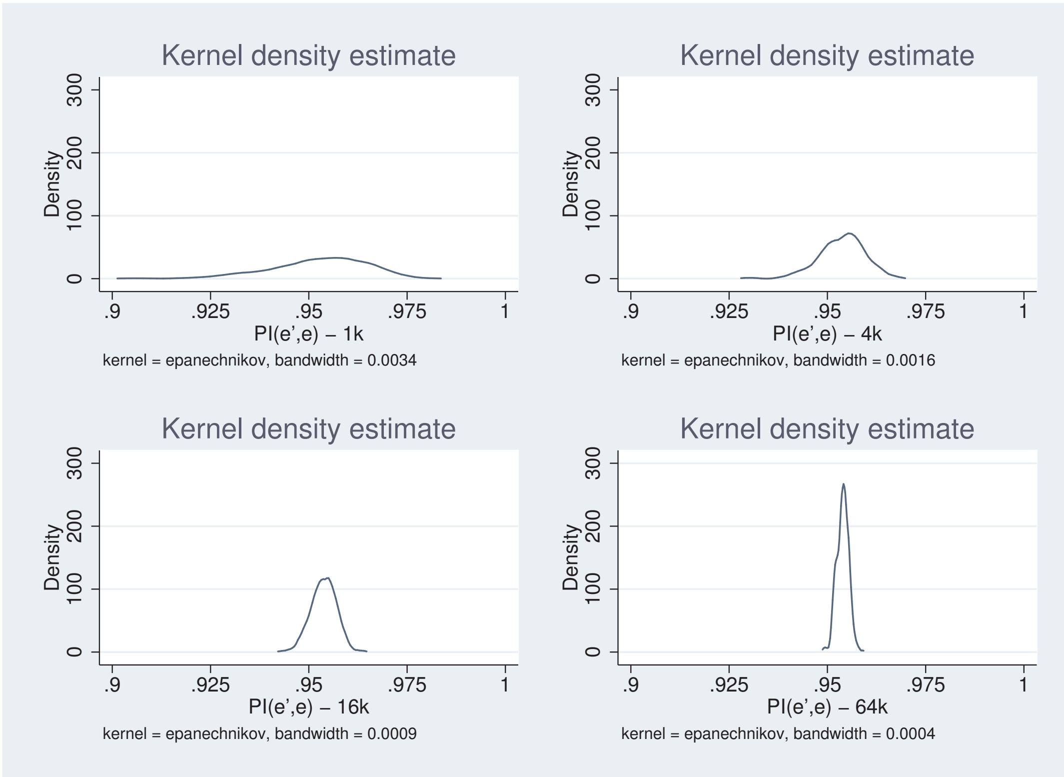

Figure 7: Calibration, ∞-lived Agents Model - "Simulated" Parameters: $\pi(e\prime, e)$

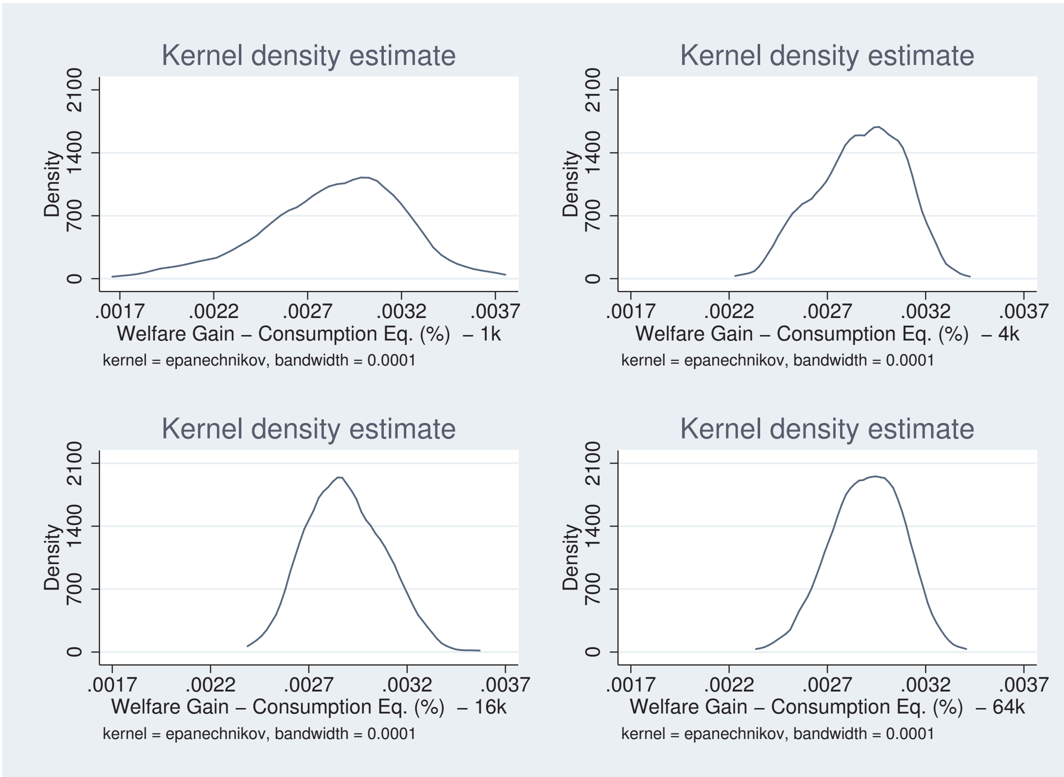

Figure 8: Overall Welfare Gains (ϖ), ∞-lived Agents Model, UI Experiment

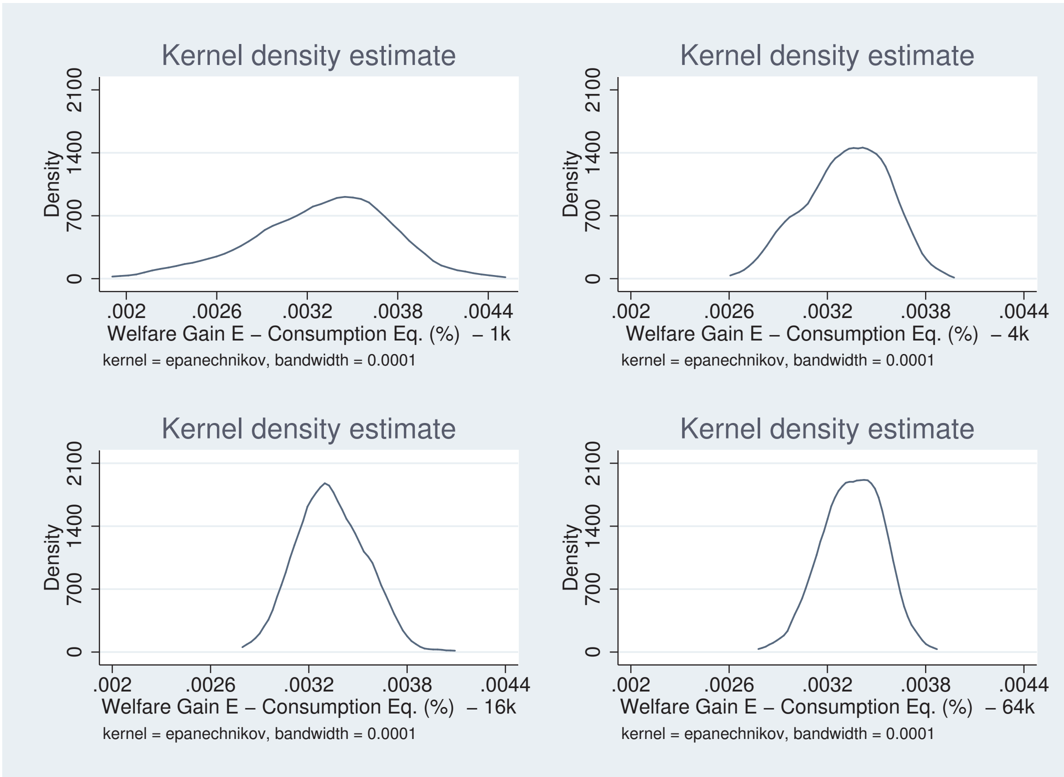

Figure 9: Welfare Gains Employed (ϖ), ∞-lived Agents Model, UI Experiment

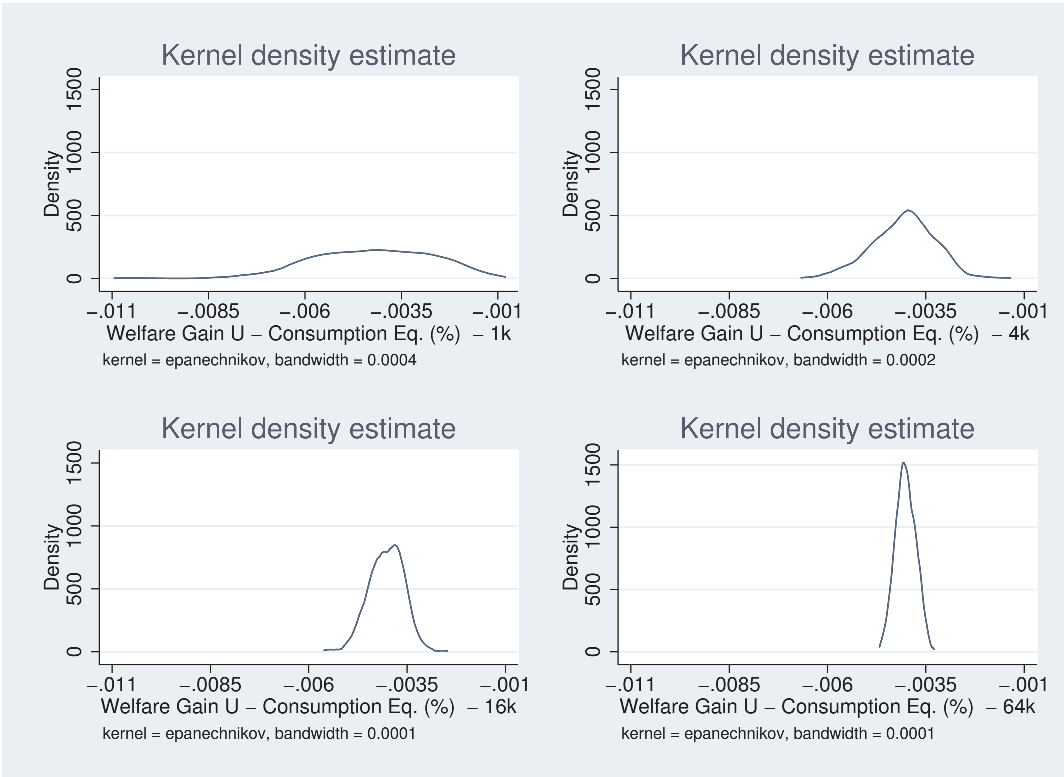

Figure 10: Welfare Gains Unemployed (ϖ), ∞-lived Agents Model, UI Experiment

Sample Size	*Min*	*Max*	*Mean*	*Med*	*C.v.*
ρ_y (*DGP*=0.988)					
0.2k	0.9379	0.999	0.9845	0.9879	0.0148
0.8k	0.9553	0.999	0.9870	0.9875	0.0088
3.2k	0.9740	0.999	0.9876	0.9877	0.0046
12.8k	0.9794	0.9952	0.9880	0.9881	0.0024
σ_y (*DGP*=0.1225)					
0.2k	0.1172	0.1299	0.1227	0.1227	0.0150
0.8k	0.1198	0.1256	0.1225	0.1225	0.0077
3.2k	0.1211	0.1238	0.1224	0.1225	0.0036
12.8k	0.1217	0.1231	0.1225	0.1225	0.0019

Table 1: Calibration, OLG Model, Guvenen (2009) DGP - "Simulated" Parameters: Descriptive Statistics

Parameter	*Value*	*Target*
Model Period	*Year*	*Frequency of PSID Data*
β *- Rate of time preference*	*See Text and Figures*	*Interest rate* $\approx 4\%$
θ *- CRRA*	$\{1.5, 3.0\}$	*Micro Estimates on the Elasticity of Intertemporal Substitution*
δ *- Capital depreciation rate*	$\{0.0698, 0.0457\}$	*Investment share of output* $\approx \{25.0\%, 22.5\%\}$
α *- Capital share*	$\{0.34, 0.40\}$	*Capital share of output* $= \{34\%, 40\%\}$
σ_f^2 *- Var. of the fixed effect*	$\{0.058, 0.211\}$	{*Guvenen* (2009), *Storesletten, Telmer and Yaron* (2004)}

Table 2: Calibration and Robustness, OLG Model - Fixed Parameters

Parameter	*Value*	*Policy*
τ_a	0.30	*Capital income tax*
$(\overline{G/Y})$	0.20	*Government expenditure (Fixed at 20% of Steady State Output)*
$(\overline{TR/Y})$	0.05	*Lump-sum public transfers (Fixed at 5% of Steady State Output)*
ϕ_R	0.39	*Pension replacement rate, from OECD (2011)*

Table 3: Policy Parameters, OLG Model

Parameter	*Value*	β	*True* ϖ (%)
δ - *Capital depreciation rate*	0.0457	0.9792	2.792
α - *Capital share*	0.40	0.9810	2.420
σ_f^2 - *Var. of the fixed effect*	0.211	0.9724	2.064
θ - *CRRA*	3.0	0.9868	0.284

Table 4: Robustness, OLG Model - Alternative Calibrations for the Fixed Parameters

Welfare Change (%)	*Min*	*Max*	*Mean*	*Med*	*C.v.*
Experiment 1a (true $\varpi = 2.175$*)*					
0.2k	2.060	2.239	2.157	2.175	0.0235
0.8k	2.073	2.231	2.169	2.173	0.0182
3.2k	2.109	2.233	2.175	2.176	0.0123
12.8k	2.128	2.219	2.178	2.178	0.0071
Experiment 1b (true $\varpi = 2.175$*)*					
0.2k	2.021	2.244	2.153	2.176	0.0287
0.8k	2.046	2.241	2.168	2.173	0.0213
3.2k	2.092	2.239	2.175	2.175	0.0140
12.8k	2.121	2.225	2.178	2.178	0.0081
Experiment 1c (true $\varpi = 2.112$*)*					
0.2k	2.063	2.240	2.141	2.126	0.0253
0.8k	2.074	2.238	2.135	2.125	0.0186
3.2k	2.092	2.223	2.129	2.127	0.0093
12.8k	2.101	2.167	2.128	2.128	0.0045
Experiment 1d (true $\varpi = 2.112$*)*					
0.2k	2.025	2.245	2.127	2.116	0.0318
0.8k	2.035	2.245	2.122	2.112	0.0229
3.2k	2.062	2.222	2.116	2.114	0.0118
12.8k	2.079	2.160	2.116	2.115	0.0063

Table 5: Monte Carlo Results, OLG model, Capital Tax Experiment - Welfare Effects

Welfare Change (%)	*Min*	*Max*	*Mean*	*Med*	*C.v.*
Robustness - $\delta = 0.0457$ (true $\varpi = 2.792$)					
0.2k	2.556	2.888	2.769	2.791	0.039
0.8k	2.584	2.885	2.783	2.786	0.028
3.2k	2.670	2.884	2.789	2.790	0.017
12.8k	2.706	2.871	2.793	2.794	0.009
Robustness: $\alpha = 0.40$ (true $\varpi = 2.420$)					
0.2k	2.117	2.575	2.405	2.417	0.061
0.8k	2.163	2.574	2.416	2.415	0.042
3.2k	2.272	2.572	2.417	2.416	0.024
12.8k	2.317	2.522	2.420	2.420	0.013
Robustness - $\sigma_f^2 = 0.211$ (true $\varpi = 2.064$)					
0.2k	1.973	2.135	2.061	2.063	0.025
0.8k	1.985	2.132	2.063	2.064	0.018
3.2k	2.012	2.130	2.063	2.065	0.011
12.8k	2.027	2.104	2.065	2.066	0.005
Robustness - $\theta = 3.0$ (true $\varpi = 0.284$)					
0.2k	0.087	0.395	0.245	0.270	0.3159
0.8k	0.119	0.385	0.267	0.277	0.2496
3.2k	0.165	0.388	0.281	0.282	0.1684
12.8k	0.200	0.360	0.286	0.286	0.0972

Table 6: Monte Carlo Results, OLG model, Robustness Analysis - Welfare Effects

Sample Size	*Min*	*Max*	*Mean*	*Med*	*C.v.*
$\pi(e', e)$ (DGP=0.9539)					
1k	0.9047	0.9802	0.9526	0.9540	0.0128
4k	0.9296	0.9682	0.9537	0.9542	0.0061
16k	0.9431	0.9638	0.9538	0.9539	0.0034
64k	0.9492	0.9588	0.9539	0.9540	0.0016
$\pi(e', u)$ (DGP=0.6946)					
1k	0.4206	1	0.7040	0.6851	0.1682
4k	0.5585	0.9416	0.6978	0.6965	0.0820
16k	0.5943	0.7958	0.6963	0.6962	0.0444
64k	0.6548	0.7330	0.6949	0.6945	0.0016

Table 7: Calibration, Infinitely-lived Agents Model - "Simulated" Parameters: Descriptive Statistics

Parameter	*Value*	*Target*
Model Period	*Quarter*	*Labor Market Flows*
θ - *CRRA*	1.5	*Micro Estimates*
b - *Borrowing limit*	0	*No Borrowing*
δ - *Capital depreciation rate*	0.02	*Investment share of output* $\approx 22.5\%$
α - *Capital share*	0.34	*Capital share of output* $= 34\%$
β - *Rate of time preference*	0.9898	*Interest rate* $\approx 4\%$ *(annual)*

Table 8: Calibration, Infinitely-lived Agents Model - Fixed Parameters

Welfare Change (%)	*Min*	*Max*	*Mean*	*Med*	*C.v.*
Experiment 2 - All (true $\varpi = 0.2881$*)*					
1k	0.1764	0.3649	0.2837	0.2877	0.1277
4k	0.2292	0.3364	0.2865	0.2880	0.0750
16k	0.2444	0.3514	0.2884	0.2866	0.0675
64k	0.2388	0.3355	0.2900	0.2907	0.0622
Experiment 2 - Employed (true $\varpi = .3334$*)*					
1k	0.2033	0.4388	0.3295	0.3365	0.1361
4k	0.2679	0.3902	0.3317	0.3330	0.0751
16k	0.2854	0.4033	0.3336	0.3318	0.0618
64k	0.2831	0.3816	0.3353	0.3357	0.0545
Experiment 2 - Unemployed (true $\varpi = -.4033$*)*					
1k	−1.0505	−0.1243	−0.4234	−0.4175	0.3590
4k	−0.6454	−0.1565	−0.4071	−0.4020	0.1928
16k	−0.5485	−0.2598	−0.4024	−0.3989	0.1085
64k	−0.4609	−0.3361	−0.4013	−0.4026	0.0621

Table 9: Monte Carlo Results, Infinitely-lived Agents Model, UI Experiment - Welfare Effects

References

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

Aiyagari, R., and McGrattan, E. (1998). "The optimum quantity of debt," *Journal of Monetary Economics,* Vol. 42, 447-469.

Alvarez, F., and Veracierto, M. (2001). "Severance Payments in an Economy with Frictions," *Journal of Monetary Economics*, Vol. 47, 477-498.

Bell, F., and Miller, M. (2002). "Life Tables for the United States Social Security Area 1900–2100." Office of the Chief Actuary, Social Security Administration, Actuarial Study 116.

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

Canova, F. (1995). "Sensitivity Analysis and Model Evaluation in Simulated Dynamic General Equilibrium Economies," *International Economic Review*, Vol. 36 , 477-501.

Carroll, C. (2000). "Requiem for the Representative Consumer? Aggregate implications of Microeconomic Consumption Behavior," *American Economic Review (P&P)*, Vol. 90, 110-115.

Castaneda, A., Diaz-Gimenez, J., and Rios-Rull, J.V. (2003). "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, Vol. 111, 818-857.

Conesa, J., Kitao, S. and Krueger, D. (2009). "Taxing Capital? Not a Bad Idea after All!," *American Economic Review*, Vol. 99, 25-48.

Cozzi, M. (2012). "Optimal Unemployment Insurance in GE: a Robust Calibration Approach," *Economics Letters*, Vol. 117, 28-31.

French, E. (2005). "The Effects of Health, Wealth and Wages on Labor Supply and Retirement Behavior," *Review of Economic Studies*, Vol. 72, 395-427.

Fuster, L., Imrohoroglu, A., and Imrohoroglu, S. (2007). "Elimination of Social Security in a Dynastic Framework," *Review of Economic Studies*, Vol. 74, 113-145.

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

Gregory, A., and Smith, G. (1990). "Calibration as Estimation," *Econometric Reviews,* Vol. 9, 57-89.

Gregory, A., and Smith, G. (1991). "Calibration as Testing: Inference in Simulated Macroeconomic Models," *Journal of Business and Economic Statistics*, Vol. 9, 297-303.

Guvenen, M.F. (2009). "An Empirical Investigation of Labor Income Processes," *Review of Economic Dynamics*, Vol. 12, 58-79.

Guvenen, F. (2007). "Learning Your Earning: Are Labor Income Shocks Really Very Persistent?," *American Economic Review*, Vol. 97, 687-712.

Hansen, G. (1993). "The Cyclical and Secular Behaviour of the Labour Input: Comparing Efficiency Units and Hours Worked." *Journal of Applied Econometrics*, Vol. 8, 71-80.

Hansen, L., and Heckman, J. (1996). "The Empirical Foundations of Calibration," *Journal of Economic Perspectives*, Vol. 10, 87-104.

Hansen, G., and Imrohoroglu, A. (1992). "The Role of Unemployment Insurance in an Economy with Liquidity Constraints and Moral Hazard," *Journal of Political Economy*, Vol. 100, 118-142.

Heathcote, J., Perri, F., and Violante, G. (2010). "Unequal we stand: An empirical analysis of economic inequality in the United States, 1967–2006," *Review of Economic Dynamics*, Vol. 13, 15-51.

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

Heathcote, J, Storesletten, K., and Violante, G. (2010). "The Macroeconomic Implications of Rising Wage Inequality in the United States," *Journal of Political Economy*, Vol. 118, 681-722.

Heckman, J., Lochner, L., and Taber, C. (1998). "Explaining Rising Wage Inequality: Explorations with a Dynamic General Equilibrium Model of Labor Earnings with Heterogeneous Agents," *Review of Economic Dynamics*, Vol. 1, 1-58.

Hoover, K. (1995). "Facts and Artifacts: Calibration and the Empirical Assessment of Real-Business-Cycle Models," *Oxford Economic Papers*, Vol. 47, 24-44.

Huggett, M. (1993). "The Risk-free Rate in Heterogeneous-agent Incomplete-insurance Economies," *Journal of Economic Dynamics and Control,* Vol. 17, 953-969.

Huggett, M. (1996). "Wealth distribution in life-cycle economies," *Journal of Monetary Economics*, Vol. 38, 469-494.

Imrohoroglu, A. (1989). "Cost of Business Cycles with Indivisibilities and Liquidity Constraints," *Journal of Political Economy*, Vol.97, 1364-1383.

Imrohoroglu, S. (1998). "A Quantitative Analysis of Capital Income Taxation," *International Economic Review*, Vol. 39, 307-328.

Kopecky, K., and Suen, R. (2010). "Finite state Markov-chain approximations to highly persistent processes," *Review of Economic Dynamics*, Vol. 13, 701-714.

Krusell, P., Mukoyama, T., Rogerson, R., and Sahin, A. (2010). "Aggregate Labor Market Outcomes: the Roles of Choice and Chance," *Quantitative Economics,* Vol. 1, 97–127.

Krusell, P., and Smith, A. (2006). "Quantitative Macroeconomic Models with Heterogeneous Agents," in Blundell, R., Newey, W., and Persson, T. (eds). Advances in Economics and Econometrics: Theory and Applications, Econometric Society Monographs, 41, Cambridge University Press, 298-340.

Michaelides, A., and Ng, S. (2000). "Estimating the Rational Expectations Model of Speculative Storage: a Monte Carlo Comparison of Three Simulation Estimators," *Journal of Econometrics*, Vol. 96, 231-266.

OECD. (2011). *Pensions at a Glance 2011: Retirement-income Systems in OECD and G20 Countries*, OECD Publishing, Paris.

Rios-Rull, J.V. (1999). "Computation of Equilibria in Heterogenous Agent Models," in *Computational Methods for the Study of Dynamic Economies: An Introduction*, (Ramon Marimon and Andrew Scott, Eds.), Oxford University Press.

Storesletten, K., Telmer, C. and Yaron, A. (2004). "Consumption and Risk Sharing over the Life Cycle," *Journal of Monetary Economics*, Vol. 51, pp. 609-633.

Ventura, G. (1999). "Flat Tax Reform: a Quantitative Exploration," *Journal of Economic Dynamics and Control*, Vol. 23, 1425-1458.

Watson, M. (1993). "Measures of Fit for Calibrated Models," *Journal of Political Economy*, Vol. 101, 1011-1041.

Appendix A - The OLG Model and its Recursive Representation

6 Stationary Equilibrium

In this Section, first the problem of the agents in their recursive representation is defined, then I provide a formal definition of the equilibrium concept used in this model, the recursive competitive equilibrium. The individual state variables are: age $t \in \mathcal{T} = \{1, ..., T\}$, the fixed effect $f \in \mathcal{F} = \{-\sigma_f, +\sigma_f\}$, the persistent labor endowment shock $\varepsilon \in \mathcal{E} = \{\varepsilon_{\min}, ..., \overline{\varepsilon}, ..., \varepsilon_{\max}\}$ and asset holdings $a \in \mathcal{A} = [-b, \overline{a}]$. Notice that ε is discretized with the Rouwenhorst method, using an 11-state Markov chain. The transition function of the labor endowment shocks is represented by the matrix $\Pi(\varepsilon', \varepsilon) = [\pi(i,j)]$, where each element $\pi(i,j)$ is defined as $\pi(i,j) = \Pr\{\varepsilon_{t+1} = i | \varepsilon_t = j\}$, $i, j \in \mathcal{E}$. In every period the total efficiency units are given by $\epsilon_{t,\varepsilon,f} = e_t \varepsilon f$. The stationary distribution of working-age agents is denoted by $\mu_t(a, \varepsilon, f)$ while that of retirees with $\mu_t^R(a)$. Φ_t denotes the share of each cohort t in the total population. These satisfy the recursion $\Phi_{t+1} = \left(\frac{1-\pi_j^d}{1+g_n}\right)\Phi_t$, and are normalized to add up to 1.

6.1 Problem of the agents

The model is solved backwards, starting from the terminal age T and with the assumption that the terminal utility value is zero, i.e. $V_{T+1} = 0$.

6.1.1 Problem of the retirees

The value function of an age-t retired agent whose current asset holdings are equal to a is denoted with $V_t^R(a)$. The problem of these agents can be represented as follows:

$$V_t^R(a) = \max_{c,a'} \left\{ u(c) + \beta\left(1 - \pi_j^d\right) V_{t+1}^R(a') \right\} \tag{7}$$

$$\begin{aligned} & s.t. \\ & c + a' = \left(\frac{1 + (1-\tau_a) r}{1 - \pi_j^d}\right) a + (1-\tau_w)\overline{y}_R + tr \\ & c \geq 0, \quad a' > 0 \end{aligned}$$

In the budget constraint notice the presence of the pension payment $\overline{y}_p$, and that the interest rate is adjusted by both the capital tax and by $\left(1 - \pi_j^d\right)$, because of the assumption on the availability of annuities. Retired agents also pay proportional pension income taxes at rate τ_w and receive a lump-sum transfer tr.

6.1.2 Problem of the workers

The value function of a working-age agent whose current asset holdings are equal to a, whose current efficiency units shock is ε and whose fixed effect is f is denoted with $V_t(a, \varepsilon, f)$. The problem of these agents can be represented as follows:

$$V_t(a,\varepsilon,f) = \max_{c,a'} \left\{ u(c) + \beta\left(1-\pi_j^d\right)\pi(\varepsilon',\varepsilon)\, V_{t+1}(a',\varepsilon',f) \right\} \tag{8}$$

$$\begin{aligned}
&s.t.\\
&c + a' = \left(\frac{1+(1-\tau_a)\,r}{1-\pi_j^d}\right) a + (1-\tau_w-\tau_R)\, w\epsilon_{t,\varepsilon,f} + tr\\
&a_0 = 0, \quad c \geq 0, \quad a' > -b
\end{aligned}$$

Non-retired agents have to set optimally their consumption/savings plans. They enjoy utility from consumption, and face some uncertain events in the future. In the next period they can still be alive, and with probability $\pi(\varepsilon',\varepsilon)$ they transit from their current efficiency units ε to the value ε'. These agents pay proportional labor and capital income taxes, τ_a and τ_w, respectively. They also pay a proportional tax τ_R on their labor earnings to finance the pension scheme. Finally, they receive a lump-sum transfer tr, they are born with no wealth and with the average shock $\overline{\varepsilon}$, and they are subject to an exogenous borrowing constraint, $b \geq 0$.

6.2 Recursive Stationary Equilibrium

Definition 1 *For given policies $\{\tau_a, G, TR, \phi_R\}$ a recursive stationary equilibrium is a set of decision rules, $\{c_t(a,\varepsilon,f), a'_t(a,\varepsilon,f)\}_{t=1}^{T_R-1}$ and $\left\{c_t^R(a), a_t^{R\prime}(a)\right\}_{t=T_R}^{T}$, value functions, $\{V_t(a,\varepsilon,f)\}_{t=1}^{T_R-1}$ and $\left\{V_t^R(a)\right\}_{t=T_R}^{T}$, prices $\{r,w\}$, proportional taxes $\{\tau_R,\tau_w\}$ and a set of stationary distributions, $\{\mu_t(a,\varepsilon,f)\}_{t=1}^{T_R-1}$ and $\left\{\mu_t^R(a)\right\}_{t=T_R}^{T}$, such that:*

- Given relative prices $\{r,w\}$, proportional taxes $\{\tau_a,\tau_R,\tau_w\}$ and pension benefits $\overline{y}_R$, the individual policy functions $\{c_t(a,\varepsilon,f), a'_t(a,\varepsilon,f)\}_{t=1}^{T_R-1}$, $\left\{c_t^R(a), a_t^{R\prime}(a)\right\}_{t=T_R}^{T}$ solve the household problems (7)-(8), and $\{V_t(a,\varepsilon,f)\}_{t=1}^{T_R-1}$, $\left\{V_t^R(a)\right\}_{t=T_R}^{T}$ are the associated value functions.

- Given relative prices $\{r,w\}$, K/L solves the firm's problem and satisfies (1)-(2).

- The labor market is in equilibrium, and the labor input L corresponds to the total supply of labor efficiency units

$$L = \sum_{t=1}^{T_R-1} \Phi_t \int_{\mathcal{A}\times\mathcal{E}\times\mathcal{F}} \epsilon_{t,\varepsilon,f}\, d\mu_t(a,\varepsilon,f)$$

- The asset market clears

$$(1+n)\,K = \sum_{t=1}^{T_R-1} \Phi_t \int_{\mathcal{A}\times\mathcal{E}\times\mathcal{F}} a'_t(a,\varepsilon,f)\, d\mu_t(a,\varepsilon,f) + \sum_{t=T_R}^{T} \Phi_t \int_{\mathcal{A}} a_t^{R\prime}(a)\, d\mu_t^R(a)$$

- The goods market clears

$$F(K,L) = C + I + G = \sum_{t=1}^{T_R-1} \Phi_t \int_{\mathcal{A}\times\mathcal{E}\times\mathcal{F}} c_t(a,\varepsilon,f)\, d\mu_t(a,\varepsilon,f) + \sum_{t=T_R}^{T} \Phi_t \int_{\mathcal{A}} c_t^R(a)\, d\mu_t^R(a) + (\delta + g_n) K + G$$

- The government's budget is balanced, that is tax revenues from (capital and labor) income taxation are equal to the sum of government purchases G and total transfers TR

$$G + TR = \sum_{t=1}^{T_R-1} \Phi_t \int_{\mathcal{A}\times\mathcal{E}\times\mathcal{F}} \tau_a r a d\mu_t(a,\varepsilon,f) + \sum_{t=T_R}^{T} \Phi_t \int_{\mathcal{A}} \tau_a r a d\mu_t^R(a) + \sum_{t=1}^{T_R-1} \Phi_t \int_{\mathcal{A}\times\mathcal{E}\times\mathcal{F}} \tau_w w \epsilon_{t,\varepsilon,f} d\mu_t(a,\varepsilon,f) + \sum_{t=T_R}^{T} \Phi_t \int_{\mathcal{A}} \tau_w \overline{y}_R d\mu_t^R(a)$$

- The pension is equal to

$$\overline{y}_R = \sum_{t=1}^{T_R-1} \Phi_t \int_{\mathcal{A}\times\mathcal{E}\times\mathcal{F}} \phi_R w \epsilon_{t,\varepsilon,f} d\mu_t(a,\varepsilon,f) = \phi_R w \overline{\epsilon}$$

notice that the pension benefits are the replacement rate ϕ_R multiplied by the average labor earnings $w\overline{\epsilon}$.

- The proportional pension tax satisfies

$$\tau_R = \frac{\sum_{t=T_R}^{T} \Phi_t \int_{\mathcal{A}} \overline{y}_R d\mu_t^R(a)}{\sum_{t=1}^{T_R-1} \Phi_t \int_{\mathcal{A}\times\mathcal{E}\times\mathcal{F}} w \epsilon_{t,\varepsilon,f} d\mu_t(a,\varepsilon,f)} = \phi_R \overline{\mu}_R$$

where $\overline{\mu}_R$ is the mass of retirees.

- It can be shown that, given G, TR and τ_a, the proportional tax τ_w satisfies

$$\tau_w = \left(\frac{1}{1-\alpha}\right)\left(\frac{1}{1+\phi_p \overline{\mu}_p}\right)\left[\frac{G}{Y} + \frac{TR}{Y} - \tau_a \left(\frac{\alpha r}{r+\delta}\right)\right] \tag{9}$$

$$\text{with } \frac{G}{Y} = \overline{\left(\frac{G}{Y}\right)} = \text{constant, and } \frac{TR}{Y} = \overline{\left(\frac{TR}{Y}\right)} = \text{constant.}$$

- The stationary distributions $\{\mu_t(a,\varepsilon,f), \mu_t^R(a)\}$ satisfy

$$\mu_{t+1}(a',\varepsilon',f) = \int \psi(a,\varepsilon,f,t,a',\varepsilon')\, d\mu_t(a,\varepsilon,f) \tag{10}$$

$$\mu_{t+1}^R(a') = \int \psi^R(a,t,a')\, d\mu_t^R(a) \tag{11}$$

In equilibrium the measure of agents in each state is time invariant and consistent with individual decisions, as given by the above two equations (10)-(11), where $\psi(.)$ and $\psi^R(.)$ are the transition functions.

- The welfare measure W_{OLG} is utilitarian, i.e. it weights agents' utilities by their mass in the steady-state

$$W_{OLG} = \sum_{t=1}^{T_R-1} \Phi_t \int_{\mathcal{A}\times\mathcal{E}\times\mathcal{F}} V_t(a,\varepsilon,f)\, d\mu_t(a,\varepsilon,f) + \sum_{t=T_R}^{T} \Phi_t \int_{\mathcal{A}} V_t^R(a)\, d\mu_t^R(a) \tag{12}$$

- The consumption based welfare measure $\widehat{\varpi}$ is the percentage increase in consumption in all states of the world that makes welfare in the counterfactual economy $W_{OLG}^1(\varpi)$ equal to welfare in the baseline one W_{OLG}^0

$$\begin{aligned} W_{OLG}^0 &= W_{OLG}^1(\varpi) \\ \widehat{\varpi} &= \left(\frac{W_{OLG}^1}{W_{OLG}^0}\right)^{\frac{1}{1-\theta}} - 1 \end{aligned} \tag{13}$$

Appendix B - The Infinitely Lived Model and its Recursive Representation

7 Stationary Equilibrium

First the problem of employed and unemployed workers is defined. The individual state variables are the employment status $s \in \mathcal{S} = \{e, u\}$, and asset holdings $a \in \mathcal{A} = [-b, \overline{a}]$.[26] The stationary distribution of employed agents is denoted by $\mu_e(a)$ whereas the distribution of unemployed agents is $\mu_u(a)$.

7.1 Problem of the agents

7.1.1 Problem of the unemployed workers

The value function of an unemployed agent whose current asset holdings are equal to a is denoted with $V_u(a)$. The problem of these agents can be represented as follows:

$$V_u(a) = \max_{c,a'} \{u(c) + \beta [\pi(u', u) V_u(a') + (1 - \pi(u', u)) V_e(a')]\} \tag{14}$$

$$s.t.$$

$$c + a' = (1 + r) a + \phi_u w$$

$$a_0 \; given, \quad c \geq 0, \quad a' > -b$$

Unemployed agents have to set optimally their consumption/savings plans. They enjoy utility from consumption, and face some uncertain events in the future. In the next period they can still be unemployed, with probability $\pi(u', u)$, or they can find a job and enjoy an employment spell, with probability $1 - \pi(u', u)$.

While unemployed these workers receive an unemployment benefit equal to $\phi_u w$. The unemployment benefit consists of the replacement rate ϕ_u (a policy parameter) of the wage w an employed worker is receiving. Finally, they are subject to an exogenous borrowing constraint, $b \geq 0$.

7.1.2 Problem of the employed workers

The recursive representation of the problem of an employed worker is as follows:

$$V_e(a) = \max_{c,a'} \{u(c) + \beta [\pi(u', e) V_u(a') + (1 - \pi(u', e)) V_e(a')]\} \tag{15}$$

$$s.t.$$

$$c + a' = (1 + r) a + (1 - \tau_u) w$$

$$a_0 \; given, \quad c \geq 0, \quad a' > -b$$

[26] A formal argument proving that $\overline{a} < \infty$ appears for a similar economy in Huggett (1993).

Employed agents enjoy utility from consumption and face some uncertain events in the future. In the next period they can still be employed, or they can experience a job separation and begin an unemployment spell. Finally, notice that these agents pay proportional labor income taxes at rate τ_u to finance the unemployment benefit scheme.

7.2 Recursive Stationary Equilibrium

Definition 2 *For a given policy ϕ_u a recursive stationary equilibrium is a set of decision rules $\{c_e(a), c_u(a), a_e'(a), a_u'(a)\}$, value functions $\{V_e(a), V_u(a)\}$, prices $\{r, w\}$, proportional taxes τ_u and a set of stationary distributions $\{\mu_e(a), \mu_u(a)\}$ such that:*

- Given relative prices $\{r, w\}$, proportional taxes τ_u and unemployment benefits $\phi_u w$, the individual policy functions $\{c_e(a), c_u(a), a_e'(a), a_u'(a)\}$ solve the household problems (14)-(15), and $\{V_e(a), V_u(a)\}$ are the associated value functions.

- Given relative prices $\{r, w\}$, K/L solves the firm's problem and satisfies (1)-(2).

- The labor market is in flow equilibrium, that is the measure of people becoming unemployed is identical to the measure of people finding a job

$$\int_{\mathcal{A}} \pi(u', e)\, d\mu_e(a) = \pi(u', e) L = \pi(e', u)(1 - L) = \int_{\mathcal{A}} \pi(e', u)\, d\mu_u(a)$$

- The asset market clears

$$K = \int_{\mathcal{A} \times \mathcal{S}} a_s'(a) d\mu_s(a)$$

- The goods market clears

$$F(K, L) = C + I = \int_{\mathcal{A} \times \mathcal{S}} c_s(a) d\mu_s(a) + \delta K$$

- The unemployment benefit scheme is self-financing, i.e. the proportional tax τ_u satisfies

$$\tau_u = \frac{\int_{\mathcal{A}} \phi_u w d\mu_u(a)}{\int_{\mathcal{A}} w d\mu_e(a)} = \frac{\phi_u (1 - L)}{L} \tag{16}$$

- The stationary distributions $\{\mu_e(a), \mu_u(a)\}$ satisfy

$$\mu_e(a') = \int_{a: a_u'(a) = a'} \pi(e', u)\, d\mu_u(a) + \int_{a: a_e'(a) = a'} \pi(e', e)\, d\mu_e(a) \tag{17}$$

$$\mu_u(a') = \int_{a: a_u'(a) = a'} \pi(u', u)\, d\mu_u(a) + \int_{a: a_e'(a) = a'} \pi(u', e)\, d\mu_e(a) \tag{18}$$

In equilibrium the measure of agents in each state is time invariant and consistent with individual decisions, as given by the above two equations (17)-(18).

- The welfare measures W, W_e and W_u are utilitarian, i.e. they weight agents' utilities by their mass in the steady-state

$$W = \int_{\mathcal{A}\times\mathcal{S}} V_s(a)d\mu_s(a),\ W_e = \int_{\mathcal{A}} V_e(a)d\mu_e(a),\ W_u = \int_{\mathcal{A}} V_u(a)d\mu_u(a) \tag{19}$$

- The consumption based welfare measure $\widehat{\varpi}$ is the percentage increase in consumption in all states of the world that makes welfare in the counterfactual economy $W^1(\varpi)$ equal to welfare in the baseline one W^0

$$\begin{aligned} W^0 &= W^1(\varpi) \\ \int_{\mathcal{A}\times\mathcal{S}} E_0 \sum_{t=0}^{\infty} \beta^t \frac{\left[c^0(.)_t\right]^{1-\theta}-1}{1-\theta} d\mu_s^0(a) &= \int_{\mathcal{A}\times\mathcal{S}} E_0 \sum_{t=0}^{\infty} \beta^t \frac{\left[(1+\varpi)\,c^1(.)_t\right]^{1-\theta}-1}{1-\theta} d\mu_s^1(a) \\ \widehat{\varpi} &= \left(\frac{W^1}{W^0}\right)^{\frac{1}{1-\theta}} - 1 \end{aligned} \tag{20}$$

Appendix C - Computation

- All codes solving the model economies and simulating samples of agents were written in the FORTRAN 95 language, relying on the Intel Fortran Compiler, build 11.1.048 (with the IMSL library). They were compiled selecting the O3 option (maximize speed), and without automatic parallelization. They were run on different 64-bit PC platforms, all running Windows 7 Professional Edition, either with an Intel $i7-870$ Quad Core 2.93 Ghz processor, or with an Intel $i7-2600k$ Quad Core processor clocked at 4.6 Ghz.

- For the OLG model, the 500 Monte Carlo replications take up to 50 hours to complete. Notice that $1,000$ equilibria have to be computed, and typically from 10 to 14 iterations on the interest rate are needed to find each equilibrium.

- For the Infinitely-lived agents model, the 500 Monte Carlo replications take up to 250 hours to complete. The high discount factor together with the extremely large asset space (calling for a large number of gridpoints) explain the severe computational costs.

- The sample simulations are performed outside the numerical solution of the model. For the OLG model, I simulate for 500 times a sample of individuals from the true stochastic process. These simulations take a few minutes to complete. For the Infinitely lived agents model, I simulate for 500 times and for $3,000$ periods (to ensure stationarity) a sample of individuals from the true stochastic process. These simulations take from 5 minutes (with $1,000$ individuals) to 5 hours (with $64,000$ ones). For a given sample size, in order to avoid sampling variability affecting the results, I rely on the same sample realizations to parameterize and solve all models.

- In the actual solution of the models I need to discretize the continuous state variable 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 the OLG model, I use 301 points, as increasing the number of points does not affect the results considerably. Differently, the Infinitely-lived agents model is more sensitive with respect to this choice. In order to allow for very good approximations of both the policy and value functions, I use $1,001$ grid points on the asset space, the lowest value being the borrowing constraint b and the highest one being a value $a_{\max} > \overline{a}$ high enough for the saving functions to cut the 45 degree line ($a_{\max} = 6,000$ for the quarterly model).

- In the OLG model, ε is discretized with the Rouwenhorst method, using an 11-state Markov chain. This method has several desirable properties, especially when working with highly persistent processes, as discussed by Kopecky and Suen (2010). Notice that in the Infinitely-lived agents model the employment status is already discrete.

- The OLG model is solved with a backward recursion on the Bellman equations. I start from the terminal value $V_{T+1}^{R} = 0$, and at each age, for every point in the state space, I solve the constrained maximization problem. I retrieve the policy functions, $a_t'(a,\varepsilon,f)$ and $a_t^{R\prime}(a)$, and I compute the vector of parameters Ω representing the Schumaker spline approximations of the value functions. Notice that I do not restrict the agents' asset holdings to belong to a discrete set. As for the approximation method, I rely on the quadratic spline approximations for the future value functions, when evaluated at the chosen saving level.

- The Infinitely-lived agents model is solved with a "successive approximation" procedure on the set of value functions. The value functions are computed relying on the Bellman equations (14)-(15). I start from a set of guesses, $V_e(a)_0$, and $V_u(a)_0$, I compute the vector of parameters Ω representing the Schumaker spline approximations of the value functions, and I solve the constrained maximization problems and retrieve the policy functions, $a_e'(a)$ and $a_u'(a)$. As for the approximation method, I rely on the quadratic spline approximations for the future value functions, when evaluated at the chosen saving level. I keep on iterating until a fixed point is reached, i.e. until two successive iterations satisfy:

$$\underset{a}{Sup}\left|V_e(a)_{n+1} - V_e(a)_n\right| < 10^{-8} \text{ and } \underset{a}{Sup}\left|V_u(a)_{n+1} - V_u(a)_n\right| < 10^{-8}.$$

  Typically, around $2,000$ iterations are needed to reach a fixed point with the initial guess (such that the agents do not save), and between 500 and $1,500$ with the more accurate guesses in the subsequent steps towards the asset market clearing.

- I do not use a sample of individuals to approximate the stationary distributions, in order to avoid sampling variability confounding the results of the Monte Carlo experiments. In the OLG model, the distributions are computed relying on their definitions (10)-(11). I rely on these recursions and compute numerically the transitions functions. In the Infinitely-lived agents model, the stationary distributions are computed relying on their definitions (17)-(18). I start from a set of guesses, $\mu_e(a)_0$, and $\mu_u(a)_0$ and keep on iterating until convergence, i.e. until two successive iterations satisfy:

$$\underset{a}{Sup}\left|\mu_e(a)_{n+1} - \mu_e(a)_n\right| < 10^{-10} \text{ and } \underset{a}{Sup}\left|\mu_u(a)_{n+1} - \mu_u(a)_n\right| < 10^{-10}.$$

  Between $36,000$ and $250,000$ iterations are needed to reach a fixed point. Finally, to better approximate the stationary distributions, I use 13 grid points between every pair of grid-points in the original grid on assets. Between grid-points, I use a linear approximation scheme. For more details, see Rios-Rull (1999).

- The asset market is in equilibrium when the current guess for the interest rate r_0 achieves a capital excess demand which is less than 0.1% of the market size. In turn, this implies that the excess demand in the final good market is always less than 0.1% of the market size. This tolerance level could seem a minor aspect of the analysis. However, it is quite important. It was found that, for the policies with a quantitatively small true welfare effect, relying on a loose market clearing criterion was frequently preventing the distribution of welfare effects to display a convergence behavior when increasing the sample size.

- The welfare measures W_{OLG}, W, W_e and W_u are just the numerical integral of the value functions, integrated with respect to the steady state distributions.

- All convergence criteria are quite strict, and more stringent than what is normally used for these models. However, they are needed in order to avoid numerical errors systematically biasing the computation of the equilibria and their associated welfare measures. In the infinitely-lived agents model, some experimentation showed that asset grids with less than 501 points led to welfare effects that differed in a non trivial way from the true ones, when these were less than 1%. The same comment applies to virtually all other convergence criteria. The required level of precision, however, drastically impacts the computational time. Finally, the normalized maximum Euler Equation errors are sufficiently small. In the baseline models they are 2×10^{-5} or less.

Appendix D - Monte Carlo Algorithms

This algorithm represents the computational procedure used to solve the Monte Carlo experiments with the OLG model:

1. Simulate for 500 times a sample of individuals of given size N for 20 periods.
2. Estimate both the persistence and the variance of the income process with the Blundell and Bond GMM estimator, and assign them to the model's parameters $\rho_{y,m}$ and $\sigma_{y,m}$.
3. Generate a discrete grid over the asset space $[-b, ..., a_{\max}]$.
4. Generate a discrete grid over the income shocks with the Rouwenhorst method $[\varepsilon_{\min}, ..., \varepsilon_{\max}]$.
5. Start the loop for the benchmark economy.
6. Set the discount factor β_0 (if applicable).
7. Get the labor supply L.
8. Get the pension benefits p.
9. Set the capital tax τ_a.
10. Guess the interest rate r_0.
11. Get the capital demand K_0 and wages w_0.
12. Get the equilibrium taxes τ_p, τ_w.
13. Get the saving functions $a_t'(a, \varepsilon, f), a_t^{R\prime}(a)$ and the value functions $V_t(a, \varepsilon, f), V_t^R(a)$.
14. Get the stationary distributions $\mu_t(a, \varepsilon, f), \mu_t^R(a)$.
15. Get the aggregate capital supply and check the asset market clearing; Get r_1.
16. Update r_0' (with a bi-section method).
17. Iterate until asset market clearing.
18. Update β_0' (with a bi-section method, if applicable).
19. Iterate until matching $r_1 \simeq r_0' = 4\%$ (if applicable).
20. Get the consumption functions $c_t(a, \varepsilon, f), c_t^R(a)$ and check the final good market clearing.
21. Compute the ex-ante welfare W_{OLG}.
22. Start the loop for the counterfactual economy (i.e. under the new policy regime) and repeat steps $5 - 22$.
23. Save the output and repeat from step 2 for 500 times.

This algorithm represents the computational procedure used to solve the Monte Carlo experiments with the ∞-lived agents model:

1. Simulate for 500 times a sample of individuals of given size N and compute two statistics: the average unemployment rate and the average unemployment duration; store them.
2. Read the two simulated moments and match them exactly with the markov chain probabilities.
3. Generate a discrete grid over the asset space $[-b, ..., a_{\max}]$.
4. Start the loop for the benchmark economy.
5. Get the employment level L.
6. Guess the interest rate r_0.
7. Get the capital demand K_0, wages w_0, and unemployment benefits $\phi_u w_0$.
8. Get the equilibrium taxes τ_u.
9. Get the saving functions $a_e'(a), a_u'(a)$.
10. Get the stationary distributions $\mu_e(a), \mu_u(a)$.
11. Get the aggregate capital supply and check the asset market clearing; Get r_1.
12. Update r_0' (with a bi-section method).
13. Iterate until asset market clearing.
14. Get the consumption functions $c_e'(a), c_u'(a)$ and check the final good market clearing.
15. Compute the equilibrium value functions $V_e(a), V_u(a)$ and the ex-ante welfare W, W_e and W_u.
16. Start the loop for the counterfactual economy (i.e. under the new policy regime) and repeat steps $4 - 16$.
17. Save the output and repeat from step 2 for 500 times.

Appendix E - Additional Results and Figures

Sample Size	*Min*	*Max*	*Mean*	*Med*	*C.v.*
ρ_y (*DGP*=0.977)					
0.2k	0.9227	0.999	0.9755	0.9770	0.0187
0.8k	0.9415	0.999	0.9765	0.9764	0.0106
3.2k	0.9620	0.9935	0.9766	0.9767	0.0051
12.8k	0.9677	0.9849	0.9770	0.9771	0.0027
σ_y (*DGP*=0.122)					
0.2k	0.1148	0.1273	0.1202	0.1202	0.0152
0.8k	0.1174	0.1231	0.1201	0.1200	0.0078
3.2k	0.1187	0.1214	0.1200	0.1200	0.0037
12.8k	0.1193	0.1207	0.1200	0.1200	0.0019

Table 10: Calibration, OLG Model, French (2005) DGP - "Simulated" Parameters: Descriptive Statistics

Welfare Change (%)	*Min*	*Max*	*Mean*	*Med*	*C.v.*
Robustness - $\delta = 0.0457$ *(true* $\varpi = 2.792$*)*					
0.2k	2.584	2.898	2.777	2.793	0.037
0.8k	2.607	2.893	2.787	2.787	0.026
3.2k	2.678	2.892	2.790	2.790	0.015
12.8k	2.720	2.865	2.793	2.793	0.008
Robustness - $\alpha = 0.40$ *(true* $\varpi = 2.420$*)*					
0.2k	2.137	2.583	2.412	2.417	0.060
0.8k	2.178	2.583	2.419	2.413	0.041
3.2k	2.276	2.583	2.418	2.417	0.023
12.8k	2.323	2.517	2.421	2.420	0.012
Robustness - $\sigma_f^2 = 0.211$ *(true* $\varpi = 2.064$*)*					
0.2k	1.990	2.141	2.067	2.062	0.021
0.8k	2.010	2.133	2.064	2.062	0.015
3.2k	2.022	2.129	2.063	2.065	0.009
12.8k	2.034	2.098	2.065	2.067	0.004
Robustness - $\theta = 3.0$ *(true* $\varpi = 0.284$*)*					
0.2k	0.005	0.406	0.230	0.282	0.4905
0.8k	0.031	0.399	0.262	0.277	0.3231
3.2k	0.127	0.400	0.280	0.281	0.1972
12.8k	0.181	0.371	0.286	0.286	0.1098

Table 11: Monte Carlo Results - Welfare Effects Robustness with fixed Discount Factor

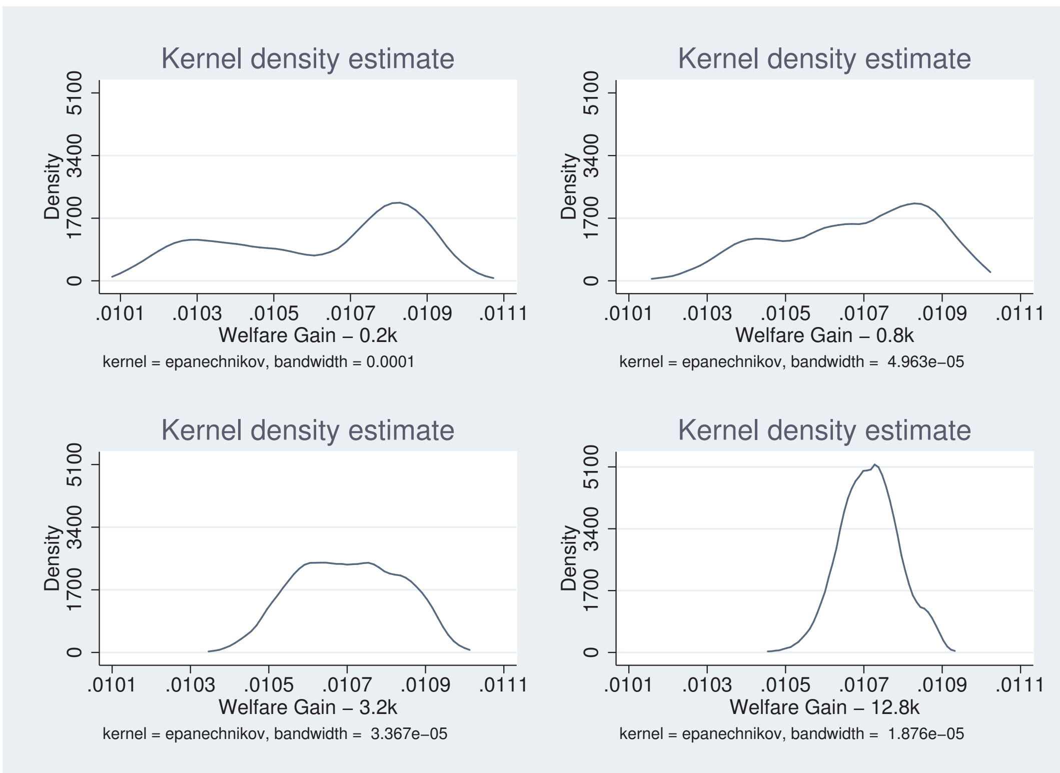

Figure 11: Overall Welfare Gains, OLG model, Capital Tax Experiment

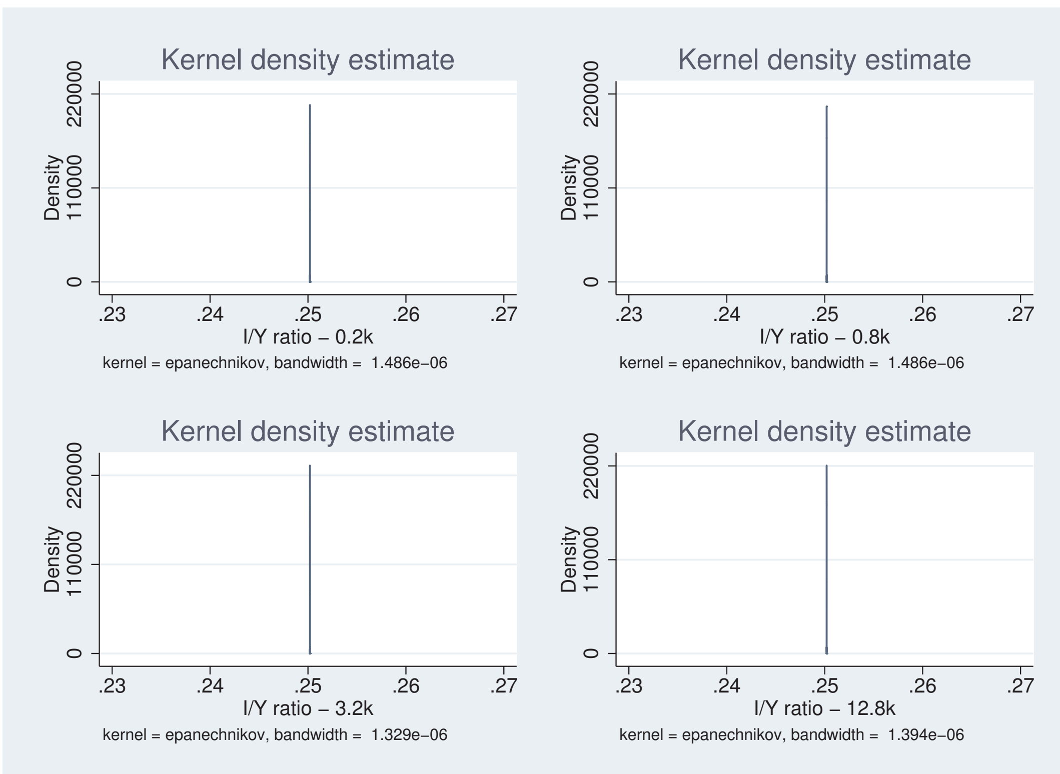

Figure 12: Investment/Output Ratio, OLG economy with recalibrated β.

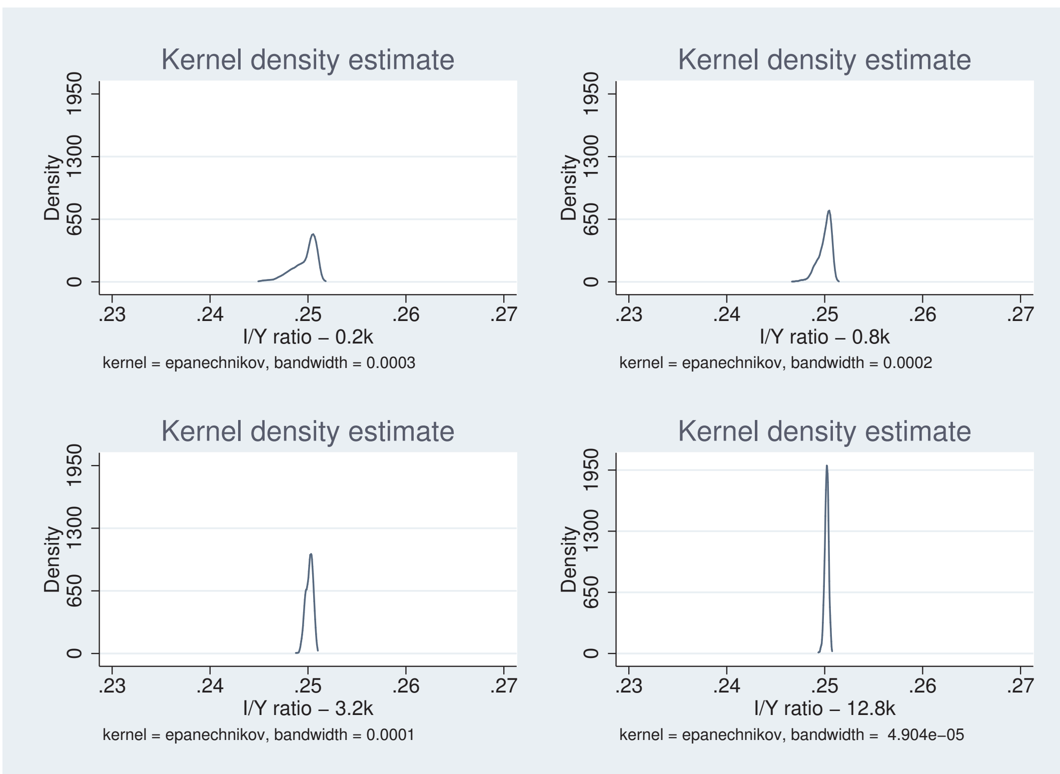

Figure 13: Investment/Output Ratio, OLG economy with fixed β.

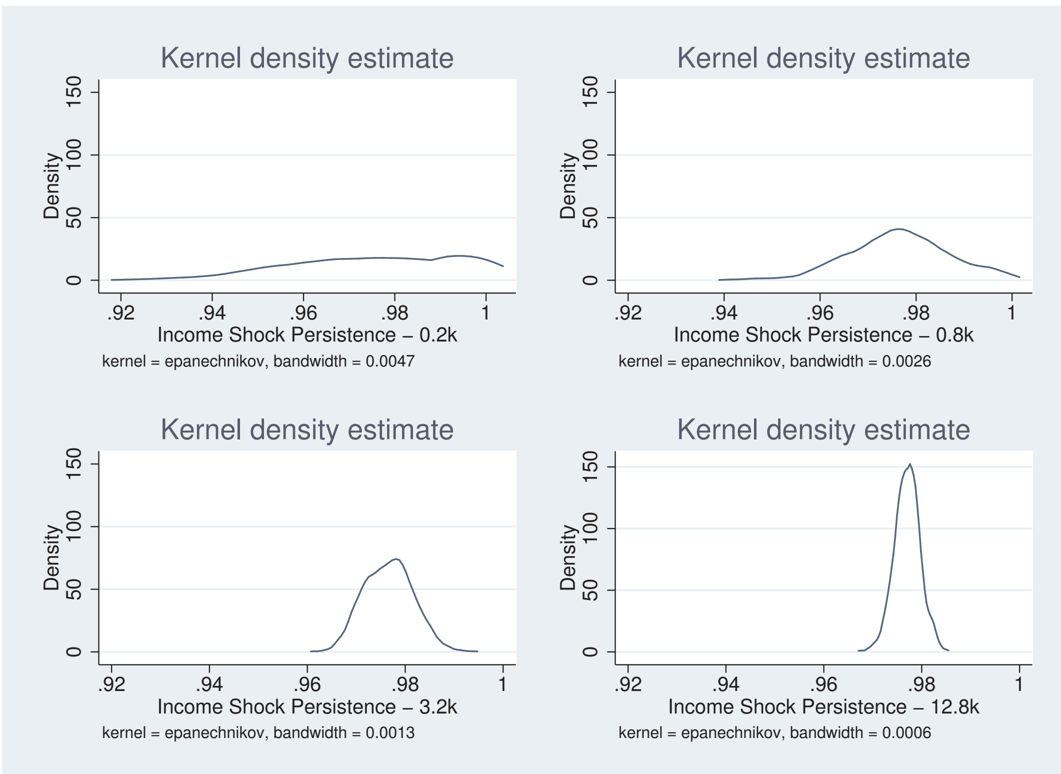

Figure 14: Calibration, OLG Model, French (2005) DGP - "Simulated" Parameters: ρ_y

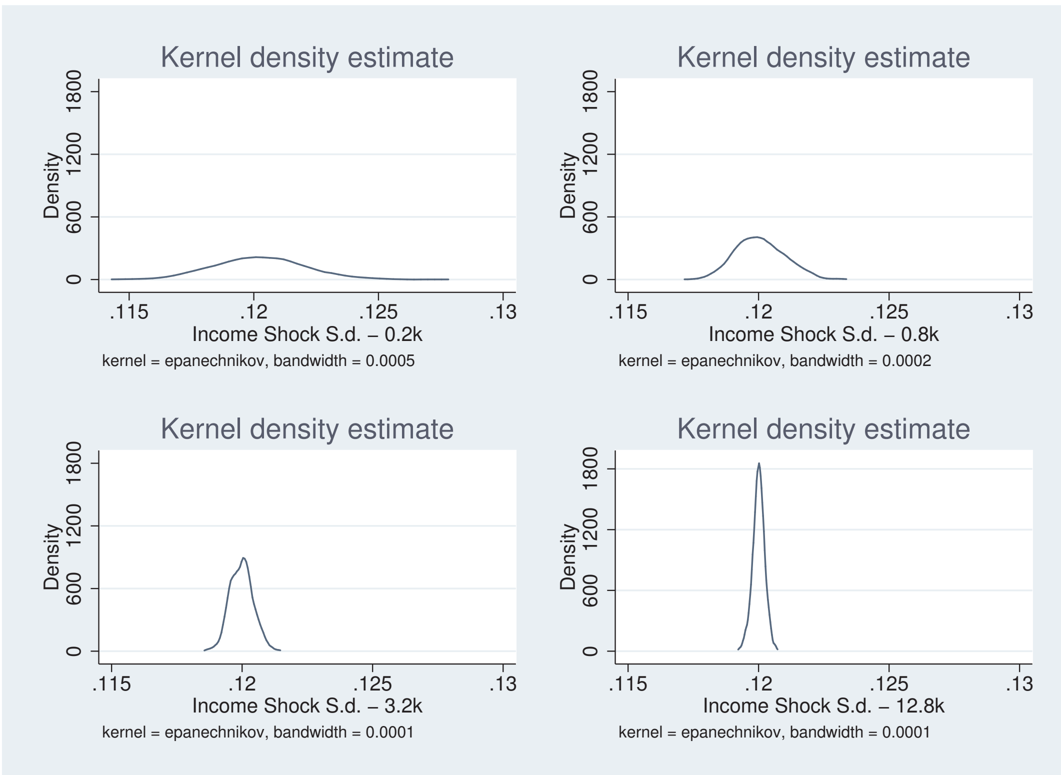

Figure 15: Calibration, OLG Model, French (2005) DGP - "Simulated" Parameters: σ_y

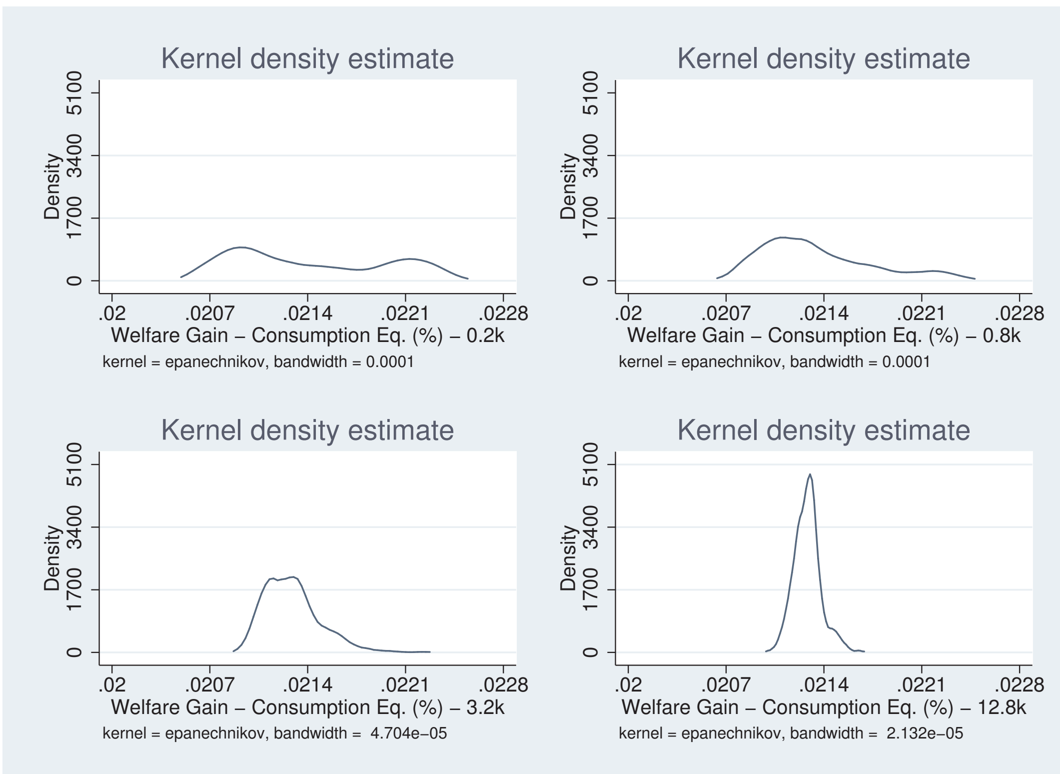

Figure 16: Overall Welfare Gains (ϖ), OLG model, Capital Tax Experiment, French (2005) DGP - Case 1c (Fixed β)

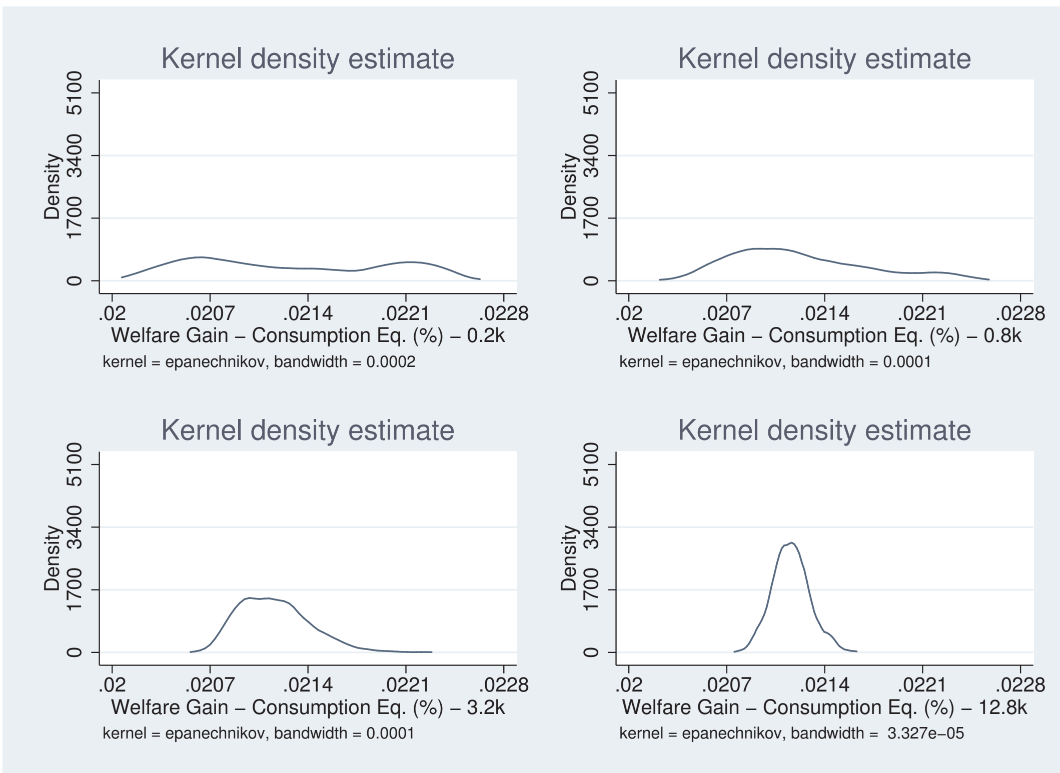

Figure 17: Overall Welfare Gains (ϖ), OLG model, Capital Tax Experiment, French (2005) DGP - Case 1d (Recalibrated β)

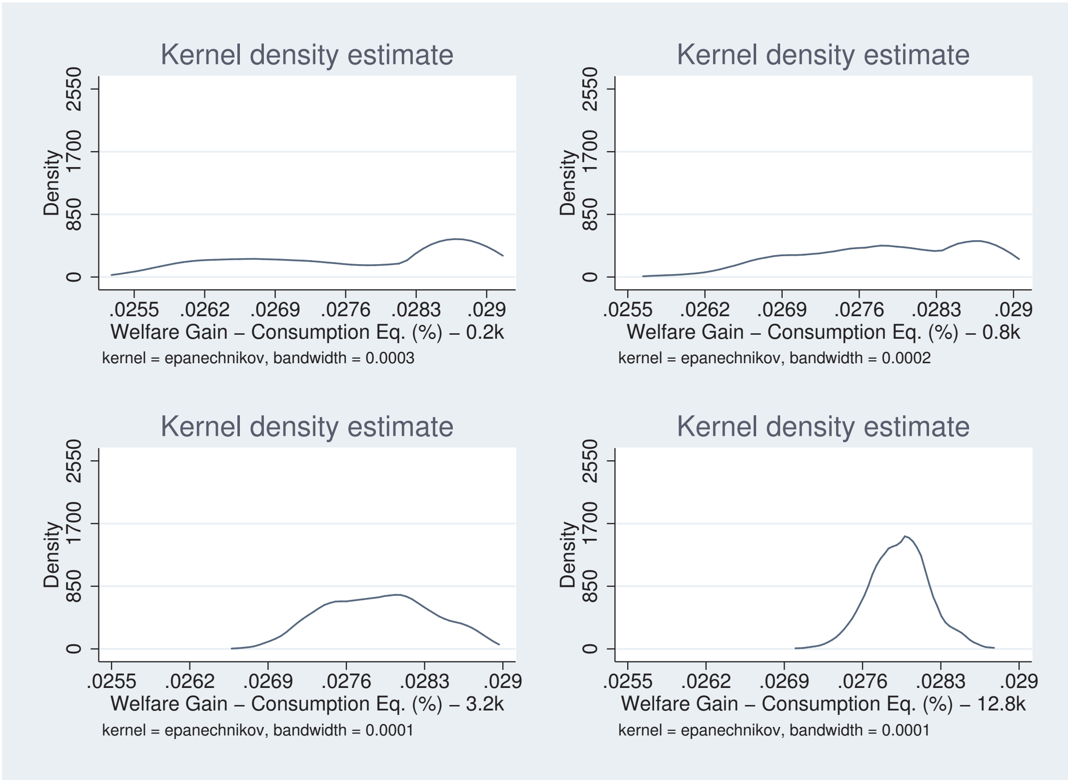

Figure 18: Overall Welfare Gains (ϖ), OLG model, Capital Tax Experiment, Robustness ($\delta = 0.0457$)

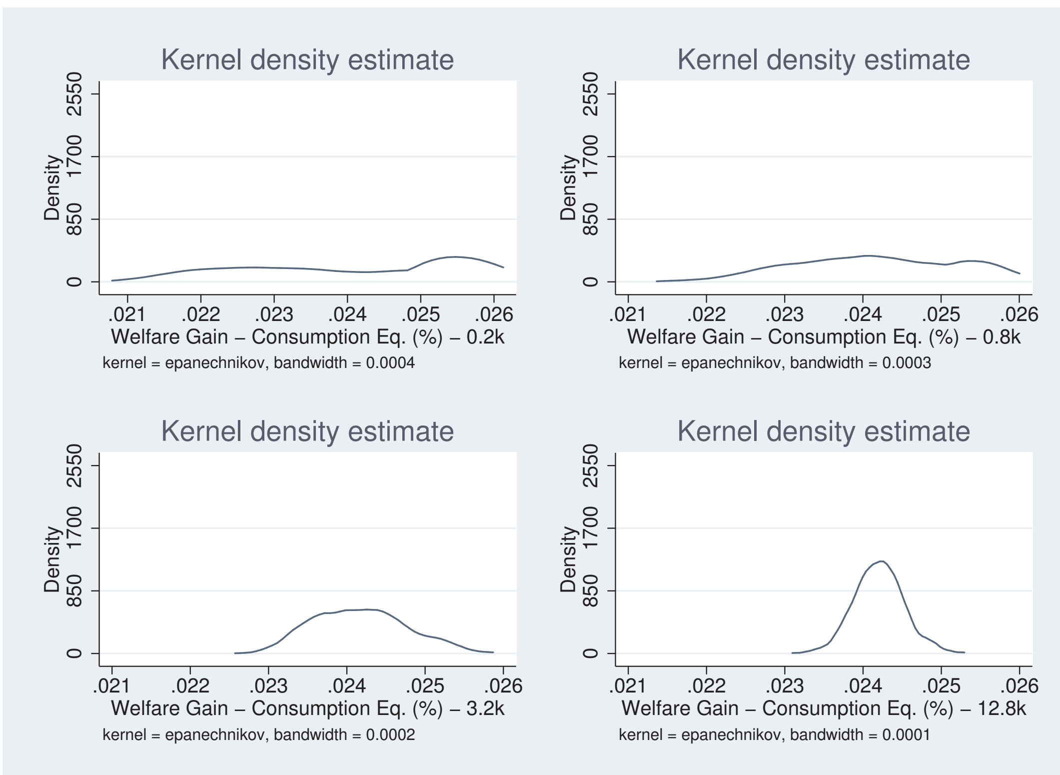

Figure 19: Overall Welfare Gains (ϖ), OLG model, Capital Tax Experiment, Robustness ($\alpha = 0.40$)

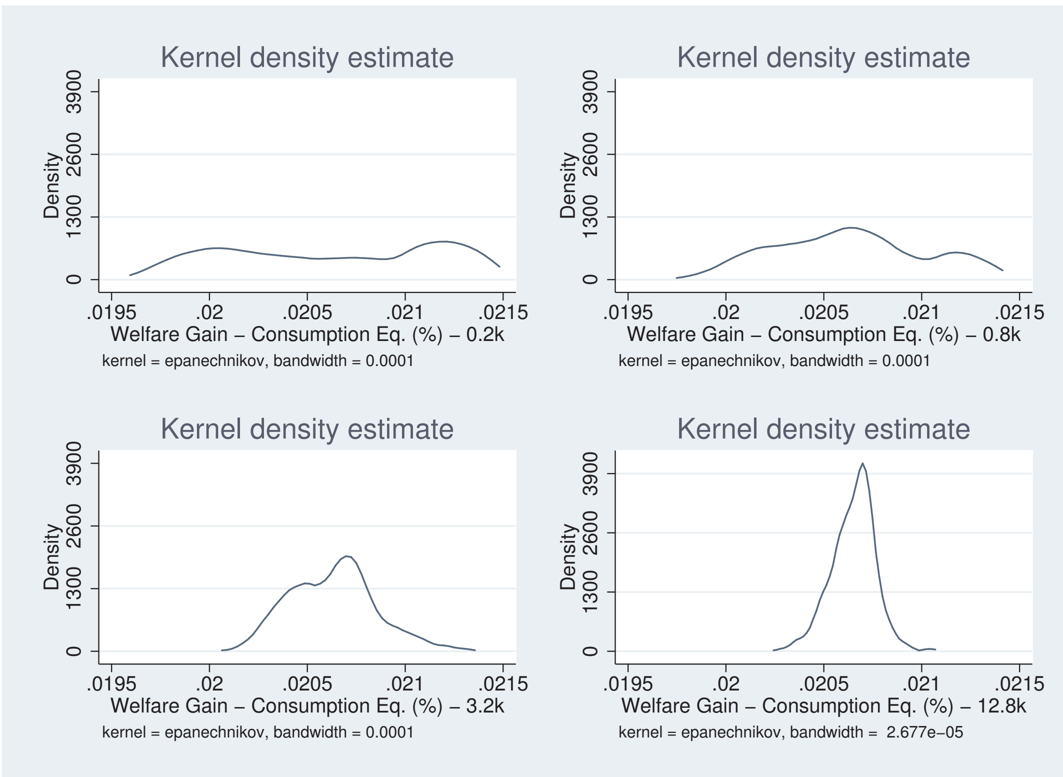

Figure 20: Overall Welfare Gains (ϖ), OLG model, Capital Tax Experiment, Robustness ($\sigma_f^2 = 0.211$)

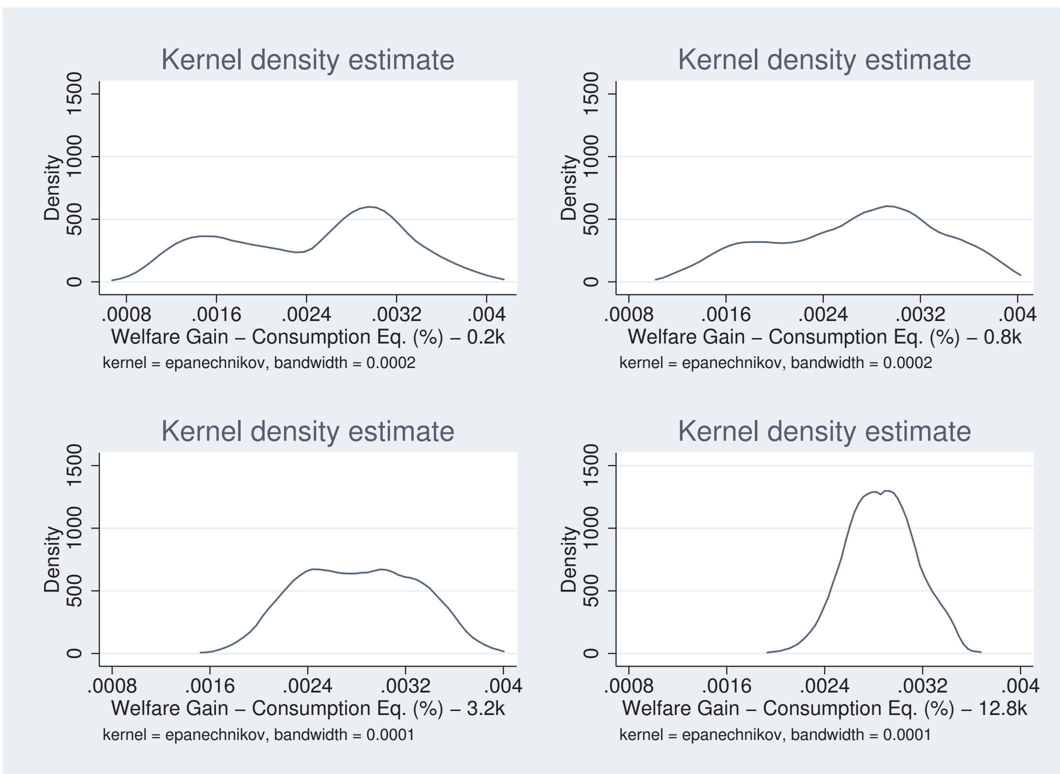

Figure 21: Overall Welfare Gains (ϖ), OLG model, Capital Tax Experiment, Robustness ($\theta = 3.0$)

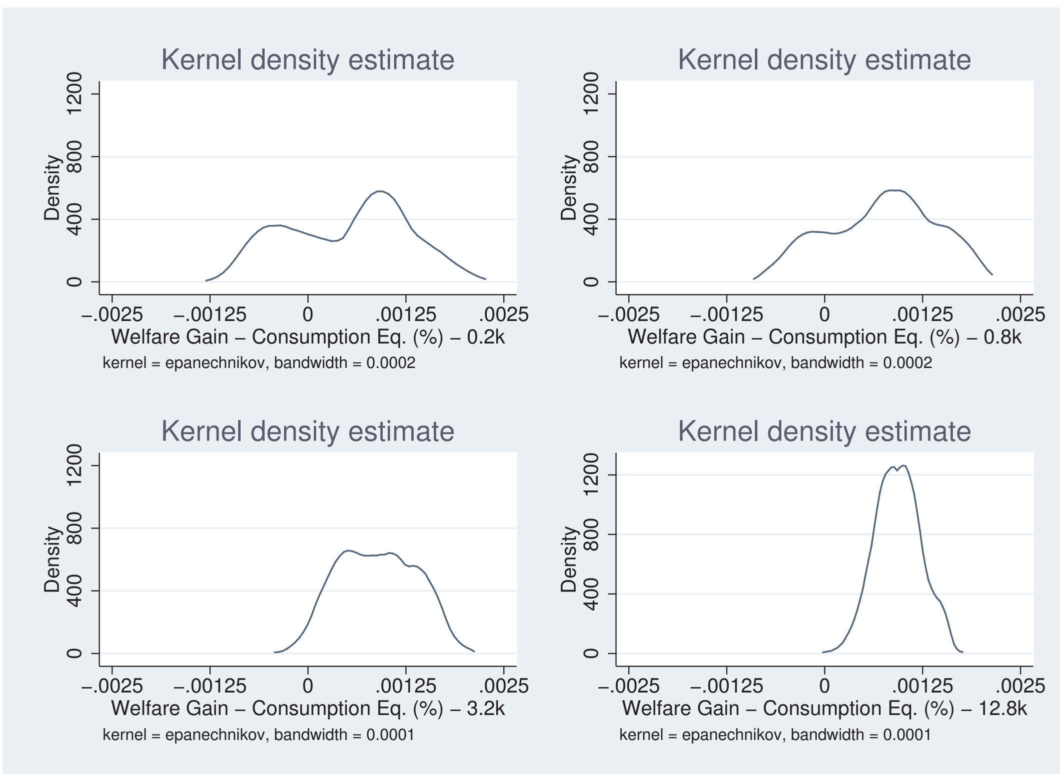

Figure 22: Overall Welfare Gains (ϖ), OLG model, Capital Tax Experiment, Robustness ($\theta = 3.15$)