



Working Paper Series

Estimating the Effects of Demographics on Interest Rates: A Robust Bayesian Perspective

WP 20-14

Paul Ho
Federal Reserve Bank of Richmond

This paper can be downloaded without charge
from: <http://www.richmondfed.org/publications/>



Richmond • Baltimore • Charlotte

Estimating the Effects of Demographics on Interest Rates: A Robust Bayesian Perspective*

Paul Ho[†]

Federal Reserve Bank of Richmond

paul.ho@rich.frb.org

October 7, 2020

Abstract

There are a vast range of estimates for the effect of demographics on interest rates. I show that these magnitudes are not well-identified without data on capital and life-cycle consumption. However, these data are often omitted. Using nonparametric prior sensitivity analysis for an overlapping generations model estimated through Bayesian methods, I show that without these data, small changes in the prior for the discount rate, intertemporal elasticity of substitution, and capital depreciation rate can shift the posterior quantiles for the effects of demographics by up to 1.5 percentage points. Data on the capital-output ratio and life-cycle consumption tighten identification.

*Download the latest version of the paper [here](#).

[†]I am indebted to Jaroslav Borovička, Chris Sims, and Mark Watson for their guidance. I thank SeHyoung Ahn, Carlos Viana de Carvalho, Jesús Fernández-Villaverde, Federico Huneus, Nobuhiro Kiyotaki, Ezra Oberfield, Mikkel Plagborg-Møller, and numerous seminar participants for comments and suggestions. The views expressed herein are those of the author and are not necessarily the views of the Federal Reserve Bank of Richmond or the Federal Reserve System.

1 Introduction

Secular changes in demographics and interest rates in developed countries over the past thirty years have made quantifying the effects of an aging population on the interest rate crucial for forecasting and policy analysis. The literature has sought estimates using various overlapping generations (OLG) models but disagrees on the magnitude of these effects. For instance, [Attanasio et al. \(2006\)](#), [Carvalho et al. \(2016\)](#), and [Fujita and Fukiwara \(2016\)](#) find that between the early-1980s and mid-2010s, demographic changes contribute to declines in the real interest rate of 2.5, 1.5, and 0.9 percentage points, respectively. For the period between 1980 and 2020, [Ikeda and Saito \(2014\)](#) predict a 0.3 percentage point decline due to demographic changes.¹ One key driver of these differences is the choice of structural parameters (see [Figure D.1](#) for the range of values of the discount rate, intertemporal elasticity of substitution, and capital depreciation rate used in the literature).

This paper shows that the effect of demographics on interest rates is not well-identified with the data typically used in the literature. Instead, data on capital and household consumption over the life cycle are important for accurate estimates. Without these data, the estimated effects of demographics on interest rates in an OLG model vary substantially with the prior for the discount rate, intertemporal elasticity of substitution, and capital depreciation rate even though other parameters are well-identified.² We establish these insights using robust Bayesian analysis. First, we use Bayesian methods to estimate the structural parameters of a parsimonious OLG model and show that the discount rate, intertemporal elasticity of substitution, and capital depreciation rate are not well-identified by the data. We then use nonparametric prior sensitivity analysis techniques from [Ho \(2020\)](#) to show how the prior for these parameters influences the estimated effects of demographics on interest rates. Finally, we show that including data on the capital-output ratio and consumption over the life cycle can tighten the likelihood and substantially change posterior mean estimates. These results show that calibrating or estimating OLG models without these data, as much of literature has done, can lead to misleading conclusions about the quantitative effect of demographics.

Our econometric framework is a vector autoregression (VAR) with a structural break that captures the secular changes in demographics, interest rates, and other macroeconomic

¹In comparison, the [Federal Reserve Bank of New York](#) estimates a 3 percentage point decline in the natural interest rate for the U.S. between 1980 and 2020.

²While the choice of model is also potentially important, these parameters are present in any model studying demographics and interest rates. [Section 5.3](#) discusses how our results extend to OLG models more generally. Cross-country differences can also lead to diverging estimates. However, if the model parameters are not well-identified, then it is hard to determine if the estimates actually arise from cross-country variation or simply the way the model is being fitted to data.

aggregates. The long-run averages of the VAR are determined by the steady state of an OLG model that captures the economic effects of demographics on the natural interest rate. We use the OLG model from [Gertler \(1999\)](#), which captures the main effects of demographics on macroeconomic variables but remains tractable and transparent. Given the structural parameters, the model implies an interest rate that we refer to as the natural interest rate, as in [Laubach and Williams \(2003\)](#). To measure the historical effect of demographics, we take the difference between the estimated natural interest rate and the counterfactual natural interest rate that would have arisen if specific demographic parameters had not experienced a structural break. In particular, we study the effects of population growth, life expectancy, and the relative productivity of retirees. No further restrictions are placed on the dynamics of the data, allowing flexibility in the high-frequency variation.

We use Bayesian methods to fit the model to eight macroeconomic and demographic time series from 1980-2013 for Japan, where secular macroeconomic and demographic changes have been especially pronounced. Bayesian estimation utilizes the full likelihood and avoids having to choose how to weight a potentially large number of overidentifying moments. Although the starkness of these changes improves identification, there remains substantial posterior uncertainty about both the effects of demographics and the underlying parameters. Our estimates of the realized and counterfactual natural interest rates have relatively wide 68% credible intervals of up to 0.8 percentage points. While the time-varying parameters and structural break date are tightly identified by the data, the likelihood for the time-invariant parameters is dispersed. In particular, the posterior standard deviations of these time-invariant parameters are at most 35% smaller than the prior standard deviations. Since the effects of demographics in the model are determined by these underlying parameters, the results suggest that the estimated effects are strongly influenced by the prior and that more data are required to pin down the effects.

To understand how the data inform the estimated effects of demographics, we use the relative entropy prior sensitivity (REPS) methodology from [Ho \(2020\)](#) to determine how much the estimated effects of demographics on interest rates depend on the prior for the three fixed parameters—the discount rate, intertemporal elasticity of substitution, and capital depreciation rate. REPS considers a nonparametric set of priors that are close to the original prior in relative entropy and seeks a *worst-case* prior that changes the posterior results the most. The worst-case priors yield bounds that contain the posterior credible intervals for any prior in that set and identify parts of the prior that are important for the posterior estimates. REPS does not limit one to parametric or infinitesimal changes in the prior, thus allowing one to check across an uncountable set of priors. We are therefore able to determine how informative the data are without having to estimate the model for every

plausible prior (i.e., changing the means, variances, correlations, parametric families, etc. used in the original prior). In addition, the worst-case posterior is derived from the same likelihood as the original estimation, ensuring that the data discipline the bounds we obtain. The methodology is thus a flexible but systematic way to measure how much the effects of demographics depend on the fixed parameters and how informative the likelihood is in the relevant directions.

The REPS analysis reveals that a small change in the prior can lead the posterior quantiles for the effects to shift by up to 1.5 percentage points. Many features of the worst-case prior are plausible *ex ante*, emphasizing the concern that the estimated effects are not well-identified by the data. In Appendix D, we show that these results are not dependent on the nonparametric nature of the exercise—reestimating the model under an alternative parametric prior that is consistent with existing calibrations in the literature can lead the posterior mean for the effects of demographics to change by over one posterior standard deviation, or up to 0.6 percentage points.

The large role of the prior for the posterior in our baseline estimation highlights the need to include additional data to discipline the parameters. Intuitively, the effects of demographics on interest rates are determined by the capital-demand and savings-supply functions. Capital and savings data provide measures of quantities, while the interest rate provides a measure of prices. By measuring savings and investment responses to changes in macroeconomic conditions, aggregate data on capital improves estimates of the intertemporal elasticity of substitution and depreciation rate. In addition, consumption and savings over the life cycle reveal how households trade off current and future consumption, which is informative about the discount rate and intertemporal elasticity of substitution.

We incorporate the capital-output ratio into the original set of time series to show how data on capital sharpens our estimates of the effects of demographics by providing information on the underlying parameters, especially the depreciation rate. The additional data reduces the standard deviation of the intertemporal elasticity of substitution and depreciation rate by two and five times, respectively, but does not change the precision of the discount rate estimate. We obtain a more precise estimate for the effect of the relative productivity of retirees on interest rates, corroborating the result from REPS analysis that the depreciation rate is especially important for determining this effect. In addition, the new data shifts the posterior means of the effects of demographics by up to 1.5 percentage points.

As the appropriate data on life-cycle patterns of consumption are not available for Japan, we show indirectly that such data would be able to improve identification. In particular, we use Monte Carlo draws from the posterior for the estimation with the capital-output ratio and run quadratic regressions of the model-implied effects of demographics on the correspond-

ing steady-state life-cycle consumption. The steady-state consumption levels of workers and retirees (relative to total output) account for between 35% and 80% of the posterior uncertainty in the effects of demographics, and the ratio of average retiree consumption to average worker consumption can account for between 4% and 74% of the posterior uncertainty in the effects of demographics. We find suggestive evidence that the life-cycle consumption data is able to account for the effects of demographics by informing the posterior estimates for the discount rate and intertemporal elasticity of substitution.

Related Literature. Our results suggest that the numerical results from calibrated OLG models (e.g., Ikeda and Saito (2014); Sudo and Takizuka (2018) for Japan and Carvalho et al. (2016); Gagnon et al. (2016); Aksoy et al. (2019); Eggertsson et al. (2019) for the rest of the world) may be sensitive to their calibration strategies, as there is a wide range of possible parameter values consistent with the time series in our baseline estimation. Including the appropriate data on capital and consumption over the life cycle can help pin down the discount rate, intertemporal elasticity of substitution, and capital depreciation rate, which determine the effects of demographics on interest rates in OLG models. In Section 5.3, we discuss how these conclusions extend to questions about other effects of demographics and models with additional channels driving interest rates.

The prior sensitivity analysis extends the literature on identification in representative agent models (e.g., Canova and Sala (2009); Iskrev (2010); Komunjer and Ng (2011)) to OLG models. In related work, Janssens (2020) finds that the labor share, the capital depreciation rate and the intertemporal elasticity of substitution³ are not well-identified in the Aiyagari model estimated using indirect inference on aggregate data. In our OLG model, we find a similar conclusion that aggregate data is insufficient for identifying the intertemporal elasticity of substitution. On the firm side, while Janssens (2020) highlights that the labor share and capital depreciation rate are jointly identified, we find that the capital depreciation rate is poorly identified even with direct observations of the labor share.

In terms of methodology, the REPS analysis nonparametrically establishes the lack of identification given the data. The Bayesian approach acknowledges the relatively short length of the available time series and does not rely on asymptotics. REPS checks across a wider range of priors than the local method of Müller (2012), which only considers a limited parametric class of infinitesimal changes in the prior. In addition, the structural break framework contributes to the literature on estimating models using specific frequencies of the data, which we discuss in Section 2.2.

³The Aiyagari model in Janssens (2020) uses a constant relative risk aversion utility function that, unlike the recursive preferences in this paper, does not allow us to disentangle the intertemporal elasticity of substitution from household risk aversion.

Outline. The organization of the paper is as follows. I introduce the econometric model in Section 2 and outline the OLG model in Section 3. Section 4 describes the Bayesian estimation and Section 5 describes the prior sensitivity analysis. In Section 6, I show how additional data can improve the estimates for the effects of demographics. Section 7 concludes.

2 Econometric Framework

We now present the econometric framework that we use to disentangle secular changes in the data from the high-frequency variation. The secular change is modeled as a change in the steady state of an OLG model that occurs in response to a structural break in the macroeconomic and demographic parameters. The data is annual and the high-frequency variation is modeled as a mean-zero VAR(1) process appended to a constant term whose values are determined by the OLG model’s steady states. The empirical analysis will examine how the data inform estimates of both the time-varying and fixed parameters, as well as how the effects of demographics on interest rates in the OLG model depend on these parameters.

2.1 Setup

We observe annual data y_t that follows the model:

$$y_t = \mu(s_t) + v_t \tag{2.1}$$

$$v_t = \Phi(s_t)v_{t-1} + u_t \tag{2.2}$$

$$u_t \sim \mathcal{N}(0, \Sigma(s_t)), \tag{2.3}$$

where $s_t = \mathbf{1}\{t \geq t^*\}$ is an indicator for the structural break.⁴ The data y_t has a mean μ and dynamics v_t , which follow a mean zero VAR(1) process. The mean and VAR process change after the structural break in period t^* . The structural change in μ captures the secular changes in levels observed in the data. The structural break in the VAR process in period t^* allows for differences in the dynamics, including changes in the volatility, persistence, or comovement of the data after the structural break.

We model the secular changes in the data by assuming that μ is determined by the steady state of the OLG model with parameters that depend on s_t . These cross-equation restrictions on μ account for economic forces that determine the secular comovements in y_t . The OLG model also allows us to compute counterfactuals for μ . On the other hand, we allow for flexibility in the dynamics by modeling v_t as a reduced-form VAR. Since the OLG model

⁴Including a second break does not materially affect results.

plays no role in the high-frequency variation of the data, we avoid modeling the frictions necessary to match business cycle dynamics, which keeps the OLG model tractable both for estimation and interpretation. Instead, we isolate features of the data that the simple OLG model is best suited to explain.⁵

Estimating the econometric framework (2.1)-(2.3) is analogous to fitting the steady state of the OLG model to data from the start and end of the sample. However, instead of using data from a specific year, we estimate these steady states based on the full time series of data. By estimating the break date t^* , we allow the data to determine which periods correspond to the steady state before and after the structural break. The VAR flexibly accounts for correlation across time and variables.

2.2 Relation to Other Approaches

The model (2.1)-(2.3) decomposes the data into a long-run mean component $\mu(s_t)$ and a high-frequency component v_t . The long-run mean captures low-frequency changes in the data. We now discuss alternative approaches to extract low-frequency variation from data.

By using the OLG model to discipline only the long-run averages instead the full variation of the data, the approach described above is similar to the literature on Bayesian limited information estimation of DSGE models. For example, [Christiano et al. \(2010\)](#) estimate a DSGE model using only impulse responses to a set of identified structural shocks, thus focusing on the effect of these shocks while ignoring predictions of the model that are of less interest to the researcher. Since we are interested in secular changes in the economy here, we only use the steady state of the model to discipline the means μ . [Sala \(2015\)](#) estimates a DSGE model in the frequency domain, restricting the estimation to components of the likelihood corresponding to particular frequencies. In (2.1)-(2.3), the VAR captures the high-frequency variation, and the structural estimates are informed by the low-frequency component consisting of two long-run means, with a structural break date that is estimated.

In related work, [Del Negro and Schorfheide \(2004\)](#) use information from a DSGE model to inform the priors of a VAR. In our econometric framework, instead of using the OLG model to construct the prior, we directly model the mean of the VAR as coming from the steady state of the OLG model. While the object of interest for [Del Negro and Schorfheide \(2004\)](#) is the VAR estimates, in this paper we are interested in the estimates of the OLG model's structural parameters and implied counterfactuals.

⁵In the absence of additional frictions, the transition path for the OLG model we consider involves jumps at the break dates for non-demographic variables. On the break date, these variables either overshoot the new steady state or go in the opposite direction from the change in steady state. As a result, the model-implied transition paths are unrealistic, leading to implausible estimates.

If we had fitted both low and high-frequency variation to the OLG model, we would have required a rich set of frictions in the OLG model to produce realistic dynamics, then incorporate these dynamics in the estimation. [Smets and Wouters \(2007\)](#) do this in the context of a representative agent DSGE model. Log-linearizing such a model would yield the system (2.1)-(2.3) without the regimes s_t , but would impose restrictions on Φ and Σ arising from the equilibrium conditions in the structural model. For a given set of structural parameters, our approach yields the same μ if the frictions affect the dynamics but not the steady state. However, we remain agnostic about the modeling of frictions. Moreover, the approach here ensures that the structural parameter estimates are primarily driven by the low-frequency variation in the data, which is the object of interest. [Sala \(2015\)](#) shows that the posterior estimates in a representative agent DSGE model depend on the frequencies used in the estimation, suggesting that the distinction between μ and v_t is important.

One could also filter out the low-frequency fluctuations in the data using weighted averages without reference to any model (e.g., [Hodrick and Prescott \(1997\)](#), [Müller and Watson \(2018\)](#)), then use these low-frequency fluctuations as input to estimate the parameters of the OLG model. However, the high-frequency observations are informative for estimating the time-varying parameters as well as the structural break date.

3 Overlapping Generations Model

The OLG model is similar to the one in [Gertler \(1999\)](#), which is frequently used in the literature (e.g., [Fujiwara and Teranishi \(2008\)](#); [Carvalho et al. \(2016\)](#); [Kara and von Thadden \(2016\)](#)). It is a neoclassical growth model with endogenous labor supply, stochastic retirement and death, and recursive preferences. The model is a parsimonious way to capture the main economic forces linking demographics to interest rates. In this section, we describe the setup of the model and focus on key equilibrium conditions for the relationship between demographics and interest rates. The full set of steady-state equilibrium conditions is listed in [Appendix A](#).

3.1 Households

3.1.1 Life Cycle and Population Growth

Each individual is born as a worker. At the end of each period, a worker has probability $1 - \omega$ of retiring and a retiree has probability $1 - \gamma$ of dying. Denoting the stock of workers by N_t , we assume $(1 - \omega + n)N_t$ new workers are born each period, so that the workforce

grows at a constant rate n . The ratio of retirees to workers ψ_t satisfies:

$$\psi_{t+1}N_{t+1} = \gamma\psi_tN_t + (1 - \omega)N_t. \quad (3.1)$$

3.1.2 Retirees

Retirees have preferences:

$$V_t^{r,i,j} = \left[\left[(C_t^{r,i,j})^\nu (1 - L_t^{r,i,j})^{1-\nu} \right]^\rho + \beta\gamma (V_{t+1}^{r,i,j})^\rho \right]^{\frac{1}{\rho}}, \quad (3.2)$$

where i and j indicate the birth and retirement cohort, respectively. $C_t^{r,i,j}$ and $L_t^{r,i,j}$ are consumption and labor of retirees, respectively. These preferences imply risk neutrality with respect to wealth, which allows for aggregation across cohorts. The intertemporal elasticity of substitution, $\sigma \equiv \frac{1}{1-\rho}$, controls the desire to smooth consumption over time, which is a key force for determining households' propensity to save and hence the interest rate. The survival rate γ augments the discount factor β , so that retirees have an effective discount rate of $\beta\gamma$.

We assume perfect annuity markets (Yaari (1965); Blanchard (1985)) that insure against the risk arising from the uncertain time of death. Each retiree places her wealth in a mutual fund that invests its proceeds. The surviving fraction of retirees γ receive all the returns, while those who die receive nothing. Retirees are therefore subject to the budget constraint:

$$A_{t+1}^{r,i,j} = \frac{R_t}{\gamma}A_t^{r,i,j} + W_t\xi L_t^{r,i,j} - C_t^{r,i,j}, \quad (3.3)$$

where $A_t^{r,i,j}$ is the level of assets, W_t is the wage per effective unit of labor, $\xi \in [0, 1]$ is the productivity of a retiree relative to a worker, and R_t/γ is the return to a surviving retiree.

3.1.3 Workers

Workers have preferences:

$$V_t^{w,i} = \left[\left[(C_t^{w,i})^\nu (1 - L_t^{w,i})^{1-\nu} \right]^\rho + \beta \left[\omega V_{t+1}^{w,i} + (1 - \omega) V_{t+1}^{r,i,t+1} \right]^\rho \right]^{\frac{1}{\rho}} \quad (3.4)$$

and are subject to the budget constraint:

$$A_{t+1}^{w,i} = R_tA_t^{w,i} + W_tL_t^{w,i} - C_t^{w,i}. \quad (3.5)$$

The presence of the retirees' value function $V_{t+1}^{r,i,t+1}$ in (3.4) implies that workers take retirement into account when making savings decisions. Therefore, the value of γ and ξ implicitly enter into the workers' decisions.

3.1.4 Aggregation

Retirees. Retiree consumption is linear in the sum of assets and discounted expected value of labor income, with marginal propensity to consume $\epsilon_t \pi_t$ that does not depend on cohort. This allows us to write aggregate retiree consumption as:

$$C_t^r = \epsilon_t \pi_t \left(\frac{R_t}{\gamma} A_t^r + H_t^r \right), \quad (3.6)$$

where human wealth:

$$H_t^r = W_t \xi L_t^r + \frac{\psi_t}{\psi_{t+1} (1+n)} \frac{H_{t+1}^r}{R_{t+1}} \quad (3.7)$$

is defined as the present discounted value of expected labor income for the entire population of retirees. The first term is the labor income earned in period t , while the $\frac{\psi_t}{\psi_{t+1}(1+n)}$ term accounts for the population growth of retirees.

Workers. Similarly, workers have a common marginal propensity to consume π_t . Aggregate worker consumption can be written as:

$$C_t^w = \pi_t (R_t A_t^w + H_t^w), \quad (3.8)$$

where human wealth is defined as:

$$H_t^w = W_t L_t^w + \omega \frac{1}{1+n} \frac{H_{t+1}^w}{R_{t+1} \Omega_{t+1}} + (1-\omega) \xi^{\nu-1} \epsilon_{t+1}^{\frac{1}{1-\sigma}} \frac{1}{\psi_{t+1} (1+n)} \frac{H_{t+1}^r}{R_{t+1} \Omega_{t+1}}. \quad (3.9)$$

The first term is the labor income of workers. The next two terms are the expected present discounted value of human wealth in period $t+1$. The variable Ω_{t+1} augments R_{t+1} to account for the possibility of retiring at the end of the period and is defined as:

$$\Omega_{t+1} = \omega + (1-\omega) \xi^{\nu-1} \epsilon_{t+1}^{\frac{1}{1-\sigma}}. \quad (3.10)$$

Workers adjust their valuation of future labor income in response to two changes that happen when they retire—they become less productive by an exogenous factor of ξ , and their marginal propensity to consume increases by an endogenously determined factor of ϵ_{t+1} .

State Variables. The linearity of the consumption decisions implies that we do not need to keep track of individual cohorts when solving for the aggregate steady state or dynamics. Instead, we can aggregate across workers and retirees, respectively, which reduces the number of states and makes Bayesian estimation feasible.

3.2 Production and Market Clearing

Aggregate output follows the Cobb-Douglas production function:

$$Y_t = (X_t L_t)^\alpha K_t^{1-\alpha}, \quad (3.11)$$

where $L_t \equiv L_t^w + \xi L_t^r$ and α determine the labor share. The labor-augmenting productivity X_t grows at a constant rate x , and capital depreciates at rate δ . Market clearing for the capital and goods markets implies:

$$A_t = K_t \quad (3.12)$$

$$Y_t = K_{t+1} - (1 - \delta) K_t + C_t^w + C_t^r. \quad (3.13)$$

3.3 Structural Break

We assume that the long-run mean $\mu(s_t)$ in equation (2.1) is determined by the steady state of the OLG model above. When the structural break occurs in period t^* , a subset of the structural parameters are redrawn from the same distribution that generated the parameters before t^* . This new set of parameters yields a new steady-state and long-run mean. The demographic parameters affected by the structural break are the working population growth n , the survival rate of retirees γ , and the relative productivity of retirees ξ . The non-demographic parameters affected are the productivity growth x , the labor share α , and the parameter ν controlling the disutility of labor.

3.4 Demographics and Interest Rates

The model captures several channels through which an aging demographic affects the steady-state interest rate. The strength of each channel depends nonlinearly on the structural parameters. In what follows, we drop the t subscript to denote steady-state variables.

Firstly, as the share of retirees increases, the supply of savings decreases, thus raising the interest rate. This *savings composition channel* arises because retirees have a higher

marginal propensity to consume, as can be seen from the steady-state expressions:

$$\epsilon\pi = 1 - \left(\frac{R}{(1+x)^{1-\nu}} \right)^{\sigma-1} \beta^\sigma \gamma \quad (3.14)$$

$$\pi = 1 - \left(\frac{R\Omega}{(1+x)^{1-\nu}} \right)^{\sigma-1} \beta^\sigma \quad (3.15)$$

for retiree and worker marginal propensities to consume $\epsilon\pi$ and π , respectively. Retirees have an effective discount rate of $\beta\gamma < \beta$, which induces them to consume more from their wealth. Moreover, workers save to smooth consumption into retirement, since retirement leads to lower productivity and hence lower expected wealth. This force is captured by the Ω adjustment defined in (3.10) to the interest rate R in equation (3.15).

Secondly, increases in the survival rate γ or decreases in the relative productivity ξ cause decreases in the marginal propensity to consume within groups, which decreases interest rates through an increase in the supply of savings. We refer to this as the *within-group savings channel*. When retirees have a higher probability of survival, their effective discount rate $\beta\gamma$ increases, which increases the incentive to save in order to smooth consumption. Workers anticipate the reduced propensity to consume as a retiree, reflected by the $V_t^{r,i,t+1}$ term in their continuation value (3.4), and respond by decreasing their marginal propensity to consume as well. A lower relative productivity of retirees ξ implies a greater drop in human wealth upon retirement, leading to increased saving by workers to smooth consumption.

Changes in demographics also affect interest rates through a *capital demand channel*, which is captured by the equilibrium condition for capital in steady state:

$$R = (1 - \alpha) k^{-1} + (1 - \delta), \quad (3.16)$$

where k is the steady-state capital-output ratio K_t/Y_t . In particular, when the share of retirees increases or the relative productivity of retirees decreases, the average household in the economy becomes less productive. In the absence of adjustments to labor supply, this lowers the marginal product of capital $(1 - \alpha) k^{-1}$, which leads to a fall in the demand for capital that pushes interest rates downward. On the other hand, when workers expect to have a longer or less productive retirement, they increase their labor supply to accumulate wealth for retirement. This can dampen or even reverse the capital demand effect.

Finally, there is a *general equilibrium channel*, as defined in Carvalho et al. (2016). In response to a decrease in interest rates from the savings composition, within-group savings, or capital demand channels, households decrease savings, and firms increase their use of capital. These forces amplify the direct effect of the savings composition, within-group savings, and

capital demand channels.

In section 5, we analyze the sensitivity of our results to the prior for the discount factor β , intertemporal elasticity of substitution σ , and capital depreciation rate δ , because these three parameters influence the effects described above. A higher discount rate reduces the incentive to smooth over time, since households place a lower weight on future utility. This greater discounting weakens the savings composition and within-group savings channels. Since the two channels produce opposite effects, the net effect of changing β depends on the parameters of the model. A higher intertemporal elasticity of substitution σ strengthens the effect of the discount factor β but decreases the sensitivity of the workers' marginal propensity to consume π to the interest rate adjustment Ω . In particular, when $\sigma = 1$, Ω no longer shows up in (3.15), and the incentive to smooth consumption into retirement vanishes, reducing the savings composition and within-group savings effects. A larger σ also decreases households' response to changes in the interest rate, thus reducing the general equilibrium effect. A higher capital depreciation rate δ increases the amount of investment necessary to maintain the steady-state level of capital, hence amplifying the capital demand effect.

Any demographic change influences the interest rate through a combination of the channels described above. The relative importance of these channels determines the equilibrium response of interest rates. To quantify the effects of demographics, one needs to use data to discipline the parameters in the model that control the strength of each channel. Without more formal analysis, it is hard to establish which combinations of parameter values are supported by the data and how these change the quantitative effects of demographics. These challenges remain or are exacerbated in larger OLG models.

4 Bayesian Estimation

We use Bayesian methods to estimate the model (2.1)-(2.3) with μ determined by the steady state of the OLG model in Section 3. Given the prior, the posterior distribution concentrates around parameter values supported by the data. The posterior credible intervals indicate the uncertainty in the estimates given the prior and data. Section 5 uses prior sensitivity analysis to distinguish the contribution of the data from that of the prior.

4.1 Estimation

4.1.1 Data

We use annual data from 1980-2013 for Japan. We focus on Japan because the macroeconomic and demographic changes there have been especially pronounced, giving the data the

		Type	Mean	Std Dev
$100(\beta^{-1} - 1)$	discount rate	Gamma	3.50	0.50
σ	intertemporal elasticity of substitution	Normal	0.35	0.15
δ	depreciation rate	Beta	0.08	0.02

Table 4.1: Prior for fixed structural parameters (β, σ, δ) .

best chance of informing the parameter estimates.

The time series we observe are GDP growth, working population growth, the share of workers in the population, the share of workers among the employed, the employment-population ratio, real wage growth, real interest rate, and labor share. In the data, we take the working population to be individuals from 15 to 64 years old, and take the retirement population to be individuals age 65 and above. To match our choice of 15 to 64-year-olds as workers, ω is calibrated to 0.98 so that workers in the model have an average working life of 50 years. The data are similar to what have been used in the literature to calibrate models quantifying the effect of demographics on interest rates. We also include data that directly inform us of the path of the time-varying parameters. Our results will show that without additional data on capital and life-cycle consumption, the calibrations in the literature can produce misleading results because the effects of demographics are not well-identified.

Several papers (e.g., [Attanasio et al. \(2007\)](#); [Kitao \(2017\)](#); [Sudo and Takizuka \(2018\)](#)) include the capital-output ratio in their calibration but do not directly target the interest rate. The results in these papers suggest that omitting interest rates may also lead to misleading results. The calibrations in [Attanasio et al. \(2007\)](#) and [Kitao \(2017\)](#) produce interest rates that are roughly 5 percentage points higher than the real return on government bonds,⁶ while [Sudo and Takizuka \(2018\)](#) find a real interest rate that is up to 2.5 percentage points higher than the natural interest rate estimated using the methodology of [Laubach and Williams \(2003\)](#). Since we wish to estimate the effects of demographics on interest rates, it is natural to include interest rates in the baseline estimation. Indeed, numerous papers calibrate the discount rate to match a given interest rate (e.g., [Ikeda and Saito \(2014\)](#); [Carvalho et al. \(2016\)](#); [Eggertsson et al. \(2019\)](#)). The results in Section 6 show that even after including data on both interest rates and the capital-output ratio, we still require life-cycle consumption data to identify the effects of demographics.

4.1.2 Prior

We focus on the prior for (β, σ, δ) , which we report in Table 4.1, and describe the rest of the prior in Appendix B. The prior for (β, σ, δ) is of particular interest because our results will

⁶[Attanasio et al. \(2007\)](#) argue that such an interest rate is comparable to the return on equity.

show that the data are especially uninformative about these three parameters, and the prior sensitivity analysis in Section 5 will show that the prior for these parameters is important for the estimated effects of demographics.

Our baseline prior is based on values used in existing calibrations of the Gertler (1999) OLG model. The prior for β implies a mean discount rate of 3.5%, which is close to existing calibrations of the same model. For σ , we choose a prior with mean 0.35 and standard deviation 0.15 to match the calibration of similar models used in papers based on the same model (e.g., Gertler (1999) and Fujiwara and Teranishi (2008) set $\sigma = 0.25$, while Carvalho et al. (2016) and Ferrero et al. (2019) set $\sigma = 0.50$). Finally, the prior for δ has a mean of 0.08 and standard deviation of 0.02, which allows for the range of calibrations in the literature (e.g., Gertler (1999), Fujiwara and Teranishi (2008), and Carvalho et al. (2016) set $\delta = 0.10$, while Kara and von Thadden (2016) set $\delta = 0.05$). We keep the prior independent across parameters, as is often done in the estimation of structural models. Figure B.1 in Appendix B compares the prior to these calibrations from the literature.

One could have picked other equally plausible priors. For example, the empirical literature has found a wide range of estimates for the intertemporal elasticity of substitution σ ranging from 0 to 2. Similarly, the measured depreciation rate depends on subjective choices about the measurement process, such as how much to aggregate across different types of capital (see, e.g., Fraumeni (1997) and Feenstra et al. (2015)). In Section 5, we show that the prior does impact posterior inference, which highlights the need for additional data.

4.1.3 Markov Chain Monte Carlo

To sample from the posterior, we use a Metropolis-within-Gibbs algorithm described in Appendix B. We take 2×10^5 burn-in draws, which we use to calibrate the proposal density. We then take 2.5×10^6 draws, keeping every 25th draw to save memory. To check for convergence, we partition the draws into four blocks and ensure that the posterior moments and marginals are similar across blocks.

4.2 Results

4.2.1 Structural Break and Long-run Means

Figure 4.1 plots the estimated long-run means $\mu(s_t)$ with 68% error bands. The structural break date t^* is estimated to be 1991. The long-run means before and after the break are distinct even after accounting for the error bands. Intuitively, μ is identified by 11 periods of data before t^* and 33 periods of data after t^* . In addition, the OLG model places cross-equation restrictions on the comovement of μ . Since variables such as the interest rate

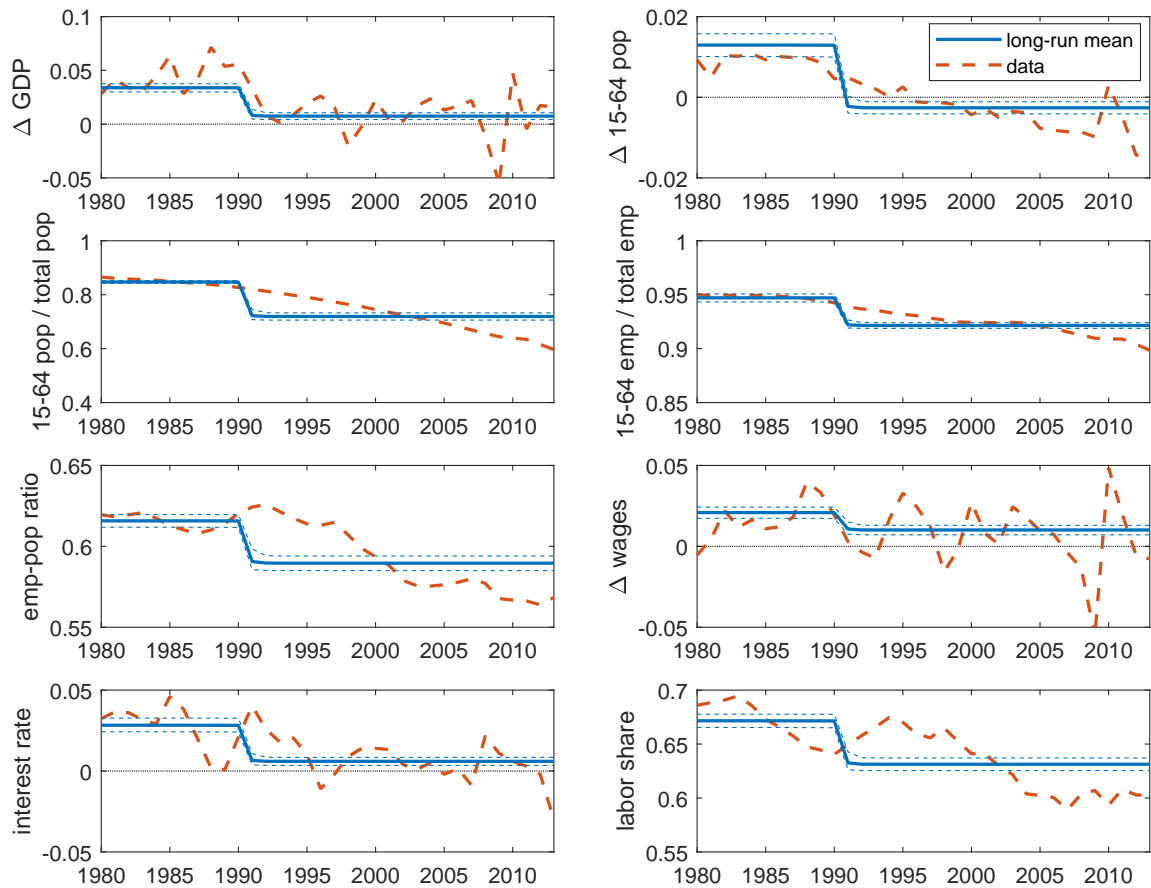


Figure 4.1: Estimated long-run means. **Blue lines:** median (solid) and 68% error bands (dashed) of long-run means; **Red dashed lines:** data.

and employment-population ratio are endogenous objects in the OLG model, the long-run means μ are therefore jointly identified by the data and the equilibrium conditions of the OLG model. The fact that the long-run means track the data indicates that parameter combinations exist that allow the model to fit the data well.

4.2.2 Natural Interest Rate and Counterfactuals

Natural Interest Rate. We define the natural interest rate to be the interest rate implied by the OLG model for given parameter values. In particular, define $\theta^* \equiv (\beta, \sigma, \delta)'$ and $\zeta \equiv (x, \nu, \alpha)'$. For any set of parameters $(\theta^*, \zeta, n, \gamma, \xi)$, we can compute the steady-state interest rate implied by the structural model, which we denote by $R(\theta^*, \zeta, n, \gamma, \xi)$. For each

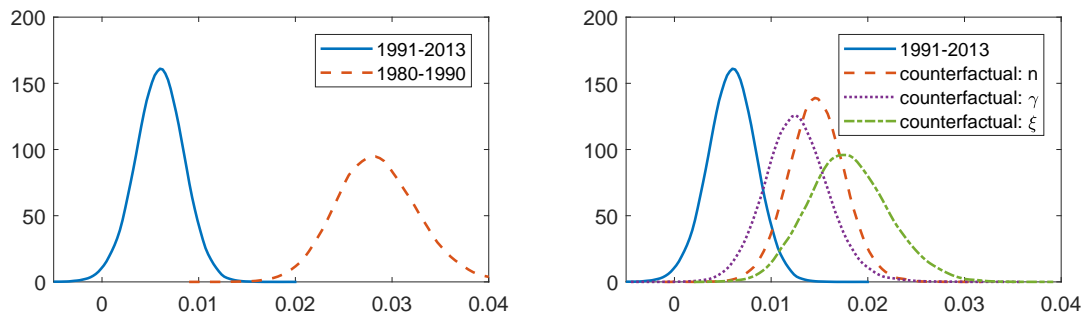


Figure 4.2: Posterior of natural interest rate and counterfactuals. **Left:** natural interest rate before and after structural break; **Right:** natural interest rate after structural break and counterfactuals.

period t , define the natural interest rate:

$$R_t \equiv R(\theta^*, \zeta(s_t), n(s_t), \gamma(s_t), \xi(s_t)).$$

This is similar to the existing literature extracting the natural interest rate (e.g., [Laubach and Williams \(2003\)](#); [Del Negro et al. \(2017\)](#); [Holston et al. \(2017\)](#)) using equilibrium conditions of a DSGE model. Here, we focus on the long-run average interest rates since we are concerned with the long-run trend in interest rates. The focus on the long run allows the OLG model to exclude frictions that are normally included in DSGE models used for extracting the natural interest rate at business cycle frequencies.

The left panel of Figure 4.2 shows that the posterior mean of the natural interest rate decreased from 2.85% to 0.60% after the structural break. The posteriors of the two interest rates are distinct, providing statistical evidence that the real interest rate has declined since the 1980s. The widths of the 68% credible intervals are 0.84 percentage points before the break and 0.49 percentage points after the break, widths that are comparable to the error bands that [Del Negro et al. \(2017\)](#) find for the low-frequency component of the natural interest rate in the United States. These credible intervals are wide enough to imply substantial uncertainty about the effects of a given path of interest rates.

Counterfactuals. We use the counterfactual natural interest rate to quantify the contribution of the population growth rate n , survival rate γ , and relative productivity of retirees

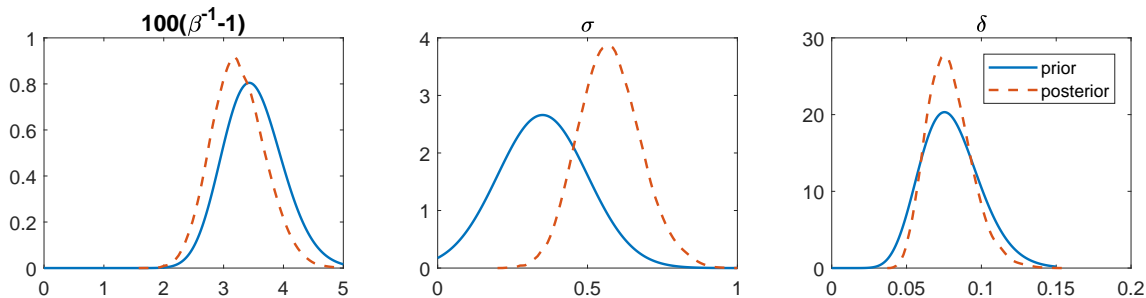


Figure 4.3: Marginal priors and posteriors of fixed parameters. **Solid blue line:** prior; **Dashed red line:** posterior.

ξ . In particular, we consider the counterfactual interest rates:

$$\begin{aligned}\widehat{R}_t^n &\equiv R(\theta^*, \zeta(s_t), \widehat{n}, \gamma(s_t), \xi(s_t)) \\ \widehat{R}_t^\gamma &\equiv R(\theta^*, \zeta(s_t), n(s_t), \widehat{\gamma}, \xi(s_t)) \\ \widehat{R}_t^\xi &\equiv R(\theta^*, \zeta(s_t), n(s_t), \gamma(s_t), \widehat{\xi}),\end{aligned}$$

where we pick the counterfactual parameter values $\widehat{\cdot}$ to be the median estimate for the parameter in 1980. These counterfactuals change one of the demographic parameters while keeping all other parameters identical. Given the Monte Carlo draws for each of the parameters, we can construct the posterior distributions for \widehat{R}_t^n , \widehat{R}_t^γ , and \widehat{R}_t^ξ .

The right panel of Figure 4.2 shows the posterior distribution of \widehat{R}_T^n , \widehat{R}_T^γ , and \widehat{R}_T^ξ , i.e., the counterfactual natural interest rate at the end of the sample had one of the three demographic parameters remained at its median value from before the structural break. They have means of 1.47%, 1.27% and 1.80% respectively, thus explaining between one-third and one-half of the decline in interest rates.⁷ The widths of the 68% credible intervals range are 0.58, 0.64, and 0.83 percentage points respectively.

4.2.3 Structural Parameters

The above estimates reveal substantial posterior uncertainty in the estimates for both the natural interest rate and the counterfactuals. To better understand the sources of uncertainty, we now turn to the estimates of the underlying structural parameters.

⁷The effects are not additive—the effect from changing more than one parameter is not the sum of the effects of changing each of those parameters individually. In addition, variation in the macroeconomic parameters also influence the change in the natural interest rate. For instance, the decline in labor share puts upward pressure on the interest rate, while the decline in productivity contributes to the decline in the interest rate.

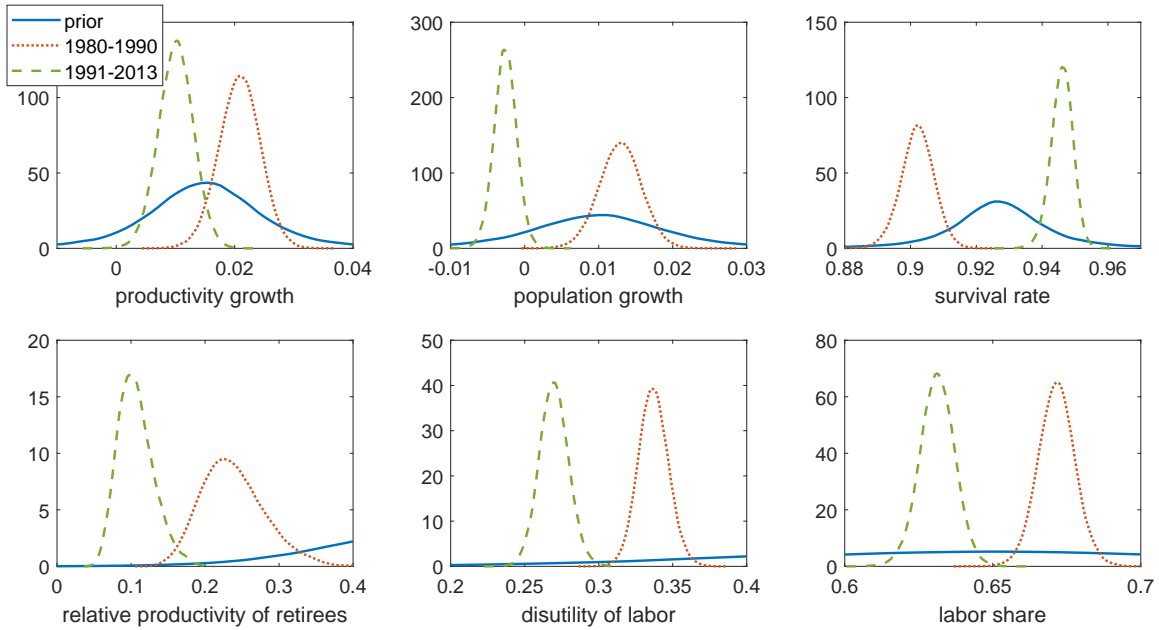


Figure 4.4: Marginal priors and posteriors of time-varying parameters. **Solid blue line:** prior; **Dotted red line:** posterior for parameter before structural break; **Dashed green line:** posterior for parameter after structural break.

Fixed Parameters. Figure 4.3 shows that the marginal posteriors for the fixed parameters (β, σ, δ) are relatively close to their priors, suggesting a dispersed likelihood. This is especially true for the discount rate β and capital depreciation rate δ . Intuitively, (β, σ, δ) are identified from two steady-state observations. The dispersed posteriors arise from the OLG model not placing substantial restrictions on the possible values of (β, σ, δ) individually given the estimated long-run means. Nevertheless, the likelihood is informative about the joint distribution of (β, σ, δ) , producing a posterior correlation between σ and δ of -0.66 .

One implication of these results is that one of the three fixed parameters could be well-identified given the other two parameters. However, fitting these parameters jointly to the data could yield a much wider range of possible values. Section 5 shows that the parameter values supported by the data can imply varied effects of demographics in the OLG model.

The dispersed marginal posteriors are not unique to the model and data here. For example, [Smets and Wouters \(2007\)](#) state that they calibrate δ because it is difficult to estimate with the data they use. In addition, they obtain relatively diffuse estimates for β and σ even though they use a longer time series of quarterly data and estimate the model using all frequencies in the data. While we lose information from business cycle fluctuations, we also have a more parsimonious model that has fewer parameters to be estimated.

Time-varying Parameters. In contrast to the fixed parameters, the time-varying parameters have marginal posteriors that are substantially tighter than their priors, as shown in Figure 4.4. Moreover, the posteriors for each of the parameters before and after the structural break are distinct from each other. The tightness of the posterior reflects the fact that the long-run averages of the data are tightly estimated relative to the prior.

Each time-varying parameter is closely connected to one of the time series. Working population growth is directly observed, while the survival rate can be inferred from the fraction of workers in the population given population growth. In the OLG model, productivity growth is equal to real wage growth as well as per capita GDP growth. The relative productivity of retirees is closely related to the fraction of workers among the employed, and the disutility of labor is similarly connected to the employment-population ratio.

In general, the estimates match the historical narrative of Japan’s economy from 1980 to 2013. The decline in productivity growth was a symptom of the lost decade. The decrease in population growth corresponds to the declining birth rates since the early 1970s, while the increase in survival rate matches the growth in life expectancy. The relative productivity of retirees is estimated to be lower after the break date due to the decrease in the fraction of the workforce below the age of 65. Even though the aggregate employment rate decreased, the employment rate by age group has increased since the 1980s, indicating a decrease in the disutility of labor. Finally the decline in the labor share has been documented by Karabarbounis and Neiman (2014) and others.

5 Prior Sensitivity

To formally establish that the data do not inform the estimated effects of demographics due to a lack of identification for (β, σ, δ) , we now analyze the sensitivity of the estimated counterfactual interest rates to changes in the prior of (β, σ, δ) using the relative entropy prior sensitivity (REPS) methodology introduced by Ho (2020). REPS explores a nonparametric set of priors that are close to the original prior in terms of relative entropy and finds bounds for the posterior results. In particular, we compute 68% *robust credible intervals*, defined as bounds that contain the equal-tailed 68% posterior credible interval for any prior in a given set of priors. In particular, we find the upper (lower) bound for the posterior 84% (16%) quantile given a set of priors that is close to the original prior. A wider robust credible interval relative to the original posterior credible interval indicates a greater dependence on the prior, which corresponds to the data being uninformative about underlying parameters and those parameters being important for the effect of demographics.

Our analysis is motivated by Section 3 showing theoretically that the three parameters

(β, σ, δ) have a role in determining the effect of demographics on interest rates and by Section 4 showing that these parameters are not well-identified by the data. Because the interest rate is determined jointly by all the equilibrium conditions, it is difficult to predict how these estimated effects change if we jointly change the prior of (β, σ, δ) . Changing the prior, as opposed to exogenously specifying new values of (β, σ, δ) , ensures that the new posterior continues to be disciplined by the data, respecting the joint likelihood of both (β, σ, δ) and the time-varying parameters. While it is impossible to reestimate the model for all possible priors on (β, σ, δ) , REPS provides bounds on our results for a nonparametric set of priors close to the original prior and identifies features of the prior that are important for the estimated effects.

5.1 Methodology

Denote the full vector of parameters by θ , the prior by π , and the posterior by p . As before, define $\theta^* \equiv (\beta, \sigma, \delta)'$, and let the marginal prior and posterior of θ^* be π^* and p^* , respectively. We are interested in how much the q th quantile of a function $\varphi(\theta)$ can change as we change the marginal prior π^* . In our setting, φ is the difference between the counterfactual and realized natural interest rate:

$$\varphi(\theta) \in \left\{ \widehat{R}_T^n - R_T, \widehat{R}_T^\gamma - R_T, \widehat{R}_T^\xi - R_T \right\} \quad (5.1)$$

These three definitions of φ measure the effects of population growth n , the survival rate γ , and the relative productivity of retirees ξ on the interest rate. Each choice of φ depends on the prior in a distinct way because n , γ , and ξ influence the interest rate in different ways.

We solve for an alternative prior $\tilde{\pi}^*$ for θ^* that minimizes the q th quantile:

$$\min_{\tilde{\pi}^*} \tilde{\varphi} \text{ s.t. } \mathcal{R} \geq \mathbb{E}_{\tilde{\pi}^*} \left[\frac{\tilde{\pi}^*(\theta^*)}{\pi^*(\theta^*)} \log \left(\frac{\tilde{\pi}^*(\theta^*)}{\pi^*(\theta^*)} \right) \right] \quad (5.2)$$

$$q = \mathbb{E}_{\tilde{p}} [\mathbf{1} \{ \varphi(\theta) \leq \tilde{\varphi} \}], \quad (5.3)$$

where \tilde{p} is the posterior arising from the worst-case marginal prior $\tilde{\pi}^*$, keeping the conditional prior of the remaining parameters and the likelihood the same. Solving (5.2)-(5.3) involves searching for a *worst-case* prior that minimizes the q th quantile of φ . Replacing the minimization with maximization instead yields the upper bound for the quantile. The constraint (5.3) states that $\tilde{\varphi}$ is the q th posterior quantile of φ . The constraint in (5.2) limits the relative entropy or Kullback–Leibler divergence of $\tilde{\pi}^*$ relative to π^* to be less than some constant $\mathcal{R} \geq 0$, restricting us to choose among priors that are statistically difficult

to distinguish from the original prior π^* . Problem (5.2)-(5.3) does not place parametric restrictions on the alternative prior. In particular, the feasible set of priors includes non-parametric distributions that could introduce correlations across parameters even though we started with a parametric and independent prior on θ^* . In our application here, we seek bounds for the 68% equal-tailed credible interval for effects of n , γ , and ξ . We therefore take $q \in \{0.16, 0.84\}$ in (5.3) with minimization replaced with maximization for $q = 0.84$.

To gauge the size of \mathcal{R} , Ho (2020) recommends the rule:

$$\mathcal{R} = 1.6^d r \frac{\pi(\mu_p)}{\pi(\mu_\pi)} \sqrt{\frac{|\Sigma_\ell|}{|\Sigma_\pi|}} \quad (5.4)$$

$$\Sigma_\ell \equiv \left(\Sigma_p^{-1} - \Sigma_\pi^{-1}\right)^{-1}, \quad (5.5)$$

where μ_p and μ_π are the posterior and prior means for θ^* , respectively; Σ_p and Σ_π are the posterior and prior variances for θ^* , respectively; and d is the dimensionality of θ^* . Σ_ℓ is a measure of how dispersed the likelihood is. The choice of r determines how large the relative entropy is, with $r < 0.05$ corresponding to small levels of relative entropy.⁸ We shall pick r to correspond to a one to two standard deviation change in the quantiles on average.

We implement the above computations using the sequential Monte Carlo and numerical approximations described in Ho (2020). See Appendix C for details.

5.2 Results

5.2.1 Robust Credible Intervals

Figure 5.1 shows that the robust credible intervals are wide, indicating that a small change in the prior for (β, σ, δ) can produce a large change in the posterior estimates for the effects of demographics. In particular, with $r = 0.005$, the robust credible intervals are between 3.7 and 5.5 posterior standard deviations wider than the corresponding original credible intervals. The quantiles shift by an economically significant amount of up to 1.5 percentage points.⁹ Relative to the $r = 0.05$ benchmark, $r = 0.005$ corresponds to a very small set of priors, which shows that the estimated effects of demographics are very sensitive to the prior.

⁸Suppose ψ is a linear combination of θ^* , and we estimate $\theta^* \equiv (\beta, \sigma, \delta)'$ from a large number of observations $\theta^* + \varepsilon_t$, with $\varepsilon_t \sim \mathcal{N}(0, \Omega)$ where Ω is known. Then $r = 0.05$ and $r = 0.005$ would correspond to changes in the quantile of approximately 0.4 and 0.1 posterior standard deviations, respectively. Ho (2020) shows that if the variance of ε_t is equal to the prior variance of θ^* , then ten observations are sufficient for this to be a good approximation.

⁹Smets and Wouters (2007) estimate that the standard deviation of the monetary policy shock is 0.2 percentage points and a one standard deviation monetary policy shock leads to a 3% decline in output.

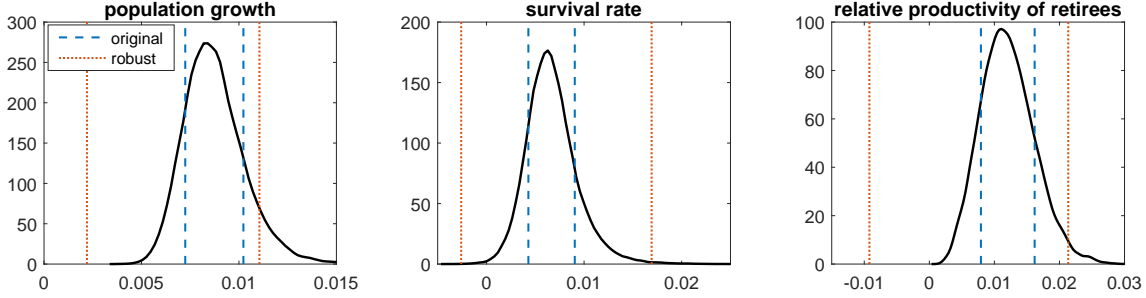


Figure 5.1: Robust credible intervals for difference between counterfactual and realized natural interest rate with $r = 0.005$. **Dashed blue line:** original 68% credible interval; **Dotted red line:** robust 68% credible interval; **Solid black line:** posterior density.

There are two reasons for the high degree of prior sensitivity. Firstly, as suggested in Figure 4.3, the likelihood for (β, σ, δ) is dispersed. Defining Σ_π and Σ_ℓ as in (5.5), we find $\sqrt{|\Sigma_\ell|/|\Sigma_\pi|} = 2.2$, which indicates that the likelihood is more dispersed than the prior. Secondly, (β, σ, δ) are strong predictors of the demographic effects $\widehat{R}_T^n - R_T$, $\widehat{R}_T^\gamma - R_T$, and $\widehat{R}_T^\xi - R_T$. Quadratic regressions of $\widehat{R}_T^n - R_T$, $\widehat{R}_T^\gamma - R_T$, and $\widehat{R}_T^\xi - R_T$ on $(\beta^{-1}, \sigma, \delta)$ yield R -squares of 0.67, 0.74, and 0.93, respectively. Changing the prior for (β, σ, δ) thus results in a large change in the posterior for both (β, σ, δ) and the effects of demographics.

These results show that the data are uninformative about structural parameters that are important for determining the effects of demographics on interest rates in the OLG model. Since the dispersion of the likelihood is a feature of the model and data only, any other function of the model parameters that is well-predicted by (β, σ, δ) would also depend heavily on the prior.

In Appendix D, we also reestimate the model using an alternative parametric prior that is motivated by the results here. The posterior estimates change substantially under the new prior, corroborating the results here.

5.2.2 Worst-case Posteriors

To understand how the estimates depend on each part of the prior, we compare the original and worst-case posteriors for the 16% and 84% quantiles, shown in Figures 5.2 and 5.3, respectively. The worst-case distortions for the 16% (84%) quantiles place greater weight on parameter values that imply smaller (larger) effects of demographics. These distortions indicate the relative importance of the savings composition, within-group savings, capital demand, and general equilibrium effects, as well as their sensitivity to changes in (β, σ, δ) .¹⁰

¹⁰While these are not necessarily the only distortions that shift the effects in the desired direction, each worst-case prior provides one example of an alternative prior that has a large impact on the estimated effects

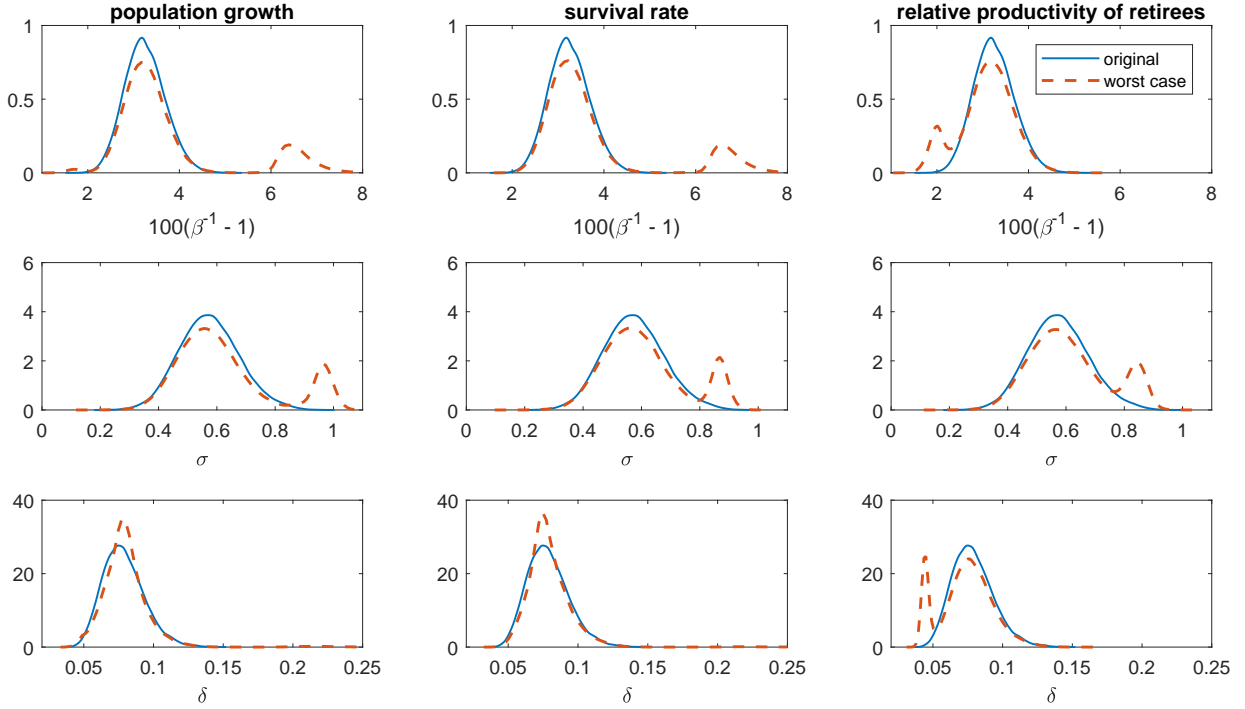


Figure 5.2: Original and worst-case posteriors generating a one posterior standard deviation decrease in the 16% quantile. **Left:** $\hat{R}_T^n - R_T$; **Center:** $\hat{R}_T^\gamma - R_T$; **Right:** $\hat{R}_T^\xi - R_T$.

The distortions are asymmetric and involve joint changes in the prior for (β, σ, δ) . These features emphasize that it is difficult to predict how one’s choice of prior may be affecting one’s estimates. Moreover, the nonlinear dependence of the estimated effects on the prior imply that reestimating the model with an ad hoc alternative prior may not give a complete picture of the sensitivity of estimates to the prior. Similarly, our results show that even though a wide range of calibrations could be consistent with the data, it is hard to know ex ante how the choice among these calibrations may affect the model’s quantitative predictions.

Population Growth. The worst-case posteriors corresponding to the effect of population growth involve especially large distortions to the marginal of the intertemporal elasticity of substitution σ , placing more weight on large (small) values of σ to decrease (increase) the estimated effect of population. A change in the working population growth n affects the interest rate through the savings composition, capital demand, and general equilibrium channels. Increasing σ reduces the worker’s incentive to save for retirement, which decreases the savings composition effect, thus strengthening the net effect of population growth. However, a large σ also dampens the effect of population growth by decreasing the general equilibrium

of demographics.

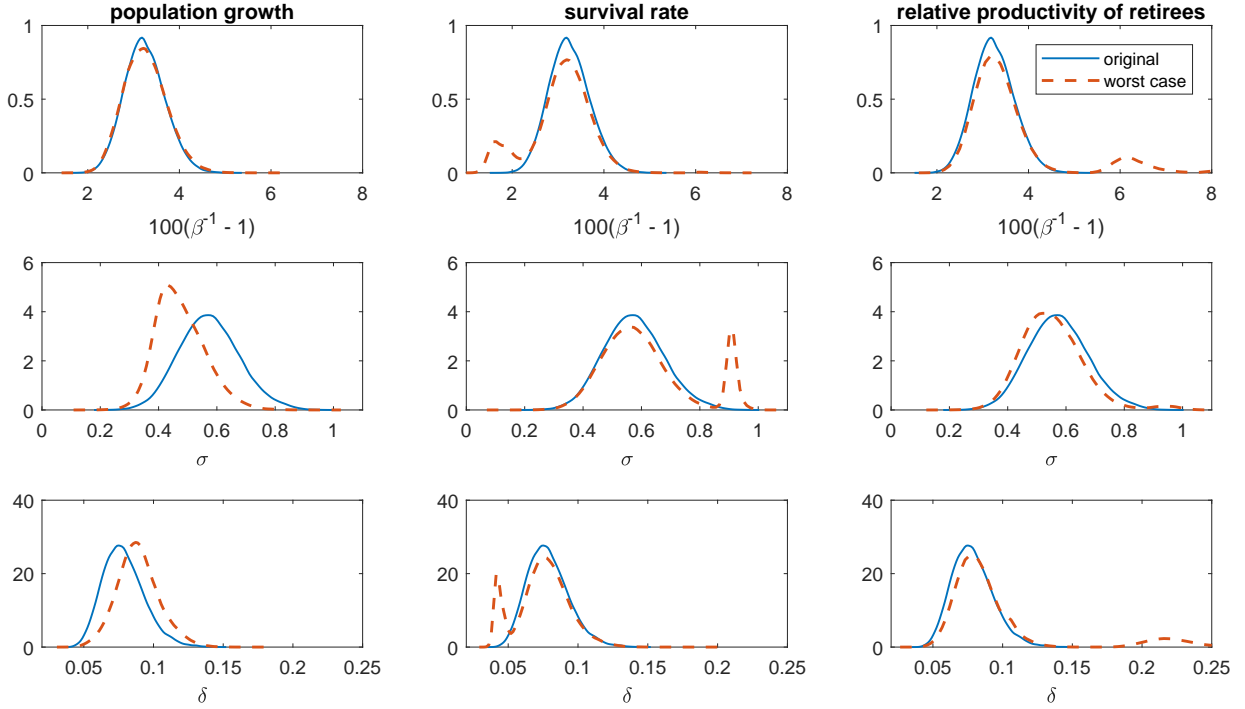


Figure 5.3: Original and worst-case posteriors generating a one posterior standard deviation increase in the 84% quantile. **Left:** $\hat{R}_T^n - R_T$; **Center:** $\hat{R}_T^\gamma - R_T$; **Right:** $\hat{R}_T^\xi - R_T$.

effect. That the worst-case posterior concentrates around large values of σ to decrease the 16% quantile suggests that the general equilibrium effect is especially sensitive to σ .

The distortions for β^{-1} are asymmetric. To decrease the 16% quantile, the worst-case distortion places more weight on large values of β^{-1} . In contrast, the marginal for β^{-1} is relatively unchanged for the 84% quantile. The asymmetry highlights the nonlinearity in the mapping from (β, σ, δ) to the effect of population growth. For example, the effect of β on the marginal propensities to consume (3.14) and (3.15) is amplified by a larger value of σ . As a result, changing β has a smaller effect when accompanied by a decrease in σ for the 84% quantile worst-case distortion.

For both the 16% and 84% quantiles, the distortion of δ is relatively small, suggesting that the capital demand channel is less important than the effects originating from household savings decisions. The small increase in δ for the 84% quantile arises due to the negative correlation of -0.66 between σ and δ .

Survival Rate. The worst-case posterior for the 16% quantile for the effect of the survival rate is similar to that of population growth, but the distortions for the 84% quantile differ, placing greater weight on small values of β^{-1} , large values of σ , and small values of δ .

Decreasing β^{-1} increases the within-group saving effect. Firstly, decreasing β^{-1} amplifies the change in the effective discount rate $\beta\gamma$ arising from a change in the survival rate γ . Retirees thus increase savings more in response to an increase in γ . Secondly, decreasing β^{-1} increases the workers' incentive to smooth consumption into retirement, causing their marginal propensity to consume to decline more in response to an increase in γ .

On its own, increasing σ dampens the effect of the survival rate on interest rates. However, increasing σ also amplifies the effects of changing β . That the worst-case distortions for the 84% quantile involves an increase in σ suggests that the interaction with β is more important than the direct effect of σ . The nonlinearity emphasizes the importance of studying the joint distribution and effect of (β, σ, δ) rather than analyzing each parameter independently.

The worst-case distortions for the 84% quantile result in a new posterior mode for δ around 0.05. The decrease in δ dampens the capital demand channel. This is consistent with an increase in the effect of the survival rate γ because the increase in worker labor supply more than offsets the increased fraction of retirees in the economy. As a result, average labor productivity rises, which increases the marginal product of capital. Decreasing δ reduces the resulting increase in capital demand, thus strengthening the effect of γ on interest rates.

Relative Productivity of Retirees. The worst-case posterior for the 16% quantile for the effect of the relative productivity ξ is similar to that of the 84% quantile for the effect of the survival rate γ . These distortions increase the effect of γ but decrease the effect of ξ . The contrasting dependence on the prior shows that the estimates of both the absolute and relative magnitudes of the different effects of demographics are sensitive to the prior.

The worst-case posterior indicates the importance of the capital demand channel for determining the effect of the relative productivity of retirees. In addition, it shows that the capital demand effect serves to decrease interest rates in response to a decrease in ξ even though it increases interest rates in response to a decrease in γ .

Decreasing β^{-1} increases the incentive for workers to increase their labor supply in response to a decrease in ξ , hence decreasing the downward pressure on interest rates from the capital demand channel. This dominates the increase in the within-group savings effect arising from the reduction of β^{-1} . The decrease in δ also dampens the capital demand effect. On the other hand, the increase in σ decreases the incentive for workers to increase their labor supply but also decreases the within-group savings effect. The effect of σ on the latter channel dominates. The worst-case posterior for the 84% quantile also suggests an important role for the capital demand channel as more weight is put on large values of β^{-1} and δ .

The fact that the marginal distortions push the various effects of ξ in different directions emphasizes the nonlinear dependence on (β, σ, δ) . The analysis here accounts for this non-

linearity while remaining disciplined by the likelihood. In particular, the endogenous labor supply response to changes in ξ and γ are disciplined by the data on the fraction of workers among the employed and the employment-population ratio.

5.3 Implications for Estimation and Calibration

Additional Data. The importance of (β, σ, δ) informs us of the data needed to tighten our estimates of the effects of demographics. Aggregate data on capital and investment, such as the capital-output ratio, can help identify σ and δ by providing the response of savings and investment to interest rate changes. Data on consumption and saving over the life cycle can help to better identify β and σ . Such data provides information on how households respond to their life-cycle trajectory of wages, which helps to quantify the savings composition effect. The evolution of these patterns over time provides information on the exact response of households' marginal propensities to consume to changes in life expectancy or future wages, which in turn helps to identify the strength of the within-group savings effect.

An alternative approach would be to include a longer time series or use a panel of countries. Increasing the length of the time series would increase the effective number of steady-state observations if we observe structural breaks prior to 1980. However, constructing a sufficiently long time series to narrow estimates on (β, σ, δ) is challenging given the relative scarcity of data before 1950. Including a larger panel of countries would similarly allow us to observe more regimes. For such an exercise to narrow our estimates of (β, σ, δ) , we require that (β, σ, δ) are identical or relatively similar across countries.

Calibration. The results here also have implications for the calibration of OLG models. In particular, when calibrating OLG models, one should consider a range of overidentifying restrictions to pin down (β, σ, δ) . For example, while it is possible to pin down β conditional on σ and δ , given a particular interest rate, our analysis has shown that the data support a range of possible interest rates. Once one acknowledges the uncertainty in σ and δ , one has an even wider range of possible combinations. Our results show that these parameter combinations can produce different conclusions. Additional data is required to determine the appropriate parameter values with more precision.

Extensions and Other Questions. The dependence of our estimates on the prior for (β, σ, δ) is not unique to the model considered here. The savings composition, within-group savings, capital demand, and general equilibrium channels are present in many OLG models used to quantify the effects of demographic changes on the interest rate. Additional channels such as financial frictions (Ikeda and Saito (2014); Wong (2018)), public pension schemes

(Muto et al. (2016); Sudo and Takizuka (2018)), or international capital flows (Brooks (2003); Attanasio et al. (2006)) also depend on (β, σ, δ) . These extensions will only tighten the likelihood of (β, σ, δ) if they place additional cross-equation restrictions that discipline the range of plausible values for (β, σ, δ) . In such cases, it would be important to understand the sensitivity of these restrictions to the details of the extensions and to show empirical support for the relevant parts of the model.

The identification of (β, σ, δ) is also important for quantifying other connections between demographics and the economy. For example, the consequences of public pension policy (e.g., Imrohoroglu et al. (1995); Attanasio et al. (2007); Kitao (2017)) or the role of demographics in the transmission of monetary policy shocks (e.g., Fujiwara and Teranishi (2008); Wong (2018)) also depend on (β, σ, δ) . The need for more data extends to quantitative analysis of such questions.

6 The Role of Additional Data

6.1 Capital-output Ratio

We now show how data on capital improves identification. First, we reestimate the model using the capital-output ratio along with the original eight time series with the original prior from Section 4. We then repeat the REPS analysis. Both exercises show that the capital-output ratio adds substantial information about the effects of demographics.

6.1.1 Bayesian Estimation

Table 6.1 shows that the inclusion of the capital-output ratio tightens the posterior estimate of the structural parameters. The inclusion of data on the capital-output ratio decreases the posterior standard deviation of δ by nearly five times, while the standard deviation of the intertemporal elasticity of substitution σ is halved. The standard deviation of the discount rate β^{-1} is roughly unchanged. The posterior means of δ and σ change by over two posterior standard deviations, while the mean of the discount rate is relatively unaffected by the inclusion of the capital-output ratio.

The changes in the posterior indicate that the capital-output ratio adds substantial information about the values of σ and δ but not β . Equation (3.16) shows that given the interest rate $R - 1$ and the labor share α , the capital-output ratio pins down the capital depreciation rate δ . Similarly, the marginal propensities to consume (3.14)-(3.15) show that the responsiveness of savings to interest rates is determined by σ . Data on interest rates measure the price of savings and capital, while the capital-output ratio measures the re-

		Original			With K/Y Ratio		
		mean	sd	68% C.I.	mean	sd	68% C.I.
Structural Parameters							
$100(\beta^{-1} - 1)$	disc. rate	3.245	0.442	(2.804, 3.690)	3.443	0.427	(3.013, 3.873)
σ	IES	0.574	0.103	(0.471, 0.677)	0.282	0.053	(0.230, 0.335)
100δ	dep. rate	7.861	1.485	(6.393, 9.317)	11.394	0.300	(11.098, 11.692)
Effects of Demographics							
$100(\widehat{R}_T^n - R_T)$	pop. growth	0.875	0.159	(0.725, 1.023)	1.480	0.421	(1.122, 1.817)
$100(\widehat{R}_T^\gamma - R_T)$	surv. rate	0.671	0.266	(0.429, 0.905)	2.140	0.808	(1.401, 2.858)
$100(\widehat{R}_T^\xi - R_T)$	rel. prod.	1.206	0.412	(0.793, 1.618)	1.423	0.202	(1.240, 1.604)

Table 6.1: Posterior statistics for fixed structural parameters and effects of demographics, including capital-output ratio in data.

sponse of the quantities. As a result, fitting the model to both of these variables gives us more precise estimates of σ and δ .

With the additional data, the posterior standard deviation of the effect of the relative productivity of retirees ξ is halved, but the posterior standard deviation of the effects of population growth n and the survival rate γ are approximately tripled. The estimated effect of ξ is more precise because of the tighter estimate of δ . In particular, Figures 5.2 and 5.3 show that the worst-case posteriors for the effect of ξ involve large distortions in the marginal of δ , suggesting that the estimate of δ is important for the estimated effect of ξ . The marginal of δ is distorted less in the worst-case posteriors corresponding to the effects of n and γ . The posterior for the effects of n and γ become more dispersed because of the new estimates for the structural parameters. The decrease in σ and increase in δ make savings and investment more responsive to changes in the interest rate, which increases the general equilibrium channel. The increased general equilibrium channel amplifies the effects of demographics on interest rates, increasing the difference between the net effect of small and large partial equilibrium effects. This amplification outweighs the effect of the increased precision in the estimates.

Besides changing the posterior standard deviations, the introduction of the capital-output ratio data also causes the posterior means of the effects to shift by up to 1.5 percentage points. The difference in estimates shows that the choice of data can alter the conclusions from quantitative exercises, such as policy analysis or forecasting that depend on the interaction between demographics and interest rates.

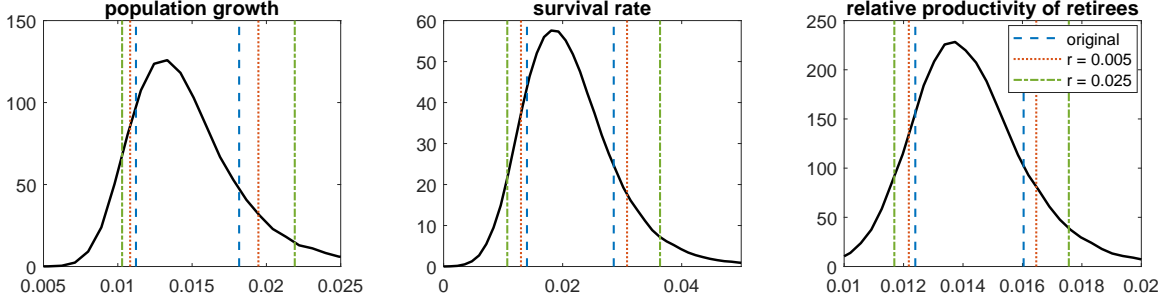


Figure 6.1: Robust credible intervals for difference between counterfactual and realized natural interest rate with $r \in \{0.005, 0.025\}$ in estimation with capital-output ratio. **Dashed blue line:** original 68% credible interval; **Dotted red line:** robust 68% credible interval with $r = 0.005$; **Dash-dot green line:** robust 68% credible interval with $r = 0.025$; **Solid black line:** posterior density.

6.1.2 Prior Sensitivity

Figure 6.1 shows that the inclusion of the capital-output ratio results in robust credible intervals that are much closer to the original credible intervals, which indicates that the data are more informative in this estimation. For comparison, we consider the robust credible intervals with $r = 0.005$ with and without the capital-output ratio. Table 6.2 reports the statistic:

$$\rho(\psi) \equiv \frac{\chi_r(\varphi)}{\chi_p(\varphi)} - 1,$$

where χ_p and χ_r are the widths of the original and robust 68% credible intervals, respectively, for each definition of φ in (5.1). The statistic ρ quantifies the sensitivity of the credible interval to the prior by measuring how much wider the robust credible interval is compared with the original credible interval. Including the capital-output ratio decreases ρ from 2.0, 3.1, and 2.7 to approximately 0.2. Even though the credible intervals for the estimated effects of population growth n and the survival rate γ are wider than in the estimation without the capital-output ratio, the REPS analysis shows that the prior now plays a smaller role in determining these estimates. These results support the earlier claim that these wider posterior credible intervals arise due to the new estimates for σ and δ .

The amount by which the robust credible intervals tighten is consistent with the worst-case distortions in Figures 5.2 and 5.3. In particular, we define ρ_k and ρ_o to be the values of ρ for the estimation with and without the capital-output ratio, respectively, and take ρ_o/ρ_k as a measure of how much additional information the capital-output ratio provides for a given φ . This ratio is largest for the effect of the relative productivity of retirees ξ . The worst-case posteriors for the effect of ξ involve largest distortions in the distribution of δ , suggesting that the importance of information about δ for the estimated effects of demographics is greatest

		68% C.I. (χ_p)		Rob. C.I. (χ_r)		Relative Change (ρ)		
		orig.	K/Y	orig.	K/Y	orig.	K/Y	ratio ($\frac{\rho_o}{\rho_k}$)
Effects of Demographics								
100	$(\widehat{R}_T^n - R_T)$ pop. growth	0.298	0.694	0.885	0.863	1.965	0.243	8.081
100	$(\widehat{R}_T^\gamma - R_T)$ surv. rate	0.475	1.458	1.951	1.782	3.105	0.223	13.954
100	$(\widehat{R}_T^\xi - R_T)$ rel. prod.	0.826	0.365	3.056	0.429	2.702	0.177	15.307

Table 6.2: Widths of 68% credible intervals and robust credible intervals with $r = 0.005$ for estimations with and without capital-output ratio. **Left panel:** width of 68% credible interval; **Center panel:** width of robust 68% credible interval; **Right panel:** difference between widths of credible interval and robust credible interval normalized by width of 68% credible interval.

		Pop. Growth (n)		Surv. Rate (γ)		Rel. Prod. (ξ)	
		16% qtl	84% qtl	16% qtl	84% qtl	16% qtl	84% qtl
Structural Parameters							
100	$(\beta^{-1} - 1)$ disc. rate	0.082	-0.039	0.701	-0.073	1.162	-0.254
	σ IES	1.364	-0.165	0.830	-0.182	0.676	-0.306
	100 δ depr. rate	2.029	-0.059	0.109	-0.049	-0.203	0.017

Table 6.3: Difference between worst-case and original posterior mean of $(\beta^{-1}, \sigma, \delta)$, normalized by original posterior standard deviation. Worst-case posteriors correspond to a one-half posterior standard deviation change in quantile.

for the effect of ξ . The change in ρ is greatest for the effect of ξ since the capital-output ratio tightens the estimate for δ more than β or σ . On the other hand, the worst-case posteriors for the effect of population growth n involve minimal distortions to the distribution of δ , consistent with ρ_o/ρ_k being the smallest for this effect.

Nevertheless, the robust credible intervals remain wide relative to the rule of thumb in Ho (2020). Figure 6.1 shows that with $r = 0.025$, which is half the $r = 0.05$ benchmark, the robust credible intervals are between 1.1 and 1.4 posterior standard deviations wider than the corresponding original credible intervals. The change is about 1.5 times as large as what one would expect from the $r = 0.05$ benchmark (see footnote 8 for details). These results show that even though the capital-output ratio greatly improves estimates, the augmented data remain relatively uninformative about the effects of demographics.

Table 6.3 shows that in the estimation with the capital-output ratio, the worst-case

distortions are concentrated in the prior of (β, σ) .¹¹ In contrast, Figures 5.2 and 5.3 show that without the capital-output ratio, the worst-case distortions in δ are relatively more important, especially for the estimated effect of the relative productivity ξ . The worst-case distortions reflect that the capital-output ratio data informs the estimate of δ , making it more difficult to change the estimated effects by changing the prior of δ . It is therefore important to incorporate additional data that can further discipline β and σ .

6.2 Consumption over the Life Cycle

One natural way to obtain information on β and σ is to use data on life-cycle patterns in consumption and savings. However, direct observations of such data are unavailable.¹² Instead of reestimating the model with missing data, we indirectly study how data on life-cycle consumption would tighten our estimates if it were available.

6.2.1 Methodology

Suppose we wish to understand the role of some data $\{y_t^*\}$, with steady states $\mu^*(s_t)$ in regime s_t . For each Monte Carlo draw j from the posterior, we compute the implied model steady states $\mu_j^*(s_t)$. Taking the Monte Carlo draws as observations, we run quadratic regressions of the parameters and effects of demographics on $(\mu^*(0), \mu^*(1))$ and use the R^2 to measure how much data on $\{y_t^*\}$ would improve identification of the parameters and effects.

Intuitively, the exercise asks how well one can predict the objects of interest if one observed $\mu^*(s_t)$. Assuming we are able to obtain tight estimates of $\mu^*(s_t)$ from $\{y_t^*\}$, a large R^2 is suggestive evidence that the data would improve identification substantially. Conversely, an R^2 that is close to zero implies that the data provides no information about the values of the parameters or effects. We use a quadratic regression as an approximation for the mapping between the objects of interest and the observables. Since we use the Monte Carlo draws from the existing posterior, our results are conditional on the existing likelihood, measuring the information contained in hypothetical data $\{y_t^*\}$ on top of the information contained in the data used for the existing estimation.

¹¹The only exception is the worst-case distortions corresponding to the 16% quantile for the effect of population growth n . However, since Figure 6.1 shows that this quantile is especially robust to changes in the prior, we focus our discussion on the other five worst-case priors.

¹²To the best of my knowledge, the closest available data is from the Family Income and Expenditure Survey, which reports average consumption per household, aggregated by head of household's age from 2003. Inferring worker and retiree consumption consistent with the model requires accounting for household composition in both the demographic-level and aggregate data. Estimating the model with missing observations is computationally much costlier, as we are no longer able to analytically integrate out the VAR parameters in (2.2) and (2.3). Given these challenges, we leave the exercise to future work.

		$(c^w, c^r/\psi)$	$\psi c^w/c^r$
Structural Parameters			
$100(\beta^{-1} - 1)$	disc. rate	0.215	0.118
σ	IES	0.439	0.046
100δ	dep. rate	0.354	0.070
Effects of Demographics			
$100(\widehat{R}_T^n - R_T)$	pop. growth	0.353	0.042
$100(\widehat{R}_T^\gamma - R_T)$	surv. rate	0.588	0.280
$100(\widehat{R}_T^\xi - R_T)$	rel. prod.	0.808	0.738

Table 6.4: R^2 from quadratic regression of parameters and effects on retiree and worker consumption.

We consider two alternatives for μ^* . First, we take $\mu^* = (c^w, c^r/\psi)$, which is the steady-state average consumption of retirees and workers respectively, scaled by output. Next, we take $\mu^* = \psi c^w/c^r$, which is the steady-state ratio of average retiree consumption to average worker consumption. The former requires consumption levels for each group, while the latter requires only the change in consumption over the life cycle. We use Monte Carlo draws for the posterior of the estimation that includes the capital-output ratio from Section 6.1.

6.2.2 Results

Table 6.4 shows that data on consumption levels $(c^w, c^r/\psi)$ over the life cycle can provide substantial information about the effects of demographics on interest rates, with R^2 s of 0.35, 0.59, and 0.81 for population growth, the survival rate, and the relative productivity, respectively. For comparison, the respective R^2 s when we take $\mu^* = k$ and use Monte Carlo draws from the estimation without the capital-output ratio are 0.43, 0.49, and 0.85.

On the other hand, data on the relative change in consumption levels $\psi c^w/c^r$ would be less informative about the effects of changes in population growth and the survival rate on interest rates, with R^2 s of 0.04 and 0.28 respectively. The low R^2 arises from reduced information about the intertemporal elasticity of substitution σ and depreciation rate δ , which are reflected in reductions in R^2 s from 0.44 and 0.35, respectively, to 0.05 and 0.07. Table 6.3 shows that σ and δ are particularly important relative to β for pinning down the effect of changes in population growth. Nevertheless, $\psi c^w/c^r$ is informative about the effects of changes in relative productivity with R^2 s of 0.74.

These results highlight the potential for life-cycle consumption data to identify the effects

of demographics on interest rates by informing estimates of the discount rate, intertemporal elasticity of substitution, and depreciation rate. Consumption levels are particularly informative relative to only knowing changes in consumption over the life cycle.

7 Conclusion

The aging populations and falling interest rates in developed countries have led to a large literature seeking to quantify the effects of demographics on interest rates through the lens of OLG models. The analysis here shows that these results may be fragile. In particular, we have shown that the estimated effects of demographics depend crucially on the discount rate, intertemporal elasticity of substitution, and capital depreciation rate. Without the appropriate data, these parameters are not well-identified. As a result, a large set of parameter values can be justified, leading to a wide range of possible estimated effects and explaining the differences across estimates in the literature. Including aggregate capital and life-cycle consumption data to discipline these parameters can help produce more accurate and precise estimates. These insights extend to more complicated models and other related empirical questions.

In terms of methodology, this paper makes two contributions. Firstly, it introduces an econometric framework that disentangles secular changes from high-frequency fluctuations in a way that uses a structural economic model to discipline only the secular changes. Secondly, the REPS analysis illustrates how prior sensitivity analysis reveals parts of a model that are not well-identified and matter for our objects of interest. Importantly, by avoiding parametric restrictions on the prior, our analysis accounts for potentially large joint changes in the parameters that could alter results without substantially worsening the fit with the data. The analysis informs us about the data needed to sharpen the estimates and suggests the appropriate moments for calibration.

References

- Aksoy, Y., H. S. Basso, R. P. Smith, and T. Grasl (2019). Demographic Structure and Macroeconomic Trends. *American Economic Journal: Macroeconomics* 11(1), 193–222.
- Attanasio, O., S. Kitao, and G. L. Violante (2007). Global Demographic Trends and Social Security Reform. *Journal of Monetary Economics* 54(1), 144–198.
- Attanasio, O. P., S. Kitao, and G. L. Violante (2006). Quantifying the Effects of the Demographic Transition in Developing Economies. *Advances in Macroeconomics* 6(1).
- Blanchard, O. J. (1985). Debt, Deficits, and Finite Horizons. *Journal of Political Economy* 93(2), 223–247.
- Brooks, R. (2002). Asset-Market Effects of the Baby Boom and Social-Security Reform. *American Economic Review* 92(2), 402–406.
- Brooks, R. (2003). Population Aging and Global Capital Flows in a Parallel Universe. *IMF Economic Review* 50, 200–221.
- Canova, F. and L. Sala (2009). Back to Square One: Identification Issues in DSGE Models. *Journal of Monetary Economics* 56(4), 431–449.
- Carvalho, C., A. Ferrero, and F. Nechio (2016). Demographics and Real Interest Rates : Inspecting the Mechanism. *European Economic Review* 88, 208–226.
- Christiano, L. J., M. Trabandt, and K. Walentin (2010). DSGE Models for Monetary Policy Analysis. In *Handbook of Monetary Economics*, Volume 3, pp. 285–367. Elsevier Ltd.
- Del Negro, M., D. Giannone, M. P. Giannoni, and A. Tambalotti (2017). Safety, Liquidity, and the Natural Rate of Interest. *Brookings Papers on Economic Activity* (Spring), 235–316.
- Del Negro, M. and F. Schorfheide (2004). Priors from General Equilibrium Models for VARs. *International Economic Review* 45(2), 643–673.
- Eggertsson, G. B., N. R. Mehrotra, and J. A. Robbins (2019). A Model of Secular Stagnation: Theory and Quantitative Evaluation. *American Economic Journal: Macroeconomics* 11(1), 1–48.
- Feenstra, R. C., R. Inklaar, and M. P. Timmer (2015). The Next Generation of the Penn World Table. *American Economic Review* 105(10), 3150–3182.

- Ferrero, A. (2010). A Structural Decomposition of the U.S. Trade Balance: Productivity, Demographics and Fiscal Policy. *Journal of Monetary Economics* 57(4), 478–490.
- Ferrero, G., M. Gross, and S. Neri (2019). On Secular Stagnation and Low Interest Rates: Demography Matters. *International Finance* 22(3), 262–278.
- Fraumeni, B. M. (1997). The Measurement of Depreciation in the U.S. National Income and Product Accounts. *Survey of Current Business* 77, 7–23.
- Fujita, S. and I. Fukiwara (2016). Declining Trends in the Real Interest Rate and Inflation: Role of Aging. Federal Reserve Bank of Philadelphia Working Paper Series 16-29.
- Fujiwara, I. and Y. Teranishi (2008). A Dynamic New Keynesian Life-Cycle Model: Societal Aging, Demographics, and Monetary Policy. *Journal of Economic Dynamics and Control* 32(8), 2507–2511.
- Gagnon, E., B. K. Johansson, and D. Lopez-Salido (2016). Understanding the New Normal: The Role of Demographics. Finance and Economics Discussion Series 2016-080, Washington: Board of Governors of the Federal Reserve System.
- Gertler, M. (1999). Government Debt and Social Security in a Life-Cycle Economy. In *Carnegie-Rochester Conference Series on Public Policy*, Volume 50, pp. 61–110. Elsevier.
- Herbst, E. and F. Schorfheide (2014). Sequential Monte Carlo Sampling for DSGE Models. *Journal of Applied Econometrics* 29, 1073–1098.
- Ho, P. (2020). Global Robust Bayesian Analysis in Large Models. Federal Reserve Bank of Richmond Working Paper 20-07.
- Hodrick, R. J. and E. C. Prescott (1997). Postwar U.S. Business Cycles: An Empirical Investigation. *Journal of Monetary Economics* 29(1), 1–16.
- Holston, K., T. Laubach, and J. C. Williams (2017). Measuring the Natural Rate of Interest: International Trends and Determinants. *Journal of International Economics* 108, S59–S75.
- Ikeda, D. and M. Saito (2014). The Effects of Demographic Changes on the Real Interest Rate in Japan. *Japan and the World Economy* 32, 37–48.
- Imrohoroglu, A., S. Imrohoroglu, and D. H. Joines (1995). A Life Cycle Analysis of Social Security. *Economic Theory* 6, 83–114.
- Iskrev, N. (2010). Local Identification in DSGE Models. *Journal of Monetary Economics* 57(2), 189–202.

- Janssens, E. (2020). Identification in Heterogeneous Agent Models. Working paper.
- Kara, E. and L. von Thadden (2016). Interest Rate Effects of Demographic Changes in a New-Keynesian Framework. *Macroeconomic Dynamics* 20(1), 120–164.
- Karabarbounis, L. and B. Neiman (2014). The Global Decline of the Labor Share. *Quarterly Journal of Economics* 129(1), 61–103.
- Kitao, S. (2017). When Do We Start? Pension Reform in Aging Japan. *Japanese Economic Review* 68(1), 26–47.
- Kitao, S. (2018). Policy Uncertainty and Cost of Delaying Reform: The Case of Aging Japan. *Review of Economic Dynamics* 27, 81–100.
- Komunjer, I. and S. Ng (2011). Dynamic Identification of Dynamic Stochastic General Equilibrium Models. *Econometrica* 79(6), 1995–2032.
- Laubach, T. and J. C. Williams (2003). Measuring the Natural Rate of Interest. *Review of Economics and Statistics* 85(4), 1063–1070.
- Müller, U. K. (2012). Measuring Prior Sensitivity and Prior Informativeness in Large Bayesian Models. *Journal of Monetary Economics* 59(6), 581–597.
- Müller, U. K. and M. W. Watson (2018). Long-Run Covariability. *Econometrica* 86(3), 775–804.
- Muto, I., T. Oda, and N. Sudo (2016). Macroeconomic Impact of Population Aging in Japan: A Perspective from an Overlapping Generations Model. *IMF Economic Review* 64(3), 408–442.
- Sala, L. (2015). DSGE Models in the Frequency Domain. *Journal of Applied Econometrics* 30, 219–240.
- Smets, F. and R. Wouters (2007). Shocks and Frictions in U.S. Business Cycles : A Bayesian DSGE Approach. *American Economic Review* 97(3), 586–606.
- Sudo, N. and Y. Takizuka (2018). Population Aging and the Real Interest Rate in the Last and Next 50 Years — A Tale Told by an Overlapping Generations Model. Bank of Japan Working Paper Series 18-E-1.
- Wong, A. (2018). Transmission of Monetary Policy to Consumption and Population Aging. Working Paper.

Yaari, M. E. (1965). Uncertain Lifetime, Life Insurance, and the Theory of the Consumer.
The Review of Economic Studies 32(2), 137–150.

Appendix

A Steady State of Overlapping Generations Model

Denote $\ell_t = L_t/N_t$ and $\ell_t^w \equiv L_t^w/L_t$. In addition, we use lowercase letters to denote variables normalized by y and drop the time subscripts for steady states. We need to solve for:

$$\{\psi, c^r, c^w, h^r, h^w, \ell, \ell^w, \epsilon, \pi, \Omega, \lambda, k, R\}.$$

A.1 Production

Firm Capital Decision. The firm capital decision is:

$$\begin{aligned} R_t &= (1 - \alpha) \frac{Y_t}{K_t} + (1 - \delta) \\ &= (1 - \alpha) k_t^{-1} + (1 - \delta), \end{aligned}$$

which yields the steady state condition:

$$k = \frac{1 - \alpha}{R - 1 + \delta}. \quad (\text{A.1})$$

Resource Constraint. We have the aggregate resource constraint:

$$Y_t = K_{t+1} - (1 - \delta) K_t + C_t^w + C_t^r.$$

Normalizing by Y_t , we have, in steady state:

$$1 = (x + n + \delta) k + c^w + c^r. \quad (\text{A.2})$$

A.2 Households

Population. The population of retirees is:

$$\psi_{t+1} N_{t+1} = \psi_{t+1} (1 + n) N_t = \gamma \psi_t N_t + (1 - \omega) N_t,$$

which yields the steady-state condition:

$$\psi = \frac{1 - \omega}{1 + n - \gamma}. \quad (\text{A.3})$$

Wealth Share of Retirees. We have:

$$\begin{aligned}\lambda_{t+1}A_{t+1} &= \lambda_t R_t A_t + W_t \xi L_t^r - C_t^r + (1 - \omega) [(1 - \lambda_t) R_t A_t + W_t L_t^w - C_t^w] \\ &= (1 - \omega + \omega \lambda_t) R_t A_t + \alpha Y_t \frac{\xi L_t^r}{L_t} - C_t^r + (1 - \omega) \left(\alpha Y_t \frac{L_t^w}{L_t} - C_t^w \right).\end{aligned}$$

Normalizing by Y_t , we have, in steady state:

$$(1 + x + n) \lambda k = (1 - \omega + \omega \lambda) R k + \alpha - \alpha \omega \ell^w - c^r - (1 - \omega) c^w. \quad (\text{A.4})$$

Consumption of Retirees. We have:

$$C_t^r = \epsilon_t \pi_t (\lambda_t R_t A_t + H_t^r).$$

Normalizing by Y_t , we have, in steady state:

$$c^r = \epsilon \pi (\lambda R k + h^r). \quad (\text{A.5})$$

Human Wealth of Retirees. If retirees choose to supply labor, we have the law of motion:

$$\begin{aligned}H_t^r &= \xi L_t^r W_t + \frac{\psi_t N_t}{\psi_{t+1} N_{t+1}} \frac{H_{t+1}^r}{R_{t+1} / \gamma} \\ &= \alpha Y_t \frac{\xi L_t^r}{L_t} + \frac{\gamma}{1 + n} \frac{\psi_t}{\psi_{t+1}} \frac{H_{t+1}^r}{R_{t+1}}.\end{aligned}$$

Normalizing by Y_t , we have, in steady state:

$$h^r \left(1 - \frac{(1 + x) \gamma}{R} \right) = \alpha (1 - l^w).$$

Consumption of Workers. We have:

$$C_t^w = \pi_t [(1 - \lambda_t) R_t A_t + H_t^w].$$

Normalizing by Y_t , we have, in steady state:

$$c^w = \pi [(1 - \lambda) R k + h^w]. \quad (\text{A.6})$$

Human Wealth of Workers. We have the law of motion:

$$\begin{aligned} H_t^w &= L_t^w W_t + \omega \frac{N_t}{N_{t+1}} \frac{H_{t+1}^w}{R_{t+1} \Omega_{t+1}} + (1 - \omega) \xi^{\nu-1} \epsilon_{t+1}^{\frac{1}{1-\sigma}} \frac{N_t}{\psi_{t+1} N_{t+1}} \frac{H_{t+1}^r}{R_{t+1} \Omega_{t+1}} \\ &= \alpha Y_t \frac{L_t^w}{L_t} + \frac{\omega}{1+n} \frac{H_{t+1}^w}{R_{t+1} \Omega_{t+1}} + \frac{1-\omega}{1+n} \frac{\xi^{\nu-1} \epsilon_{t+1}^{\frac{1}{1-\sigma}}}{\psi_{t+1}} \frac{H_{t+1}^r}{R_{t+1} \Omega_{t+1}}. \end{aligned}$$

Normalizing by Y_t , we have, in steady state:

$$h^w \left(1 - \frac{(1+x)\omega}{R\Omega} \right) = \alpha l^w + \frac{(1+x)(1-\omega) \xi^{\nu-1} \epsilon^{\frac{1}{1-\sigma}} h^r}{\psi R \Omega}. \quad (\text{A.7})$$

Labor Supply. Define $\phi \equiv \frac{1-\nu}{\nu}$. The first-order condition for an individual worker's labor is (abusing notation):

$$L_t^w = 1 - \phi \frac{C_t^w}{W_t}.$$

Aggregating, we have:

$$L_t^w = N_t - \frac{\phi}{\alpha} \frac{L_t}{Y_t} C_t^w.$$

Dividing by L_t , we have, in steady state:

$$\ell^{-1} = \ell^w + \frac{\phi}{\alpha} c^w. \quad (\text{A.8})$$

The first-order condition for an individual retiree's labor is (abusing notation):

$$L_t^r = 1 - \phi \frac{C_t^r}{\xi W_t}.$$

Aggregating, we have:

$$L_t^r = \psi_t N_t - \frac{\phi}{\xi \alpha} \frac{L_t}{Y_t} C_t^r.$$

Dividing by L_t/ξ , we have:

$$\xi \psi_t \frac{N_t}{L_t} = 1 - \ell_t^w + \frac{\phi}{\alpha} c_t^r.$$

Therefore, in steady state:

$$(1 + \xi \psi) \ell^w = 1 + \frac{\phi}{\alpha} (c^r - \xi \psi c^w). \quad (\text{A.9})$$

Retiree Propensity to Consume. The retiree propensity to consume follows:

$$\epsilon_t \pi_t = 1 - \left[\left(\frac{W_t}{W_{t+1}} \right)^{1-\nu} R_{t+1} \right]^{\sigma-1} \beta^\sigma \gamma \frac{\epsilon_t \pi_t}{\epsilon_{t+1} \pi_{t+1}}.$$

The firm labor decision $W_t L_t = \alpha Y_t$ implies

$$\frac{W_t}{W_{t+1}} = \frac{Y_t}{Y_{t+1}} \frac{L_{t+1}}{L_t} = \frac{1}{1+x+n} \frac{\ell_{t+1}}{\ell_t}.$$

Hence, in steady state,

$$\epsilon \pi = 1 - \left(\frac{R}{(1+x)^{1-\nu}} \right)^{\sigma-1} \beta^\sigma \gamma. \quad (\text{A.10})$$

Worker Propensity to Consume. The worker propensity to consume follows:

$$\pi_t = 1 - \left[\left(\frac{W_t}{W_{t+1}} \right)^{1-\nu} R_{t+1} \Omega_{t+1} \right]^{\sigma-1} \beta^\sigma \frac{\pi_t}{\pi_{t+1}},$$

where

$$\Omega_{t+1} = \omega + (1-\omega) \xi^{\nu-1} \epsilon_{t+1}^{\frac{1}{1-\sigma}}.$$

This yields steady-state conditions:

$$\pi = 1 - \left(\frac{R\Omega}{(1+x)^{1-\nu}} \right)^{\sigma-1} \beta^\sigma \quad (\text{A.11})$$

$$\Omega = \omega + (1-\omega) \xi^{\nu-1} \epsilon^{\frac{1}{1-\sigma}}. \quad (\text{A.12})$$

B Bayesian Estimation

B.1 Data

Table B.1 lists the data series that we use as observables, as well as their mapping to the variables or parameters in the OLG model. For labor quantity variables, we focus on employment at the extensive margin, so that labor supply in the model is taken to be the fraction of workers and retirees respectively who are employed. For wages, we use hourly earnings for manufacturing, which is close to the series for monthly private sector earnings.

Data Series	Model Counterpart	Source
GDP growth	$x + n$	FRED
age 15-64 pop. growth	n	OECD
age 15-64 pop. / total 15+ pop.	$1 / (1 + \psi)$	OECD
age 15-64 emp. / total 15+ emp.	$\xi \ell^w / [1 - (1 - \xi) \ell^w]$	OECD
employment-population ratio	$[1 - (1 - \xi) \ell^w] / \xi$	OECD
real wage growth	x	FRED
real interest rate	$R - 1$	Bank of Japan, IMF DSBB
labor share	α	Penn World Table 9.0
capital-output ratio	k	Penn World Table 9.0

Table B.1: Observables, model counterparts, and data sources

B.2 Prior

Structural Parameters. The structural parameters and their priors are summarized in Table B.2. For each regime, we estimate the time-varying parameters $(x, n, \gamma, \xi, \nu, \alpha)$. These are drawn iid for each regime from a distribution whose mean and variance are estimated. The priors for these means and variances are also reported in Table B.2. Figure B.1 compares the prior for (β, σ, δ) to existing calibrations of the Gertler (1999) model.

VAR Parameters. For the VAR parameters $\{\Phi(s_t), \Sigma(s_t)\}$, we use a normal-inverse-Wishart prior that shrinks toward white noise, to ensure that v_t captures primarily high-frequency variation. The normal-inverse-Wishart prior makes it straightforward to integrate out the VAR parameters. The estimation results are not significantly affected by the particular Normal-inverse Wishart prior used. The coefficients are drawn iid across regimes.

Conditional on $\mu(s_t)$, we have a prior that $v_{t^*} \sim \mathcal{N}(\mu(0) - \mu(1), \Omega)$ and $v_T \sim \mathcal{N}(0, \Omega)$, where Ω is a diagonal matrix of variances that we calibrate to be equal to the variances of each series in $\{y_t\}$. This prior on the initial and terminal condition implies that we expect the measurement error to be large immediately after the regime change but to shrink toward the end of the sample. Intuitively, y_t starts out of steady state in t^* , but converges toward its steady state. The vector autoregression allows this convergence to be modeled flexibly. The results do not change substantially if we ignore the prior for v_{t^*} and v_T .

Structural Break. The prior for t^* is flat for $t^* \in [1985, 2009]$ and zero otherwise. The restriction ensures that t^* does not lie in the initial or final periods of the sample. In the estimation, t^* is tightly estimated to lie around 1991.

		Type	Mean	Std Dev
<u>Structural Parameters</u>				
Fixed				
ω	one minus retirement rate	calibrated	0.98	–
$100(\beta^{-1} - 1)$	discount rate	Gamma	3.50	0.50
σ	intertemporal elasticity of substitution	Normal	0.35	0.15
δ	depreciation rate	Beta	0.08	0.02
Time-varying				
$100x$	productivity growth	Normal	μ_x	σ_x
$100n$	population growth	Normal	μ_n	σ_n
$100(\gamma^{-1} - 1)$	probability of death	Gamma	μ_γ	σ_γ
ξ	relative productivity of retirees	Beta	μ_ξ	σ_ξ
ν	one minus disutility of labor	Beta	μ_ν	σ_ν
α	labor share	Normal	μ_α	σ_α
<u>Hyperparameters</u>				
Means				
μ_x	productivity growth	Normal	1.50	0.50
μ_n	population growth	Normal	1.00	0.50
μ_γ	probability of death	Gamma	8.00	0.50
μ_ξ	relative productivity of retirees	Beta	0.50	0.10
μ_ν	one minus disutility of labor	Beta	0.50	0.10
μ_α	labor share	Normal	0.65	0.05
Standard Deviations				
σ_x	productivity growth	Inv. Gamma	1.00	1.00
σ_n	population growth	Inv. Gamma	1.00	1.00
σ_γ	probability of death	Inv. Gamma	2.00	2.00
σ_ξ	relative productivity of retirees	Inv. Gamma	0.10	0.10
σ_ν	one minus disutility of labor	Inv. Gamma	0.10	0.10
σ_α	labor share	Inv. Gamma	0.10	0.10

Table B.2: Prior for structural parameters.

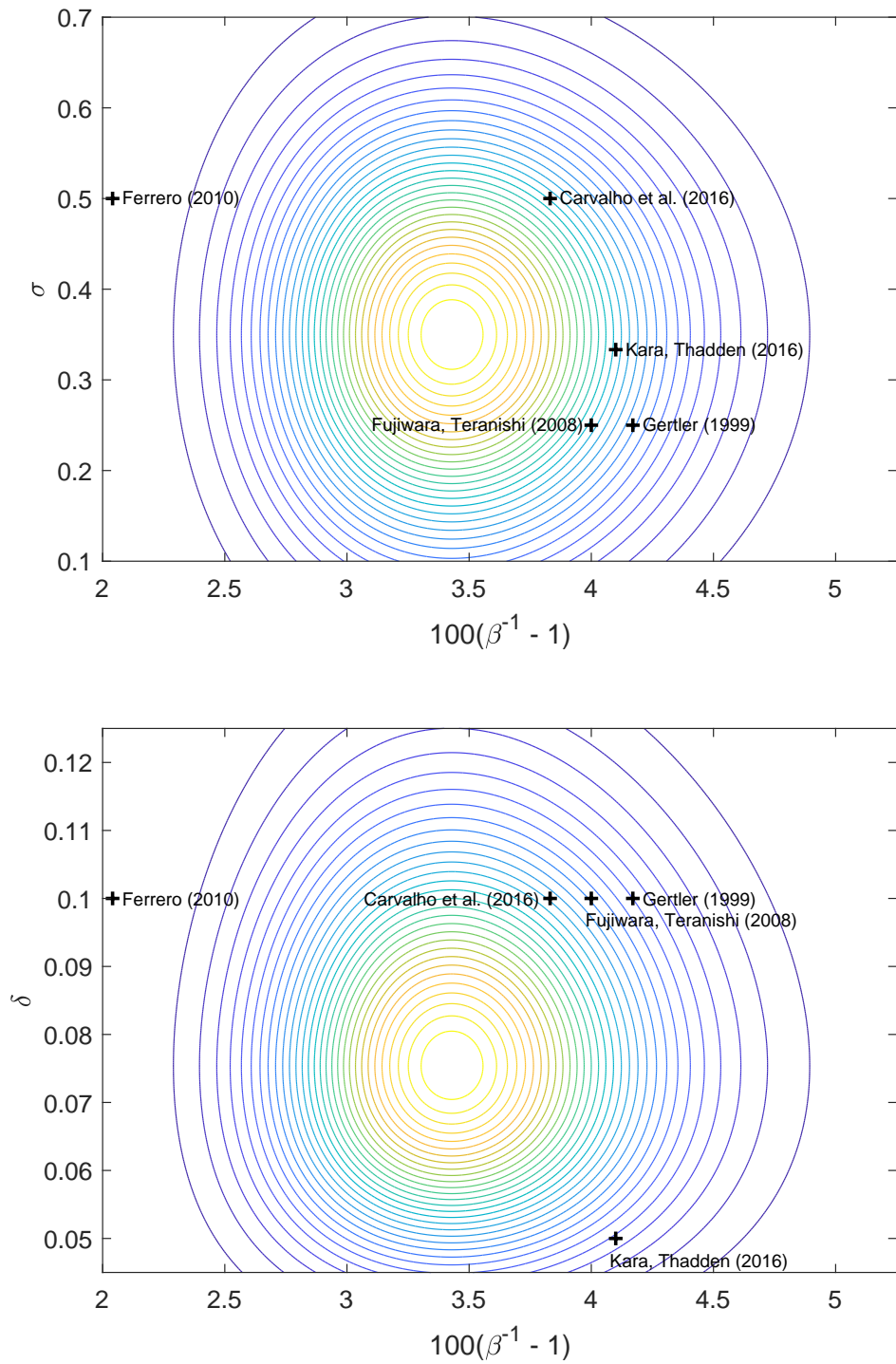


Figure B.1: Original prior and existing calibrations of (β, σ, δ) for the [Gertler \(1999\)](#) model. Black crosses indicate calibrations; lines are level curves for the joint distribution under the prior. **Top:** β and σ ; **Bottom:** β and δ .

B.3 Markov Chain Monte Carlo

The MCMC algorithm to estimate the model (2.1)-(2.3) has two main blocks. First we draw t^* given the structural parameters. Next, we draw the structural parameters given t^* .

Each step involves one or more Metropolis-Hastings draws. Given the structural parameters, we compute \bar{y} by solving numerically solving for the steady state of the OLG model given those parameters. Given \bar{y} , we obtain v_t . Since v_t follows a VAR(1) with a normal-inverse-Wishart prior, we can analytically integrate out the VAR parameters when evaluating the posterior. This reduces the size of the parameter space and improves convergence.

For the burn-in draws, we break the second step into several blocks. We make draws for the fixed and time-varying parameters in separate blocks, and break the time-varying parameters into two blocks corresponding to their values before and after the structural break. The blockwise draws improve convergence when our proposal density is not yet optimized. For the main draws, we draw all the structural parameters in a single block and scale the covariance of the burn-in draws for the proposal density.

C Prior Sensitivity

We use the SMC algorithm from Ho (2020), which is an extension of the SMC algorithm for Bayesian estimation in Herbst and Schorfheide (2014). We use 10^5 particles, 500 SMC steps and 5 Metropolis-Hastings steps, and repeat this 20 times in parallel. This takes approximately 15 hours to complete.¹³ As a reference, the main Bayesian estimation takes approximately five hours to complete.

To produce Figures 5.1 and 6.1, we extrapolate the output from the SMC algorithm. First, we take the median relative entropy of the 20 runs of the SMC algorithm for each SMC step. The i th SMC step corresponds to the same worst-case quantile across runs. Next, we run a polynomial regression of the worst-case quantile on the median relative entropy. Finally, we use the regression to predict the worst-case quantiles for the $r = 0.005$ or $r = 0.025$.

D An Alternative Parametric Prior

To demonstrate that the sensitivity of the estimated effects to the prior does not depend on the nonparametric nature of REPS, we now show that the posterior results also change substantially when we estimate the model using a different parametric prior. The new prior

¹³We obtain similar results with half the SMC steps, which requires less than half the time.

incorporates the worst-case distortions for the 16% quantile for the effect of the relative productivity of retirees ξ and the 84% quantile for the effect of the survival rate γ .

The alternative prior is reasonable in two dimensions. Firstly, it remains in a parametric family, retaining potentially desirable smoothness properties that can be violated by the worst-case prior from REPS. Secondly, the distortions put greater weight on parameter values that are consistent with calibrations of OLG models in the literature.

D.1 Prior

We take the new prior of (β, σ) to be:

$$\begin{bmatrix} 100(\beta^{-1} - 1) \\ \sigma \end{bmatrix} \sim \mathcal{N} \left(\begin{bmatrix} 2.00 \\ 0.50 \end{bmatrix}, \begin{bmatrix} 2.00^2 & -0.315 \\ -0.315 & 0.45^2 \end{bmatrix} \right).$$

The new prior for δ is an independent Beta distribution that has mean 0.06 instead of 0.08, and has the same standard deviation as original prior of 0.02.

The new prior places greater weight on regions that the REPS analysis suggests are particularly important for the estimated effects of the survival rate γ and the relative productivity of retirees ξ . In particular, it places more weight on regions of the parameter space with small values of β^{-1} and δ , as well as large values of σ . These regions coincide with the worst-case distortions for the 16% quantile for the effect of ξ , and the 84% quantile for the effect of γ . The REPS analysis also showed that increasing the prior mass on large values of σ reduces the estimated effect of population growth n .

While the worst-case distortions indicate features of the prior that are important for our estimates, some of these distortions appear implausible. For example, a discount rate $\beta < 0.95$ or depreciation rate $\delta > 0.20$ are unlikely. To discipline the prior, we ensure that it is consistent with the range of values used in the OLG literature, shown in Figure D.1.

We decrease the prior mean and increase the prior variance of β^{-1} . In addition, the Gaussian prior on β^{-1} no longer imposes the $\beta < 1$ restriction of the original Gamma distribution prior. The new support is reasonable because the restriction that $\beta < 1$ in representative agent models does not apply in OLG models. Instead, the stochastic death introduces discounting on top of the time preference β . [Imrohoroglu et al. \(1995\)](#) argue that a value of $\beta > 1$ matches the empirical evidence and allows their model to fit the US wealth-income ratio. Nevertheless, we keep a mean discount rate of 2% to reflect the positive discount rates used by a majority of the literature.

We increase the prior mean and variance of σ to reflect the wide range of values in the literature beyond versions of the [Gertler \(1999\)](#) model. The increased mean is consistent

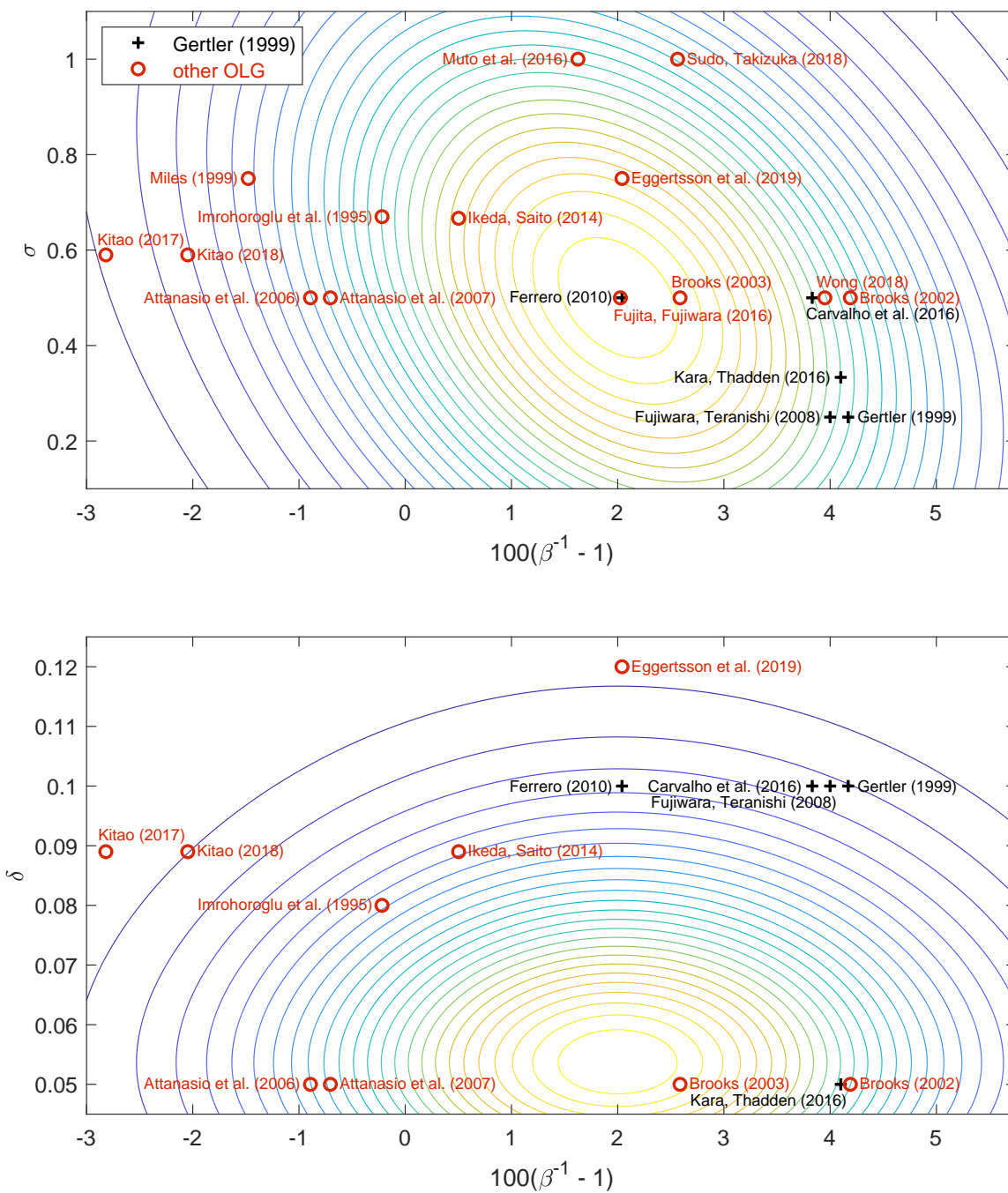


Figure D.1: Existing calibrations of (β, σ, δ) . Black crosses indicate calibrations of models based on Gertler (1999); red circles indicate calibrations of other OLG models; lines are level curves for the joint distribution under the alternative prior. **Top:** β and σ ; **Bottom:** β and δ .

with Figure D.1, which shows sixteen papers with $\sigma \geq 0.5$. Furthermore, there is a wide range of empirical estimates of σ between 0 and 2.

Under the new prior, β and σ have a correlation of -0.35 , which matches the correlation for the sample of calibrations in Figure D.1. Many papers pick σ independently and calibrate β to match a steady-state statistic, such as the interest rate in a given year or the capital-output ratio. The correlation reflects the fact that β and σ jointly determine these moments. For example, in a representative agent neoclassical growth model with constant relative risk aversion, the household’s Euler equation implies a negative relationship between β and σ to jointly match consumption growth and interest rates. In the model here, the dependence is summarized by the expressions (3.14) and (3.15) for the marginal propensities to consume. Since the REPS analysis has shown that the original assumption of independence between β and σ is not innocuous, we introduce this correlation to our prior.

To illustrate the importance of δ for the estimated effects of the survival rate γ and relative productivity ξ , we reduce the prior mean for δ . The prior mean of 0.06 is closer to the World Penn Tables estimate of 0.05 for the depreciation rate in Japan. Figure D.1 shows that this is a plausible prior mean, with values of δ ranging between 0.05 and 0.12 in the literature. We keep δ independent from β and σ for two reasons. Firstly, the correlation between δ and (β, σ) among the papers considered in Figure D.1 is smaller than the correlation between β and σ . Secondly, β and σ are both preference parameters that directly influence households’ savings decisions, but δ governs the production side of the economy.

While this prior has been chosen to emphasize the results from the REPS analysis, it is one that would be plausible without observing the data. The prior for (β, σ) was chosen based on existing calibrations with no reference to the data. The prior for δ is also plausible given existing empirical work that does not use the data we consider. We do not take a stand on which prior should be preferred. Instead, we emphasize that a complete analysis of the data requires understanding the range of results that can arise from different possible priors and take the alternative prior as an illustrative example. In particular, although there are numerous other plausible priors, we pick this particular one because the REPS results show that these specific changes in the prior can lead to large changes in the posterior estimates.

D.2 Results

Table D.1 shows that changing the prior substantially changes the estimates for the effects of demographics on the interest rate. In all three cases, the mean changes by more than one posterior standard deviation, and the new mean is not contained in the original 68% credible interval. The change in the posterior mean of up to 0.6 percentage points is also

		Original			With Alternative Prior		
		mean	sd	68% C.I.	mean	sd	68% C.I.
Structural Parameters							
$100(\beta^{-1} - 1)$	disc. rate	3.245	0.442	(2.804, 3.690)	0.022	1.353	(-1.365, 1.425)
σ	IES	0.574	0.103	(0.471, 0.677)	0.790	0.149	(0.624, 0.959)
100δ	dep. rate	7.861	1.485	(6.393, 9.317)	3.328	0.890	(2.433, 4.282)
Effects of Demographics							
$100(\widehat{R}_T^n - R_T)$	pop. growth	0.875	0.159	(0.725, 1.023)	0.696	0.136	(0.571, 0.829)
$100(\widehat{R}_T^\gamma - R_T)$	surv. rate	0.671	0.266	(0.429, 0.905)	1.110	0.225	(0.891, 1.332)
$100(\widehat{R}_T^\xi - R_T)$	rel. prod.	1.206	0.412	(0.793, 1.618)	0.635	0.465	(0.184, 1.118)

Table D.1: Posterior statistics for effects of demographics with original and alternative parametric priors.

economically significant.

The changes in the estimated effects are driven by large changes in the posterior estimates for (β, σ, δ) . The mean for the discount rate falls from 3% to 0%; the mean for the intertemporal elasticity of substitution rises from 0.57 to 0.79; and the mean depreciation rate falls from 8% to 3%.¹⁴ Relative to the posterior standard deviations, these are statistically large changes. Moreover, the resulting changes in the estimated effects of demographics indicate that these are also economically significant. In contrast to the posterior for (β, σ, δ) , the estimates for the long-run mean μ and the time-varying parameters are relatively unaffected by the change in the prior.

The directions in which the estimates change are consistent with the worst-case posteriors in Figures 5.2 and 5.3. The estimate for the effect of population growth decreases because it depends negatively on σ . The new prior is chosen in line with the distortions that increase the 84% quantile for the effect of the survival rate and thus increase the estimate of that effect. The estimate for the effect of the relative productivity of retirees decreases as the alternative prior places greater weight on similar regions to the worst-case distortions for the 16% quantile for that effect.

We have attained these changes in the posterior using an economically plausible alternative prior that lies in a commonly used distributional family. The prior sensitivity identified by REPS therefore remains an issue even when we restrict ourselves to parametric alternative

¹⁴The fact that the new posterior estimate for the depreciation rate is smaller than typical values in the literature should not be viewed as a criticism of a prior that was deemed plausible before observing the data. Rather, the result is further evidence that the current data are insufficient to discipline the estimate of δ .

priors. In addition, many of the calibrations considered in Figure D.1 fall around the high prior density region for (β, σ, δ) , suggesting that the results in the literature are also sensitive to the exact calibration strategy. For example, while it is common to calibrate β to match the interest rate in a given year (e.g., Ikeda and Saito (2014); Carvalho et al. (2016)), we find a large dispersion in the likelihood for β here that translates into imprecise estimates for the effect of demographics. Once we account for uncertainty in the natural interest rate as well as the joint dependence of all the observables on the full parameter vector, we find that the likelihood supports a wide range of parameter combinations that are consistent with the data. Our analysis has shown that the estimated effects of demographics vary substantially with parameter values in this range.