# The Distributional Effects of Redistributional Tax Policy *

Jason DeBacker† Richard W. Evans‡ Evan Magnusson§
Kerk L. Phillips¶ Shanthi P. Ramnath‖ Isaac Swift**

March 2018

**Abstract**

This paper uses a large scale overlapping generations model with heterogeneity across the life cycle and over lifetime income groups to evaluate the distributional effects of tax policy. The model is calibrated to the U.S. economy and includes realistic demographics, mortality risk, and progressive income taxes. The model generates distributions of hours worked, earnings, and wealth that are consistent with those observed in the U.S. We consider the effects of two policies that have the same steady-state revenue effect: (i) a progressive wealth tax and (ii) a progressive increase in income tax rates. We find that the wealth tax is extremely effective at reducing inequality relative to an increase in the progressivity of the income tax with the same steady-state tax revenue. The costs of reducing inequality using the wealth tax are primarily borne by the top 10 percent of wage earners and by individuals over the age of 60. The reductions in wealth and consumption from the income tax are concentrated among the top 20 percent of wage earners and among middle-aged individuals between the ages of 40 and 70.

*keywords:* inequality, wealth tax, income tax, overlapping generations.

*JEL classification:* C61, C63, D91, E21, H30, J11

*We are grateful to Kevin Hassett, Alan Viard, Alex Brill, Matt Jensen, Alan Auerbach, Aspen Gorry, Frank Caliendo, Richard W. Evans, Sr., and Sam Dodini for helpful comments and suggestions. This research benefited from support from the Brigham Young University Macroeconomics and Computational Laboratory and from the Open Source Policy Center at the American Enterprise Institute. All Python code and documentation for the computational model is available at https://github.com/rickecon/wealthtax.

†Darla Moore School of Business, University of South Carolina, Department of Economics, DMSB 427B, Columbia, SC 29208, (803) 777-1649, jason.debacker@moore.sc.edu.

‡University of Chicago, Becker Friedman Institute, Saieh Hall, Room 250, Chicago, IL 60637, (773) 702-9169, rwevans@uchicago.edu.

§Stanford University, Department of Economics, 579 Serra Mall, Stanford, CA 94305, evan.magnusson@stanford.edu.

¶Congressional Budget Office, FHOB, 2nd & D Streets, SW, Washington, DC 20002, kerk.phillips@cbo.gov.

‖U.S. Department of the Treasury, Office of Tax Analysis, 1500 Pennsylvania Ave. NW, Washington, DC 20220, (202) 622-1789, shanthi.ramnath@treasury.gov.

**University of Minnesota, Department of Economics, 3-154 Hansen Hall, 1925 Fourth Street South, Minneapolis, Minnesota 55455, swift105@umn.edu.

# 1 Introduction

Over the last several decades, most developed countries have seen steady increases in measures of wealth and income inequality. These trends are documented in Piketty (2014, 2011), Saez and Zucman (2016), Kopczuk (2015), Piketty and Saez (2003), Cagetti and De Nardi (2008), and Wolff (1998, 1992). De Nardi (2015) surveys current empirical and theoretical research on the underlying causes of inequality. She notes the difficulty in modeling the striking inequality observed in the data and the number of causal factors that have been modeled in attempts to do so.[1] Because there is not consensus as to which fundamental causes of inequality are most important, the normative implications of these trends are not yet clear.

The goal of our paper is to build a model that allows for some of the key potential causes of inequality, calibrate the model to match the current U.S. economy, and simulate the effects of two different redistributive tax policies on inequality and other economic outcomes. Piketty (2014) proposes two main policies to reduce inequality: (i) a global tax on wealth, and (ii) a more progressive income tax. In this paper, we test the effects of a wealth tax similar to the one proposed by Piketty versus an increase in the progressivity of the income tax with the same steady-state change in tax revenues.

We do not address the issues of what is the optimal wealth tax and labor income tax based on functional analysis of a model with an assumed social welfare function, as is the subject of a large literature beginning with Mirrlees (1971) and surveyed more recently by Kocherlakota (2006). Instead, we narrowly focus on the effect of a wealth tax on measures of inequality and compare those effects to those of an "equally sized" income tax increase. However, we do present the effects of these two taxes on one measure of social welfare—the weighted sum of the steady-state period utility values.

To perform these experiments, we develop a large scale overlapping generations

[1] The main causes of inequality from the recent literature explored by De Nardi (2015) include heterogeneity in patience, transmission of human capital, voluntary and involuntary bequests across generations, entrepreneurship or high returns to capital, borrowing constraints, and earnings risk.

model with heterogeneity across the life cycle and over lifetime earnings potential. A detailed lifecycle model is important to understanding our key questions of how taxes on wealth and income impact inequality. It is only through such a model that one can separate the effects of tax policy on lifetime, life-cycle, and cross-sectional inequality. Moreover, the detailed nature of the model and calibration allow us to speak to the quantitative significance of such distributional effects.

We calibrate the model's parameters to match the behavior of the U.S. economy. Using microeconomic data from U.S. income tax returns, we calibrate lifetime earnings profiles for the entire distribution of earners, including the top one-percent. The importance of these high income and high wealth individuals in driving the trends in inequality is highlighted by Piketty and Saez (2003). And we find that our qualitative results hinge critically on the behavior and responses of the top one-percent of earners in our model. Our calibration also includes rich demographic dynamics and mortality risk. Population dynamics allow us to consider distributional effects in both the short and long run, as well as effects in the cross-section and across the lifetimes of agents, as the demographics of the population shift.

We also carefully model the U.S. personal income tax system, inclusive of the progressive tiers of increasing marginal tax rates, exemptions, deductions, and choice of filing status, as well as the social security system. As such, our model captures many of the important channels through which personal income and wealth taxes affect economic efficiency and inequality through life-cycle savings and labor supply responses.

Our first policy experiment is a progressive wealth tax, similar to that suggested by Piketty (2014, pp. 515-539). Our second policy experiment is an increase in the progressivity of the current U.S. personal income tax. We make these two policies comparable by choosing tax schedules for each that give equivalent steady-state tax revenue. We find that the wealth tax is extremely effective at reducing inequality in wealth, income, and consumption relative to an increase in the progressivity of the income tax. Although the reductions in inequality across lifetime income groups from the wealth tax are significant, the reduction in inequality over the life cycle (within lifetime income groups) is even more stark. The costs of reducing inequality using the wealth tax are primarily borne by the top 10 percent of wage earners and especially by individuals over the age of 60.

The income tax policy experiment results in a smaller reduction in inequality. In this case, the reduction in income, wealth, and consumption inequality come entirely from reductions in inequality across lifetime income groups. The reductions in wealth and consumption from the income tax are focused primarily among the top 20 percent of wage earners and among middle aged individuals between the ages of 40 and 70. In addition, the income tax change only has a small effect on the steady-state distribution of labor supply, with the exception of the top one percent of wage earners who reduce their labor supply significantly after age 55. We also find that the wealth tax imposes a smaller distortion on aggregate income, capital, labor, and consumption.

The approach we take to modeling the economy builds on the work of Auerbach and Kotlikoff (1987). It is this structure with which we evaluate income and wealth taxes as redistributional tools. Kambourov et al. (2013) also evaluate wealth taxes, but with a different modeling approach. Most notably, they assume that individuals are heterogeneous in their returns to capital income. Given these heterogeneous returns, they find that a wealth tax provides efficiency gains, in that it reallocates capital to those who have the highest returns. This increases economic efficiency and reduces equality in income and wealth. The result is that a wealth tax may find support from a wide range of social welfare functions, but for its efficiency, rather than redistribution properties.

A paucity of data make it extremely difficult to test whether, and to what extent, individuals have heterogeneous earnings ability for capital income. Saez and Zucman (2016) use capital income to impute wealth over a long U.S. time series and argue that there is no significant heterogeneity in returns to capital. We follow this and assume that all individual realize the same returns to capital income. We also follow the evidence of many others (for example, DeBacker et al. (2013) and Lochner and Shin (2014)), who find substantial heterogeneity in labor income earnings. This leads to inequality in both labor and capital income as individuals with higher labor income earning ability will tend to accumulate more wealth and thus have higher capital income. These differences in income processes lead us to a different conclusions about a wealth tax than Kambourov et al. (2013).

The remainder of the paper is organized as follows. Section 2 presents the baseline model, and Section 3 discusses the model calibration and fit. Section 4 presents the results from the two policy experiments. Section 5 concludes.

## 2 Baseline Model

Our model is comprised of heterogeneous individuals, perfectly competitive firms, and a government with a balanced budget requirement. A unit measure of identical firms make a static profit maximization decision in which they rent capital and hire labor to maximize profits given a Cobb-Douglas production function. The government levies taxes on individuals and makes lump sum transfers to individuals according to a balanced budget constraint.

Individuals are assumed to live for a maximum of $E + S$ periods. We define an age-$s$ individual as being in youth and out of the workforce during ages $1 \leq s \leq E$. We implement this dichotomy of being economically relevant by age in order to more easily match true population dynamics. Individuals enter the workforce at age $E+1$ and remain in the workforce until they die or until the maximum age $E+S$. Because of mortality risk, they leave both intentional bequests at the end of life ($s = E + S$) as well as accidental bequests if they die before the maximum age of $E + S$.

When individuals are born at age $s = 1$, they are randomly assigned to one of $J$ lifetime income (ability) types. Individuals remain deterministically in their assigned lifetime income group throughout their lives. The hourly earnings process is calibrated to match the wage distribution by age in the United States, and labor is endogenously supplied by individuals. Our calibration of the hourly earnings process allows for a skewed distribution of earnings that fits U.S. life-cycle hourly earnings data. The economic environment is one of incomplete markets because the overlapping generations structure prevents households from perfectly smoothing consumption.

We calibrate the population demographics and deterministic lifetime earnings profiles from external sources. We then calibrate parameters of the model to match the steady-state distributions of labor supply and wealth from the model to those from the data.

## 2.1 Population dynamics and lifetime earnings profiles

We define $\omega_{s,t}$ as the number of individuals of age $s$ alive at time $t$. A measure $\omega_{1,t}$ of individuals with heterogeneous working ability is born in each period $t$ and live for up to $E+S$ periods, with $S \geq 4$.[2] Individuals are termed "youth", and do not participate in market activity during ages $1 \leq s \leq E$. The individuals enter the workforce and economy in period $E+1$ and remain in the workforce until they unexpectedly die or live until age $s = E+S$.[3] The population of agents of each age in each period $\omega_{s,t}$ evolves according to the following function,

$$\begin{aligned} \omega_{1,t+1} &= \sum_{s=1}^{E+S} f_s \omega_{s,t} \quad \forall t \\ \omega_{s+1,t+1} &= (1 + i_s - \rho_s)\omega_{s,t} \quad \forall t \quad \text{and} \quad 1 \leq s \leq E+S-1 \end{aligned} \tag{1}$$

where $f_s \geq 0$ is an age-specific fertility rate, $i_s$ is an age-specific net immigration rate, $\rho_s$ is an age specific mortality hazard rate,[4] and $1 + i_s - \rho_s$ is constrained to be nonnegative. The total population in the economy $N_t$ at any period is simply the sum of individuals in the economy, the population growth rate in any period $t$ from the previous period $t-1$ is $g_{n,t}$, $\tilde{N}_t$ is the working age population, and $\tilde{g}_{n,t}$ is the working age population growth rate in any period $t$ from the previous period $t-1$.[5]

[2] Theoretically, the model works without loss of generality for $S \geq 3$. However, because we are calibrating the ages outside of the economy to be one-fourth of $S$ (e.g., ages 21 to 100 in the economy, and ages 1 to 20 outside of the economy), we need $S$ to be at least 4.

[3] We model the population with individuals age $s \leq E$ outside of the workforce and economy in order most closely match the empirical population dynamics. Appendix A-1 gives more detail on the population process and its calibration.

[4] The parameter $\rho_s$ is the probability that a individual of age $s$ dies before age $s+1$.

[5] Appendix A-1 describes in detail the exogenous population dynamics.

These parameters are defined as:

$$N_t \equiv \sum_{s=1}^{E+S} \omega_{s,t} \quad \forall t \tag{2}$$

$$g_{n,t+1} \equiv \frac{N_{t+1}}{N_t} - 1 \quad \forall t \tag{3}$$

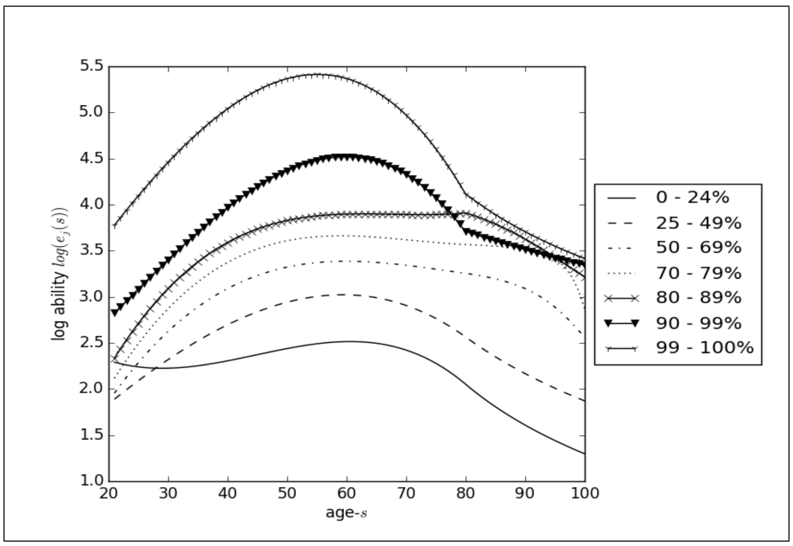
$$\tilde{N}_t \equiv \sum_{s=E+1}^{E+S} \omega_{s,t} \quad \forall t \tag{4}$$

$$\tilde{g}_{n,t+1} \equiv \frac{\tilde{N}_{t+1}}{\tilde{N}_t} - 1 \quad \forall t \tag{5}$$

At birth, a fraction $\lambda_j$ of the $\omega_{1,t}$ measure of new agents is randomly assigned to each of the $J$ lifetime income groups, indexed by $j = 1, 2, ...J$, such that $\sum_{j=1}^{J} \lambda_j = 1$. Note that lifetime income is endogenous in the model, therefore we define lifetime income groups by a particular path of earnings abilities. For each lifetime income group, the measure $\lambda_j \omega_{s,t}$ of individuals' effective labor units (which we also call ability) evolve deterministically according to $e_{j,s}$. This gives a different life cycle profile of earnings to each lifetime income group. An individual's working ability evolves over its working-age lifetime $E + 1 \leq s \leq E + S$ according to this age-dependent deterministic process. The processes for the evolution of the population weights $\omega_{s,t}$ as well as lifetime earnings are exogenous inputs to the model.

Figure 1 shows the calibrated trajectory of effective labor units (ability) $e_{j,s} \in \mathcal{E} \subset \mathbb{R}_{++}$ by age $s$ for each type $j$ for lifetime income distribution $\{\lambda_j\}_{j=1}^{7} = [0.25, 0.25, 0.20, 0.10, 0.10, 0.09, 0.01]$. We show effective labor units in logarithms because the difference in levels between the top one percent and the rest of the distribution is so large. We have calibrated this process by imputing U.S. Current Population Survey annual hours data to a panel of tax returns complied by the Internal Revenue Service's (IRS) Statistics of Income (SOI) program. This imputation allows us to measure hourly earnings for a large sample with earnings that are not top-coded. Hourly earnings in these data correspond to effective labor units in our model. That is, all individuals have the same time endowment and receive the same wage per effective labor unit, but some are endowed with more effective labor units. We utilize a measure of lifetime income, by using potential lifetime earnings, that allows us to define income groups in a way that accounts for the fact that earnings of individuals observed in the data are endogenous. It is in this way that we are able to calibrate the exogenous lifetime earnings profiles form the model with their data counterparts. Appendix A-2 details this calibration.

**Figure 1: Exogenous life cycle income ability paths** $\log(e_{j,s})$ **with** $S = 80$ **and** $J = 7$




## 2.2 Individual problem

Individuals are endowed with a measure of time $\tilde{l}$ in each period $t$, and they choose how much of that time to allocate between labor $n_{j,s,t}$ and leisure $l_{j,s,t}$ in each period. That is, an individual's labor and leisure choice is constrained by his total time endowment, which constraint is identical across all individuals.

$$n_{j,s,t} + l_{j,s,t} = \tilde{l} \tag{6}$$

At time $t$, all age-$s$ individuals with ability $e_{j,s}$ know the real wage rate, $w_t$, and know the one-period real net interest rate, $r_t$, on bond holdings, $b_{j,s,t}$, that mature at the beginning of period $t$. They also receive accidental and intentional bequests. They choose how much to consume $c_{j,s,t}$, how much to save for the next period by loaning capital to firms in the form of a one-period bond $b_{j,s+1,t+1}$, and how much to work $n_{j,s,t}$ in order to maximize expected lifetime utility of the following form,

$$
\begin{aligned}
U_{j,s,t} &= \sum_{u=0}^{E+S-s} \beta^u \left[\prod_{v=s}^{s+u-1} (1-\rho_v)\right] u\left(c_{j,s+u,t+u}, n_{j,s+u,t+u}, b_{j,s+u+1,t+u+1}\right) \\
\text{and} \quad & u\left(c_{j,s,t}, n_{j,s,t}, b_{j,s+1,t+1}\right) = \frac{\left(c_{j,s,t}\right)^{1-\sigma}-1}{1-\sigma}... \\
& + e^{g_y t(1-\sigma)}\chi^n_s \left(b\left[1-\left(\frac{n_{j,s,t}}{\tilde{l}}\right)^{\upsilon}\right]^{\frac{1}{\upsilon}} + k\right) + \rho_s \chi^b_j \frac{\left(b_{j,s+1,t+1}\right)^{1-\sigma}-1}{1-\sigma} \\
& \qquad \forall j,t \quad \text{and } E+1 \leq s \leq E+S
\end{aligned}
\tag{7}
$$

where $\sigma \geq 1$ is the coefficient of relative risk aversion on consumption and on intended (precautionary) bequests, $\beta \in (0,1)$ is the agent's discount factor, and the product term in brackets depreciates the individual's discount factor by the cumulative mortality rate. The disutility of labor term in the period utility function looks nonstandard, but is simply the upper right quadrant of an ellipse that closely approximates the standard CRRA utility of leisure functional form.[6] The term $\chi^n_s$ is a constant term that varies by age $s$ influencing the disutility of labor relative to the other arguments in the period utility function,[7] and $g_y$ is a constant growth rate of labor augmenting technological progress, which we explain in Section 2.3.[8]

The last term in (7) incorporates a warm-glow bequest motive in which individuals value having savings to bequeath to the next generation in the chance they die before

[6]Appendix A-3 describes how the elliptical function closely matches the more standard utility of leisure of the form $\frac{(\tilde{l}-n_{j,s,t})^{1+\theta}}{1+\theta}$. This elliptical utility function forces an interior solution that automatically respects both the upper and lower bound of labor supply, which greatly simplifies the computation of equilibrium. In addition, the elliptical disutility of labor has a Frisch elasticity that asymptotes to a constant rather than increasing to infinity as it does in the CRRA case. For a more in-depth discussion see Evans and Phillips (2018)

[7]In Section 3, we discuss our calibration of $\chi^n_s$ and $\chi^b_j$ to match average labor hours by age and some moments of the distribution of wealth.

[8]The term with the growth rate $e^{g_y t(1-\sigma)}$ must be included in the period utility function because consumption and bequests will be growing at rate $g_y$ and this term stationarizes the individual Euler equation by making the marginal disutility of labor grow at the same rate as the marginal benefits of consumption and bequests. This is the same balanced growth technique as that used in Mertens and Ravn (2011).

the next period. Including this term is essential to generating the positive wealth levels across the life cycle and across abilities that exist in the data. In addition, the term $\chi^b_j$ is a constant term that varies by lifetime income group $j$ influencing the marginal utility of bequests, $b_{j,s+1,t+1}$ relative to the other arguments in the period utility function. Allowing the $\chi^b_j$ scale parameter on the warm glow bequest motive vary by lifetime income group is critical for matching the distribution of wealth. As was mentioned in Section 2.1, individuals in the model have no income uncertainty because each lifetime earnings path $e_{j,s}$ deterministic, model agents thus hold no precautionary savings. As we will show in Section 3, calibrating the $\chi^b_j$ for each income group $j$ allows us to recapture in a reduced form way some of the characteristics that individual income risk provides.

The parameter $\sigma \geq 1$ is the coefficient of relative risk aversion on bequests, and the mortality rate $\rho_s$ appropriately discounts the value of this term.[9] Note that, because of this bequest motive, individuals in the last period of their lives ($s = S$) will die with positive savings $b > 0$. Also note that the CRRA utility of bequests term prohibits negative wealth holdings in the model, but is not a strong restriction since none of the wealth data for the lifetime income group $j$ and age $s$ cohorts is negative except for the lowest quartile. And their wealth is only slightly negative as is shown later in Figure 5.

The per-period budget constraints for each agent normalized by the price of consumption are the following,

$$\begin{split}
c_{j,s,t} + b_{j,s+1,t+1} \leq (1+r_t)\, b_{j,s,t} + w_t e_{j,s} n_{j,s,t} + \frac{BQ_{j,t}}{\lambda_j \tilde{N}_t} - T_{j,s,t} \\
\text{where} \quad b_{j,E+1,t} = 0 \quad \text{for} \quad E+1 \leq s \leq E+S \quad \forall j,t
\end{split} \tag{8}$$

where $\tilde{N}_t$ is the total working age population at time $t$ defined in (4) and $\lambda_j \tilde{N}_t$ is the number of the total working individuals of type $j$ in period $t$. Note that the price of consumption is normalized to one, so $w_t$ is the real wage and $r_t$ is the net real interest

[9] It is necessary for the coefficient of relative risk aversion $\sigma$ to be the same on both the utility of consumption and the utility of bequests. If not, the resulting Euler equations are not stationarizable.

rate. The term $BQ_{j,t}$ represents total bequests from individuals in income group $j$ who died at the end of period $t-1$. $T_{j,s,t}$ are net taxes paid, which we specify more fully below in equations (10) through (14).

Implicit in the period budget constraint (8) is a strong assumption about the distribution of bequests. We assume that bequests are distributed evenly across all ages to those in the same lifetime income group. It is difficult to precisely calibrate the distribution of bequests from the data, both across income types $j$ and across ages $s$. However, the assumptions about the bequest motive as well as how bequests are distributed are clearly important modeling decisions. Our current specification of bequests is the most persistent, which should make wealth inequality more persistent relative to other bequest specifications.[10] A large number of papers study the effects of different bequest motives and specifications on the distribution of wealth, though there is no consensus regarding the true bequest transmission process.[11]

Because the form of the period utility function in (7) ensures that $b_{j,s,t} > 0$ for all $j$, $s$, and $t$, total bequests will always be positive $BQ_{j,t} > 0$ for all $j$ and $t$.

$$BQ_{j,t+1} = (1 + r_{t+1})\lambda_j \left( \sum_{s=E+1}^{E+S} \rho_s \omega_{s,t} b_{j,s+1,t+1} \right) \quad \forall j, t \tag{9}$$

In addition to each the budget constraint in each period, the utility function (7) imposes nonnegative consumption through infinite marginal utility, and the elliptical utility of leisure ensures individual labor and leisure must be strictly nonnegative $n_{j,s,t}, l_{j,s,t} > 0$. Because individual savings or wealth is always strictly positive, the aggregate capital stock is always positive.[12] An interior solution to the individual's problem (7) is assured.

The individual is subject to four types of taxes: an income tax on both capital and

[10] Another allocation rule at the opposite extreme would be to equally divide all bequests among all surviving individuals. An intermediate rule would be some kind of distribution of bequests with most going to ones own type and a declining proportion going to the other types.

[11] See De Nardi and Yang (2014), De Nardi (2004), Nishiyama (2002), Laitner (2001), Gokhale et al. (2000), Gale and Scholz (1994), Hurd (1989), Venti and Wise (1988), Kotlikoff and Summers (1981), and Wolff (2015).

[12] An alternative would be to allow for individual borrowing as long as an aggregate capital constraint $K_t > 0$ for all $t$ is satisfied.

labor income $T^I_{j,s,t}$, a payroll tax $T^P_{j,s,t}$, an inheritance tax $T^{BQ}_{j,t}$, and a wealth tax $T^W_{j,s,t}$. Every individual also receives an equal lump sum transfer $T^H_t$. The specifications of the tax functions are the following,

$$
\begin{aligned}
&T^I_{j,s,t} = \tau^I(\hat{a}_{j,s,t})a_{j,s,t} && (10)\\
&\qquad \text{where} \quad \hat{a}_{j,s,t} \equiv \frac{a_{j,s,t}}{e^{g_y t}} \quad \text{and} \quad a_{j,s,t} \equiv (r_t b_{j,s,t} + w_t e_{j,s} n_{j,s,t}) &&\\
&T^P_{j,s,t} = \begin{cases} \tau^P w_t e_{j,s} n_{j,s,t} & \text{if} \quad s < R \\ \tau^P w_t e_{j,s} n_{j,s,t} - \theta_j w_t & \text{if} \quad s \geq R \end{cases} && (11)\\
&T^{BQ}_{j,t} = \tau^{BQ}\frac{BQ_{j,t}}{\lambda_j \tilde{N}_t} && (12)\\
&T^W_{j,s,t} = \tau^W(\hat{b}_{j,s,t})b_{j,s,t}, \quad \text{where} \quad \hat{b}_{j,s,t} \equiv \frac{b_{j,s,t}}{e^{g_y t}} && (13)
\end{aligned}
$$

where $a_{j,s,t}$ is total income from labor and capital, $\hat{a}_{j,s,t}$ is stationarized total income, and $\tau^I(\hat{a}_{j,s,t})$ is the effective income tax rate on labor and capital income as a function of stationarized total income.[13] Similarly, $\tau^W(\hat{b}_{j,s,t})$ is the effective wealth tax rate as a function of stationary individual wealth $\hat{b}_{j,s,t}$.[14]

Because individual lifetime income type (and thus their life cycle earnings profile) are deterministic from birth, the Social Security replacement rate $\theta_j$ in the payroll tax (11) can be thought of as simply a percent of the age $R-1$ labor earnings. This replacement rate $\theta_j$ is indexed to current average wage $w_t$, and then the ability $j$-specific $\theta_j$ captures the percent consistent with the average replacement amount of each type. In this way, $e_{j,s}$ is incorporated into each $\theta_j$. $R$ is the age at which the individual becomes eligible to receive the retirement benefit from the payroll

[13] In Appendix A-4, we describe how we fit increasing and concave functions to the current data on deduction ratios and effective tax rates in the United States. We restrict the tax functions to be concave because any non-increasing or non-concave segments would result in multiple potential local maxima in the computation of the solution. Our approximation is very close to the actual effective tax rate schedule as a function of income. In addition, the tax rate function must be a function of stationary income $\hat{a}_{j,s,t}$, otherwise all individuals would eventually be at the highest effective income tax rate as time goes to infinity. The definition of stationary income $\hat{y}_{j,s,t}$ is in Table 1.

[14] We use the following functional form $\tau^W(b) = P\frac{Hb}{Hb+M}$. Our baseline calibration value for $P$ is zero because there is currently no wealth tax. But in our policy experiment, we calibrate the values of $H$, $M$, and $P$ such that the tax rate on average wealth in the steady state is 1 percent, the rate on highest steady state wealth is 2 percent, and the highest possible rate is 2.5 percent.

tax.[15] Given the perfect foresight ability process, this is equivalent to wage indexing an average index of monthly earnings (AIME) for each lifetime income group. The payroll tax rate is $\tau^P$, and the estate tax rate is $\tau^{BQ}$. Net taxes $T_{j,s,t}$ from a given individual in each period's budget constraint (8) are given by the following equation.

$$T_{j,s,t} = T^I_{j,s,t} + T^P_{j,s,t} + T^{BQ}_{j,t} + T^W_{j,s,t} - T^H_t \tag{14}$$

The solution to the lifetime maximization problem (7) of individual with ability $j$ subject to the per-period budget constraint (8) and the specification of taxes in (14) and (10) through (13) is a system of $2S$ Euler equations. The $S$ static first order conditions for labor supply $n_{j,s,t}$ are the following,

$$\begin{aligned}
&(c_{j,s,t})^{-\sigma}\left(w_t e_{j,s} - \frac{\partial T_{j,s,t}}{\partial n_{j,s,t}}\right) = e^{g_y t(1-\sigma)}\chi^n_s\left(\frac{b}{\tilde{l}}\right)\left(\frac{n_{j,s,t}}{\tilde{l}}\right)^{\upsilon-1}\left[1-\left(\frac{n_{j,s,t}}{\tilde{l}}\right)^{\upsilon}\right]^{\frac{1-\upsilon}{\upsilon}} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\forall j,t, \quad \text{and} \quad E+1\leq s\leq E+S \\
&\text{where} \quad c_{j,s,t} = (1+r_t)\,b_{j,s,t} + w_t e_{j,s}n_{j,s,t} + \frac{BQ_{j,t}}{\lambda_j\tilde{N}_t} - b_{j,s+1,t+1} - T_{j,s,t} \\
&\text{and} \quad \frac{\partial T_{j,s,t}}{\partial n_{j,s,t}} = w_t e_{j,s}\left[\tau^I\big(F\hat{a}_{j,s,t}\big) + \frac{F\hat{a}_{j,s,t}CD\big[2A(F\hat{a}_{j,s,t})+B\big]}{\big[A(F\hat{a}_{j,s,t})^2 + B(F\hat{a}_{j,s,t}) + C\big]^2} + \tau^P\right] \\
&\text{and} \quad b_{j,E+1,t} = 0 \quad \forall j,t
\end{aligned} \tag{15}$$

where the parameters of the effective tax rate function $A$, $B$, $C$, and $D$ are defined in (A.4.1) in Appendix A-4. The parameter $F$ is a positive constant that multiplies the stationary disposable income $\hat{a}_{j,s,t}$ to make the model income units match up with the data income units.

An individual also has $S-1$ dynamic Euler equations that govern his saving decisions, $b_{j,s+1,t+1}$, with the included precautionary bequest saving in case of unexpected

[15] We calibrate $R$ to be equivalent to age $E+s=65$ in the population. See Appendix A-5 for a description of the calibration of the payroll tax parameters $\tau^P$ and $\{\theta_j\}_{j=1}^J$.

death. These are given by:

$$
\begin{aligned}
(c_{j,s,t})^{-\sigma} = \rho_s\chi^b_j\bigl(b_{j,s+1,t+1}\bigr)^{-\sigma} + \beta(1-\rho_s)(c_{j,s+1,t+1})^{-\sigma}\Biggl[(1+r_{t+1}) - \frac{\partial T_{j,s+1,t+1}}{\partial b_{j,s+1,t+1}}\Biggr] \\
\forall j,t, \quad \text{and} \quad E+1\leq s\leq E+S-1 \\
\text{where} \quad \frac{\partial T_{j,s+1,t+1}}{\partial b_{j,s+1,t+1}} = ... \\
r_{t+1}\Biggl(\tau^I(F\hat{a}_{j,s+1,t+1}) + \frac{F\hat{a}_{j,s+1,t+1}CD\left[2A(F\hat{a}_{j,s+1,t+1})+B\right]}{\left[A(F\hat{a}_{j,s+1,t+1})^2+B(F\hat{a}_{j,s+1,t+1})+C\right]^2}\Biggr)... \\
+\tau^W(\hat{b}_{j,s+1,t+1}) + \frac{\hat{b}_{j,s+1,t+1}PHM}{\Bigl(H\hat{b}_{j,s+1,t+1}+M\Bigr)^2}
\end{aligned}
\tag{16}
$$

The parameters $P$, $H$, and $M$ characterize the progressive wealth tax function $\tau^W(\hat{b}) = P\frac{H\hat{b}}{H\hat{b}+M}$. In the baseline, the wealth tax is zero.[16] Lastly, Each individual also has one static first order condition for the last period of life $s = E + S$, which governs how much to bequeath to the following generation given that the individual will die with certainty. This condition is:

$$
(c_{j,E+S,t})^{-\sigma} = \chi^b_j(b_{j,E+S+1,t+1})^{-\sigma} \quad \forall j,t \tag{17}
$$

Define $\hat{\mathbf{\Gamma}}_t$ as the distribution of stationary individual savings across individuals at time $t$, including the intentional bequests of the oldest cohort.

$$
\hat{\mathbf{\Gamma}}_t \equiv \Bigl\{\bigl\{\hat{b}_{j,s,t}\bigr\}_{j=1}^J\Bigr\}_{s=E+2}^{E+S+1} \quad \forall t \tag{18}
$$

As will be shown in Section 2.5, the state in every period $t$ for the entire equilibrium system described in the stationary, non-steady-state equilibrium characterized in Definition 2 is the stationary distribution of individual savings $\hat{\mathbf{\Gamma}}_t$ from (18). Because individuals must forecast wages, interest rates, and aggregate bequests received in ev-

[16] In the wealth tax experiment, the parameters are calibrated so that the tax rate on average wealth in the steady state is one percent, the rate on the highest steady-state wealth is two percent, and the rate on the highest steady-state wealth is 2.5 percent.

ery period in order to solve their optimal decisions and because each of those future variables depends on the entire distribution of savings in the future, we must assume some individual beliefs about how the entire distribution will evolve over time. Let general beliefs about the future distribution of capital in period $t+u$ be characterized by the operator $\Omega(\cdot)$ such that:

$$\boldsymbol{\hat{\Gamma}^e_{t+u}} = \Omega^u\left(\boldsymbol{\hat{\Gamma}_t}\right) \quad \forall t, \quad u \geq 1 \tag{19}$$

where the $e$ superscript signifies that $\hat{\Gamma}^e_{t+u}$ is the expected distribution of wealth at time $t+u$ based on general beliefs $\Omega(\cdot)$ that are not constrained to be correct.[17]

## 2.3 Firm problem

A unit measure of identical, perfectly competitive firms exist in the economy. The representative firm is characterized by the following Cobb-Douglas production technology,

$$Y_t = ZK_t^\alpha\left(e^{g_y t}L_t\right)^{1-\alpha} \quad \forall t \tag{20}$$

where $Z$ is the measure of total factor productivity, $\alpha \in (0,1)$ is the capital share of income, $g_y$ is the constant growth rate of labor augmenting technological change, and $L_t$ is aggregate labor measured in efficiency units. The firm uses this technology to produce a homogeneous output which is consumed by individuals and used in firm investment. The interest rate $r_t$ paid to the owners of capital is the real interest rate net of depreciation. The real wage is $w_t$. The real profit function of the firm is the following.

$$\text{Real Profits} = ZK_t^\alpha\left(e^{g_y t}L_t\right)^{1-\alpha} - (r_t+\delta)K_t - w_t L_t \tag{21}$$

As in the individual budget constraint (8), note that the price output has been normalized to one.

Profit maximization results in the real wage, $w_t$, and the real rental rate of capital

[17] In Section 2.5 we will assume that beliefs are correct (rational expectations) for the stationary non-steady-state equilibrium in Definition 2.

$r_t$ being determined by the marginal products of labor and capital, respectively:

$$w_t = (1-\alpha)\frac{Y_t}{L_t} \quad \forall t \tag{22}$$

$$r_t = \alpha\frac{Y_t}{K_t} - \delta \qquad \forall t \tag{23}$$

## 2.4 Government fiscal policy

The government is represented by a balanced budget constraint. The government collects taxes from four sources ($T^I_{j,s,t}$, $T^P_{j,s,t}$, $T^{BQ}_{j,s,t}$ and $T^W_{j,s,t}$) from all individuals and divides total revenues equally among individuals in the economy to determine the lump-sum transfer.

$$T^H_t = \frac{1}{\tilde{N}_t}\sum_s\sum_j \omega_{s,t}\lambda_j\left(T^I_{j,s,t} + T^P_{j,s,t} + T^{BQ}_{j,s,t} + T^W_{j,s,t}\right) \tag{24}$$

Lump sum transfers have an impact on the distribution of income and wealth. However, since we constrain our policy experiments to have the same steady-state revenue impact, the changes in inequality in economic outcomes due to changes in government transfers is equivalent in each policy experiment in the steady-state.

## 2.5 Market clearing and stationary equilibrium

Labor market clearing requires that aggregate labor demand $L_t$ measured in efficiency units equal the sum of individual efficiency labor supplied $e_{j,s}n_{j,s,t}$. Capital market clearing requires that aggregate capital demand $K_t$ equal the sum of capital investment by individuals $b_{j,s,t}$. Aggregate consumption $C_t$ is defined as the sum of all individual consumptions, and aggregate investment is defined by the resource constraint $Y_t = C_t + I_t$ as shown in (27). That is, the following conditions must hold:

$$L_t = \sum_{s=E+1}^{E+S}\sum_{j=1}^{J}\omega_{s,t}\lambda_j e_{j,s}n_{j,s,t} \quad \forall t \tag{25}$$

$$K_t = \sum_{s=E+2}^{E+S+1}\sum_{j=1}^{J}\omega_{s-1,t-1}\lambda_j b_{j,s,t} \quad \forall t \tag{26}$$

$$\begin{split} Y_t = C_t + K_{t+1} - (1-\delta)K_t \quad \forall t \\ \text{where} \quad C_t \equiv \sum_{s=E+1}^{E+S}\sum_{j=1}^{J}\omega_{s,t}\lambda_j c_{j,s,t} \end{split} \tag{27}$$

The usual definition of equilibrium would be allocations and prices such that individuals optimize (15), (16), and (17), firms optimize (22) and (23), and markets clear (25) and (26). However, the variables in the equations characterizing the equilibrium are potentially non-stationary due to the growth rate in the total population $g_{n,t}$ each period coming from the cohort growth rates in (1) and from the deterministic growth rate of labor augmenting technological change $g_y$ in (20).

**Table 1: Stationary variable definitions**

| Sources of growth | | | Not |
|---|---|---|---|
| $e^{g_y t}$ | $\tilde{N}_t$ | $e^{g_y t}\tilde{N}_t$ | growing[a] |
| $\hat{c}_{j,s,t} \equiv \frac{c_{j,s,t}}{e^{g_y t}}$ | $\hat{\omega}_{s,t} \equiv \frac{\omega_{s,t}}{\tilde{N}_t}$ | $\hat{Y}_t \equiv \frac{Y_t}{e^{g_y t}\tilde{N}_t}$ | $n_{j,s,t}$ |
| $\hat{b}_{j,s,t} \equiv \frac{b_{j,s,t}}{e^{g_y t}}$ | $\hat{L}_t \equiv \frac{L_t}{\tilde{N}_t}$ | $\hat{K}_t \equiv \frac{K_t}{e^{g_y t}\tilde{N}_t}$ | $r_t$ |
| $\hat{w}_t \equiv \frac{w_t}{e^{g_y t}}$ | | $\hat{BQ}_{j,t} \equiv \frac{BQ_{j,t}}{e^{g_y t}\tilde{N}_t}$ | |
| $\hat{y}_{j,s,t} \equiv \frac{y_{j,s,t}}{e^{g_y t}}$ | | | |
| $\hat{T}_{j,s,t} \equiv \frac{T_{j,s,t}}{e^{g_y t}}$ | | | |

[a] The interest rate $r_t$ in (23) is already stationary because $Y_t$ and $K_t$ grow at the same rate. Individual labor supply $n_{j,s,t}$ is stationary.

Table 1 characterizes the stationary versions of the variables of the model in terms of the variables that grow because of labor augmenting technological change, population growth, both, or none. With the definitions in Table 1, it can be shown that the equations characterizing the equilibrium can be written in stationary form in the following way. The static and intertemporal first-order conditions from the individual's optimization problem corresponding to (15), (16), and (17) are the following:

$$
\begin{aligned}
&(\hat{c}_{j,s,t})^{-\sigma}\left(\hat{w}_t e_{j,s} - \frac{\partial \hat{T}_{j,s,t}}{\partial n_{j,s,t}}\right) = \chi^n_s\left(\frac{b}{\tilde{l}}\right)\left(\frac{n_{j,s,t}}{\tilde{l}}\right)^{\upsilon-1}\left[1-\left(\frac{n_{j,s,t}}{\tilde{l}}\right)^{\upsilon}\right]^{\frac{1-\upsilon}{\upsilon}} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall j,t, \quad \text{and} \quad E+1\leq s\leq E+S \\
&\quad\text{where}\quad \hat{c}_{j,s,t} = (1+r_t)\,\hat{b}_{j,s,t} + \hat{w}_t e_{j,s} n_{j,s,t} + \frac{\hat{BQ}_{j,t}}{\lambda_j} - e^{g_y}\hat{b}_{j,s+1,t+1} - \hat{T}_{j,s,t} \\
&\quad\text{and}\quad \frac{\partial \hat{T}_{j,s,t}}{\partial n_{j,s,t}} = \hat{w}_t e_{j,s}\left[\tau^I\big(F\hat{a}_{j,s,t}\big) + \frac{F\hat{a}_{j,s,t}CD\big[2A(F\hat{a}_{j,s,t})+B\big]}{\big[A(F\hat{a}_{j,s,t})^2+B(F\hat{a}_{j,s,t})+C\big]^2} + \tau^P\right] \\
&\quad\text{and}\quad \hat{b}_{j,E+1,t}=0 \quad \forall j,t
\end{aligned}
\tag{28}
$$

$$
\begin{aligned}
&(\hat{c}_{j,s,t})^{-\sigma} = ... \\
&e^{-g_y\sigma}\left(\rho_s\chi^b_j\big(\hat{b}_{j,s+1,t+1}\big)^{-\sigma} + \beta(1-\rho_s)(\hat{c}_{j,s+1,t+1})^{-\sigma}\left[(1+r_{t+1}) - \frac{\partial T_{j,s+1,t+1}}{\partial b_{j,s+1,t+1}}\right]\right) \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad \forall j,t, \quad \text{and} \quad E+1\leq s\leq E+S-1 \\
&\quad\text{where}\quad \frac{\partial T_{j,s+1,t+1}}{\partial b_{j,s+1,t+1}} = ... \\
&\qquad r_{t+1}\left(\tau^I(F\hat{a}_{j,s+1,t+1}) + \frac{F\hat{a}_{j,s+1,t+1}CD\,[2A(F\hat{a}_{j,s+1,t+1})+B]}{[A(F\hat{a}_{j,s+1,t+1})^2+B(F\hat{a}_{j,s+1,t+1})+C]^2}\right)... \\
&\qquad + \tau^W(\hat{b}_{j,s+1,t+1}) + \frac{\hat{b}_{j,s+1,t+1}PHM}{\Big(H\hat{b}_{j,s+1,t+1}+M\Big)^2}
\end{aligned}
\tag{29}
$$

$$
(\hat{c}_{j,E+S,t})^{-\sigma} = \chi^b_j e^{-g_y\sigma}(\hat{b}_{j,E+S+1,t+1})^{-\sigma} \quad \forall j,t \tag{30}
$$

The stationary firm first order conditions for optimal labor and capital demand corresponding to (22) and (23) are the following.

$$
\hat{w}_t = (1-\alpha)\frac{\hat{Y}_t}{\hat{L}_t} \quad \forall t \tag{31}
$$

$$r_t = \alpha\frac{\hat{Y}_t}{\hat{K}_t} - \delta = \alpha\frac{Y_t}{K_t} - \delta \quad \forall t \tag{23}$$

And the two stationary market clearing conditions corresponding to (25) and (26)—with the goods market clearing by Walras' Law—are the following.

$$\hat{L}_t = \sum_{s=E+1}^{E+S}\sum_{j=1}^{J}\hat{\omega}_{s,t}\lambda_j e_{j,s}n_{j,s,t} \quad \forall t \tag{32}$$

$$\hat{K}_t = \frac{1}{1+\tilde{g}_{n,t}}\left(\sum_{s=E+2}^{E+S+1}\sum_{j=1}^{J}\hat{\omega}_{s-1,t-1}\lambda_j\hat{b}_{j,s,t}\right) \quad \forall t \tag{33}$$

where $\tilde{g}_{n,t}$ is the growth rate in the working age population between periods $t-1$ and $t$ described in (5). It is also important to note the stationary version of the characterization of total bequests $BQ_{j,t+1}$ from (9) and for the government budget constraint in (24).

$$\hat{BQ}_{j,t+1} = \frac{(1+r_{t+1})\lambda_j}{1+\tilde{g}_{n,t}}\left(\sum_{s=E+1}^{E+S}\rho_s\hat{\omega}_{s,t}\hat{b}_{j,s+1,t+1}\right) \quad \forall j,t \tag{34}$$

$$\hat{T}^H_t = \sum_s\sum_j\hat{\omega}_{s,t}\lambda_j\left(\hat{T}^I_{j,s,t} + \hat{T}^P_{j,s,t} + \hat{T}^{BQ}_{j,s,t} + \hat{T}^W_{j,s,t}\right) \tag{35}$$

We can now define the stationary steady-state equilibrium for this economy in the following way.

---

**Definition 1** (**Stationary steady-state equilibrium**). A non-autarkic stationary steady-state equilibrium in the overlapping generations model with $S$-period lived agents and heterogeneous ability $e_{j,s}$ is defined as constant allocations $n_{j,s,t} = \bar{n}_{j,s}$ and $\hat{b}_{j,s+1,t+1} = \bar{b}_{j,s+1}$ and constant prices $\hat{w}_t = \bar{w}$ and $r_t = \bar{r}$ for all $j$, $s$, and $t$ such that the following conditions hold:

i. individuals optimize according to (28), (29), and (30),

ii. Firms optimize according to (31) and (23),

iii. Markets clear according to (32) and (33), and

iv. The population has reached its stationary steady state distribution $\bar{\omega}_s$ for all ages $s$, characterized in Appendix A-1.

The steady-state equilibrium is characterized by the system of $2JS$ equations and $2JS$ unknowns $\bar{n}_{j,s}$ and $\bar{b}_{j,s+1}$. Appendix A-6 details how to solve for the steady-state equilibrium. Because our qualitative results and conclusions are unchanged across the equilibrium time path of the economy from the baseline steady state to the new steady state after the policy change, we confine our definition of the non-steady-state equilibrium and its computational solution to Appendix A-7.[18]

# 3 Calibration

Table 2 shows the calibrated values for the exogenous variables and parameters. The appendices contain the details for the calibration of many of these values. Most of these values come from outside the model. However, two sets of parameters, the utility weight on the disutility of labor $\chi^n_s$ and the utility weight on bequests $\chi^b_j$, are chosen to match the steady-state values of the model with their real-world counterparts in the U.S. economy. We describe the calibration process for $\{\chi^n_s\}_{s=E+1}^{E+S}$ and $\{\chi^b_j\}_{j=1}^{7}$ in this Section.

In this model, if $\chi^n_s = 1$ for all ages $s$ and $\chi^b_j = 1$ for all $j$, then, relative to U.S. data, model individuals work too much in their old age and do not accumulate enough wealth through savings. With regard to labor supply, it is intuitive to allow the scale parameter $\chi^n_s$ to increase with age, representing an increasing disutility of labor that is not modeled anywhere else in the utility function. An hour of labor for an older person becomes more costly due to biological reasons related to aging. With regard to the distribution of wealth, it is intuitive that individuals at the high end of the earnings distribution in our model do not save as much as their real world counterparts given the deterministic earnings process in our model. They have no precautionary savings motive, only the warm-glow bequest motive for savings. One can view the assumption of heterogeneous utility weights as not just variation in preference across

[18] We can provide equilibrium time path solutions for each policy experiment upon request.

**Table 2: List of exogenous variables and baseline calibration values**

| Symbol | Description | Value |
|---|---|---|
| $\hat{\boldsymbol{\Gamma}}_1$ | Initial distribution of savings | $\bar{\boldsymbol{\Gamma}}$ |
| $N_0$ | Initial population | 1 |
| $\{\omega_{s,0}\}_{s=1}^{S}$ | Initial population by age | (see App. A-1) |
| $\{f_s\}_{s=1}^{S}$ | Fertility rates by age | (see App. A-1) |
| $\{i_s\}_{s=1}^{S}$ | Immigration rates by age | (see App. A-1) |
| $\{\rho_s\}_{s=1}^{S}$ | Mortality rates by age | (see App. A-1) |
| $\{e_{j,s}\}_{j,s=1}^{J,S}$ | Deterministic ability process | (see App. A-2) |
| $\{\lambda_j\}_{j=1}^{J}$ | Lifetime income group percentages | (see App. A-2) |
| $J$ | Number of lifetime income groups | 7 |
| $S$ | Maximum periods in economically active individual life | 80 |
| $E$ | Number of periods of youth economically outside the model | round $\left(\frac{S}{4}\right)$ |
| $R$ | Retirement age (period) | round $\left(\frac{9}{16}S\right)$ |
| $\tilde{l}$ | Maximum hours of labor supply | 1 |
| $\beta$ | Discount factor | $(0.96)^{\frac{80}{S}}$ |
| $\sigma$ | Coefficient of constant relative risk aversion | 3 |
| $b$ | Scale parameter in utility of leisure | (see App. A-3) |
| $\upsilon$ | Shape parameter in utility of leisure | (see App. A-3) |
| $k$ | constant parameter in utility of leisure | (see App. A-3) |
| $\chi^n_s$ | Disutility of labor level parameters | (see Sec. 3) |
| $\chi^b_j$ | Utility of bequests level parameters | (see Sec. 3) |
| $Z$ | Level parameter in production function | 1 |
| $\alpha$ | Capital share of income | 0.35 |
| $\delta$ | Capital depreciation rate | $1-(1-0.05)^{\frac{80}{S}}$ |
| $g_y$ | Growth rate of labor augmenting technological progress | $(1+0.03)^{\frac{80}{S}}-1$ |
| $A$ | Coefficient on squared term in $\tau^I(\cdot)$ | (see App. A-4) |
| $B$ | Coefficient on linear term in $\tau^I(\cdot)$ | (see App. A-4) |
| $C$ | Constant coefficient in $\tau^I(\cdot)$ | (see App. A-4) |
| $D$ | Level parameter for $\tau^I(\cdot)$ | (see App. A-4) |
| $F$ | Income factor for $\tau^I(\cdot)$ | (see App. A-4) |
| $\tau^P$ | Payroll tax rate | 0.15 |
| $\{\theta^j\}_{j=1}^{J}$ | Replacement rate by average income | (see App. A-5) |
| $\tau^{BQ}$ | Bequest (estate) tax rate | 0 |
| $P$ | Level parameter for $\tau^W(\cdot)$ | 0 |
| $H$ | Coefficient on linear term in $\tau^W(\cdot)$ | 1 |
| $M$ | Constant coefficient in $\tau^W(\cdot)$ | 1 |
| $T$ | Number of periods to steady state | 160 |
| $\nu$ | Dampening parameter for TPI | 0.2 |

households, but also as reflecting differences in family size, expectations of income growth, or other variations that are not explicitly modeled here.

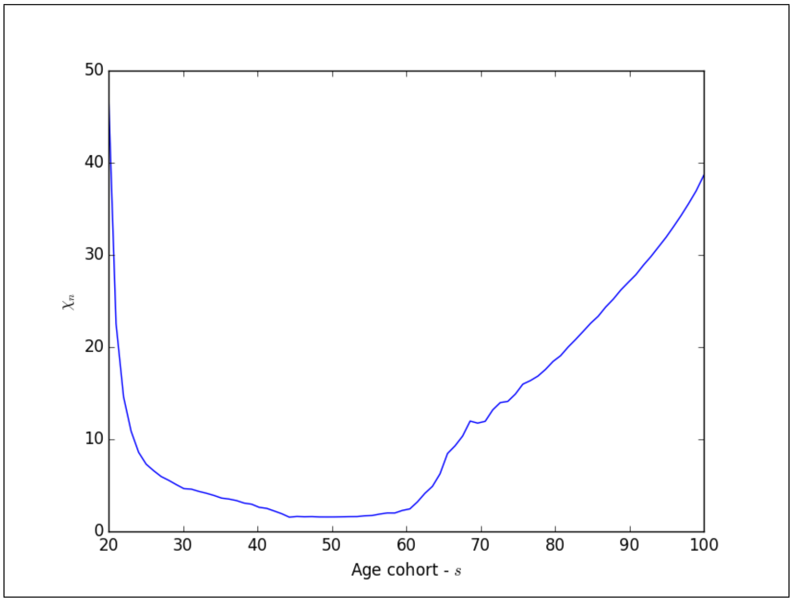
For the above two reasons, we choose the 87 parameter values $\{\chi^n_s\}_{s=E+1}^{E+S}$ and $\{\chi^b_j\}_{j=1}^{7}$ to match 94 moments from the data. We match the steady-state labor supply distribution by age from our model to the distribution of average hours by age in the United States (80 moments) and match the average wealth for individuals ages 21 to 45 and for ages 46 to 65 from our model for each of the seven income quantiles $\{\lambda_j\}_{j=1}^{7}$ to the corresponding average wealth levels from the data (14 moments). Specifically, we choose 87 parameters $\{\chi^n_s\}_{s=E+1}^{E+S}$ and $\{\chi^b_j\}_{j=1}^{7}$ to minimize the sum of squared percent deviations between 80 average steady-state labor supply values for each age 21 through 100 and 14 average steady-state wealth values (2 for each of the 7 lifetime earnings group quantiles).

Figure 2 shows the 80 calibrated values for $\{\chi^n_s\}_{s=E+1}^{E+S}$. The 7 calibrated values of $\{\chi^b_j\}_{j=1}^{7}$ are $[9.264\times 10^{-5}; 10.052; 90.841; 373.180; 1,738.031; 22,758.547; 118,648.915]$.[19] Figure 3 shows how closely the average steady-state labor supply by age from the model matches the average hours by age in the data, where the data for average labor supply after age 77 is a linear extrapolation (see Figure 4). Figure 5 shows how closely the distribution of wealth by age matches the data. Recall that the seven $\chi^b_j$ values were chosen to just match the average wealth for the two age groups (ages 21 to 45 and ages 46 to 65) in each income category. The calibration does not attempt to match the distribution of wealth from the data for individuals past age 65. As is described in De Nardi (2015), most models of the distribution of wealth have difficulty matching the wealth distribution of the highest income earners. Our model's ability to match the wealth concentration of the top one-percent shown in Figure 5 improves upon other studies.[20]

---

[19] Although the range of the values for $\{\chi^b_j\}_{j=1}^{7}$ is large, recall that those values in the warm glow bequest motive are discounted by the one-period conditional mortality rate $\rho_s$ (probability of dying next period), which is less than 0.01 for individuals alive at age $s <= 60$ and only rises above 0.10

**Figure 2: Calibrated values of $\chi^n_s$**




for ages $s >= 85$ with $\rho_{99}$=0.29 and $\rho_{100} = 1$ by force in the final period of life.

[20]Our goal in choosing the number of parameters to calibrate in order to match data moments is to use as few parameters as possible and still match key data moments. Theoretically, we could have chosen as many parameters as we have moments to identify. If those moments are identified from each other, then we would be able to exactly match the moments. However, this would be an example of overmodeling, the parameters would be less defensible as policy invariant, and potentially cause the model to lose relevance and predictive power for policy experiments.

**Figure 3: Life-cycle Average Labor Supply: Model vs. Data**

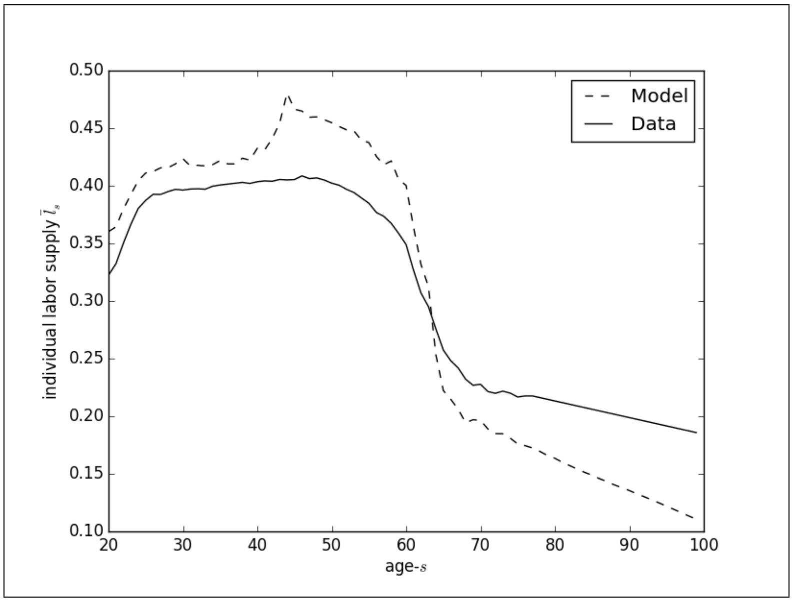



**Figure 4: Labor Distribution from Data, with extrapolation**

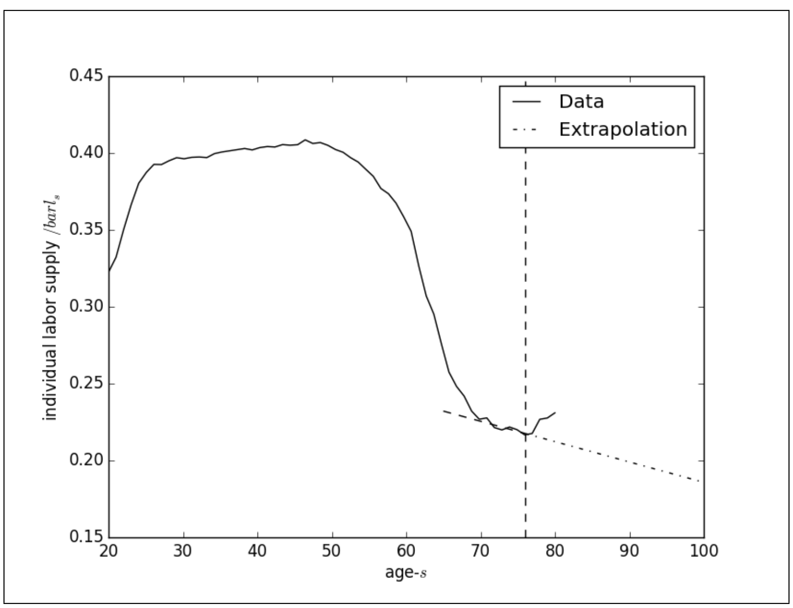
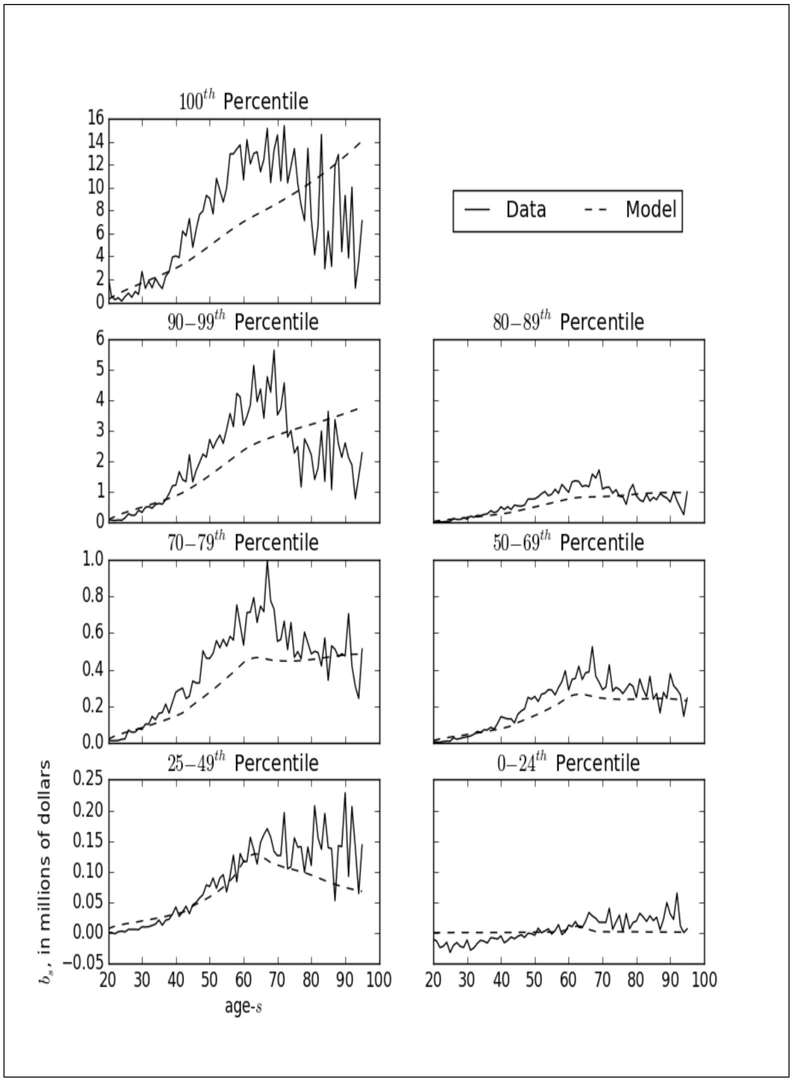


**Figure 5: Wealth over the life-cycle by age for each lifetime earnings group: Model vs. Data**



The average labor hours data from Figures 3 and 4 come from the Current Population Survey (CPS) March Supplement from 1992 to 2013. We determined hours worked in a year by the average hours worked per week in the last year and the number of weeks worked in the last year. We then compute mean hours worked by year of age using population weights. CPS on those near age 80 are noisy, thus we smooth the hours worked data for these year in the following way. We linearly fit the hours works fro ages 76 to 100 using the slope from hours worked for ages 60 to 76. For the average wealth by age and income group in Figure 5, we use the 2007, 2010, and 2013 Survey of Consumer Finances and obtain the distribution of total net worth by age using population weights.

Figure 6 shows the stationary steady-state distribution of individual labor supply $\bar{n}_{j,s}$ and Figure 7 shows the steady-state distribution of consumption $\bar{c}_{j,s}$ for the baseline calibration of the model described in Table 2. Notice from Figure 7 the hump-shaped pattern of consumption over the life cycle for each ability type, which is consistent with consumption data.

**Figure 6: Stationary steady-state distribution of individual labor supply $\bar{n}_{j,s}$ for $S = 80$ and $J = 7$** from baseline model

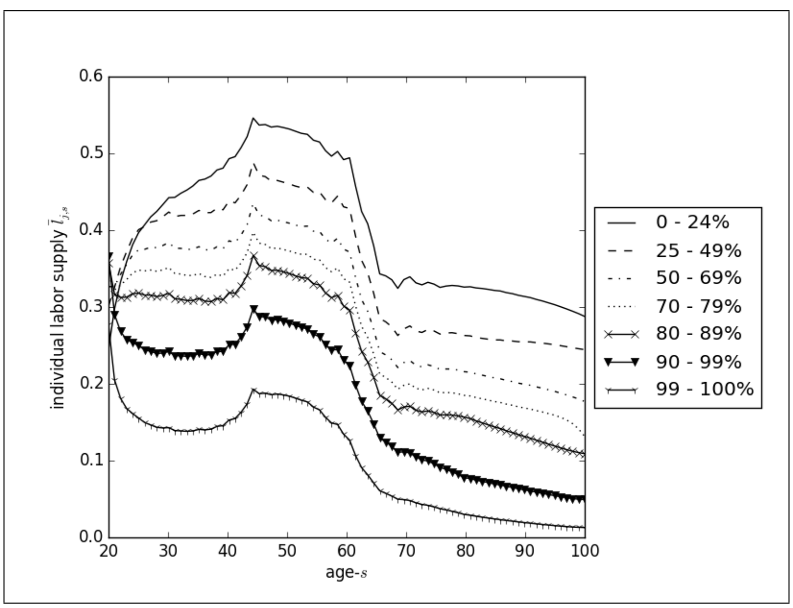
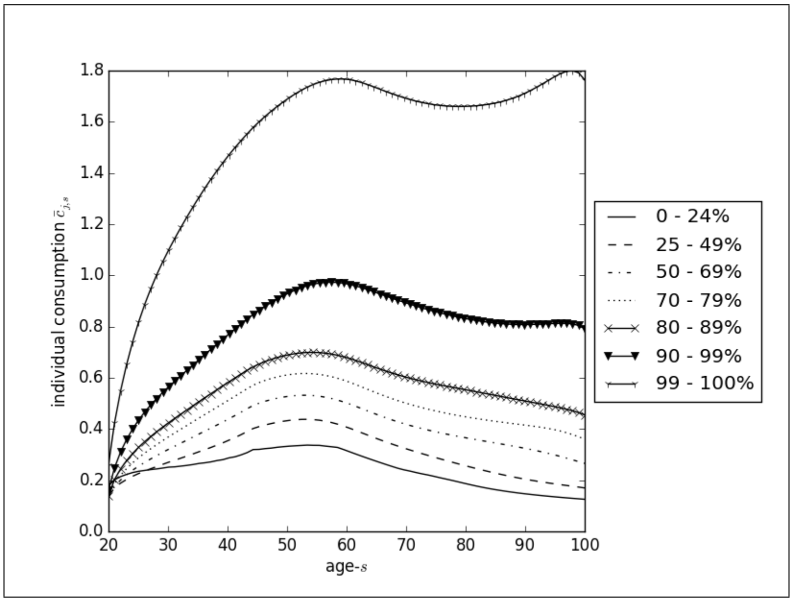


**Figure 7: Stationary steady-state distribution of consumption $\bar{c}_{j,s}$ for $S = 80$ and $J = 7$ in baseline model**




# 4 Wealth Tax versus the Income Tax

With our baseline model calibrated to the current characteristics of the U.S. economy, we test the distributional effects of two policy experiments. We compare the effectiveness at reducing inequality of the implementation of a progressive wealth tax schedule versus a more progressive income tax schedule. The wealth tax we study is similar to that proposed by Piketty (2014) where the tax basis is the individual's stock of wealth. We then study the effects of a separate and independent increase in the progressivity of the current income tax schedule which produces the same increase in steady-state revenue as the wealth tax experiment.

We study how each of these policies reduces inequality in wealth, income, consumption, and labor as measured by Gini coefficients. We assume that reducing inequality is a potential objective of a policymaker without specifying a particular social welfare function.[21] Because our model has individuals who are heterogeneous in terms of both ability ($e_{j,s}$, lifetime income profiles) and age $s$, total inequality is

[21]This is one potential weakness of much the inequality literature. Many studies do not carefully motivate why inequality should be reduced.

generated by variance across both of these dimensions. For this reason, we decompose the total inequality into two components; one reflecting inequality across lifetime income groups (averaging across ages) and the other component due to inequality across the life cycle (averaging across lifetime income groups).

Income and wealth taxes have potentially very different effects on individual decisions across the life-cycle and result in differential impacts to inequality across lifetime income groups and over the life-cycle. Our model will allow us to disentangle these effects, allowing us to not only characterize the extent to which tax policy affects total cross-sectional inequality, but also how taxes affect inequality across the two different dimensions of heterogeneity in our model.

## 4.1 Progressive wealth tax

As described in Section 2.2, the baseline wealth tax rate schedule is $\tau^W(\hat{b}_{j,s,t}) = 0$ for all $j$, $s$, and $t$ because the U.S. currently has no broad-based tax on wealth.[22] Piketty (2014, pp. 515-539) proposes a global tax on wealth in the form of a three-tier set of progressive wealth tax rates with a 2.0-percent tax rate on the highest levels of wealth, a 1.0- or 0.5-percent rate on middle incomes, and a 0.1- or 0.0-percent rate on the lowest levels of wealth.

In our experiment we calibrate a smooth progressive wealth tax rate function of stationary individual wealth $\hat{b}_{j,s,t}$ with the following functional form.

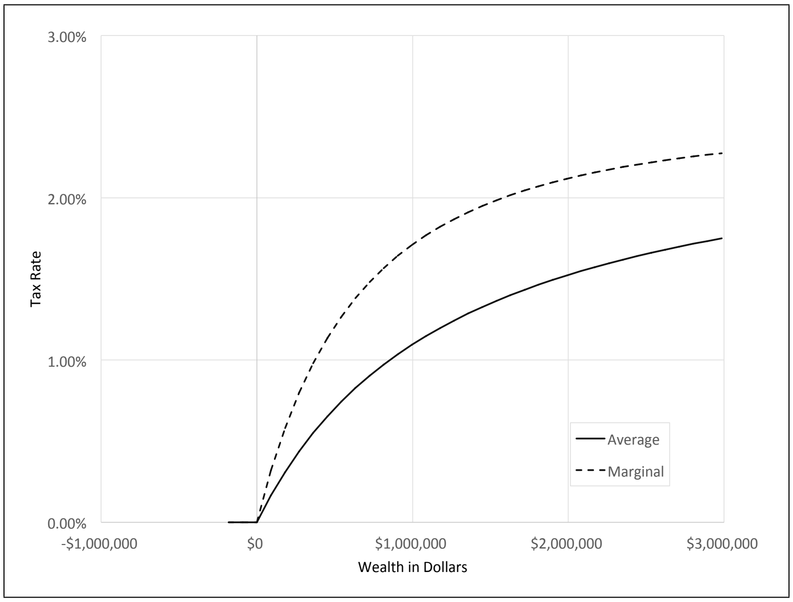
$$\tau^W\left(\hat{b}_{j,s,t}\right) = P\frac{H\hat{b}_{j,s,t}}{H\hat{b}_{j,s,t} + M} \tag{36}$$

As with the income tax, a smooth functional form is a computational necessity. We calibrate the parameters of the wealth tax (36) such that wealth equal to the top steady-state wealth level from Figure 5 has a tax rate of 2 percent, wealth equal to the average steady-state level has a tax rate of 1 percent, wealth equal to zero or less has a 0-percent rate, and the wealth tax asymptotes at 2.5-percent as wealth goes to

[22] The property tax is actually a tax on a particular type of wealth, but the U.S. does not have any broad based wealth tax. This is distinct from the portion of the income tax $\tau^I$ that taxes interest on wealth $r \times b$.

infinity. Figure 8 shows our calibrated wealth tax rate with $P = 0.025$, $H = 0.305509$, and $M = 2.16051$, as well as the implied marginal wealth tax rate.

**Figure 8: Calibrated wealth tax rate $\tau^W\left(\hat{b}_{j,s,t}\right)$ with $P = 0.025$, $H = 0.305509$, and $M = 2.16051$**




Note that the wealth tax and the capital income portion of the income tax are proportional in the steady state of this model where individuals cannot borrow, there is no aggregate uncertainty, and where returns on capital are identical across individuals. In particular, the steady-state relationship between the capital and labor tax is $\tau^W = \tau^k \bar{r}$, where $\tau^k$ is the tax on capital income. Thus, when the interest rate is low, wealth tax is equivalent to a very high capital income tax (perhaps in excess of 100%). This is relevant here since the income tax function we use represents taxes on total income. Therefore, when we consider a policy experiment that changes the progressivity of the income tax, there are changes to both the tax on capital and on labor income. The wealth tax represents only a tax on capital.

## 4.2 Progressive income tax

Figures 9 and 10 show our baseline income tax calibration and our more progressive income tax rate from the experiment. An increases in income taxes is also proposed by Piketty (2014, pp. 493-514) as a means of reducing inequality. This policy is accurately extolled as the most practical and politically feasible of his proposals. In other work, Piketty et al. (2014) estimate that the optimal marginal tax rate on the highest incomes is 82 percent when the social welfare function includes economic efficiency and distributional concerns.

As described in Appendix A-4, our baseline income tax specification calibrates a nonlinear continuously differentiable and concave function to the current effective tax rates implied by the U.S. tax code. The functional form for our baseline income tax takes the following form.

$$\tau^I\left(\hat{y}_{j,s,t}\right) = D\frac{A\hat{y}_{j,s,t}^2 + B\hat{y}_{j,s,t}}{A\hat{y}_{j,s,t}^2 + B\hat{y}_{j,s,t} + C} \tag{37}$$

We fit the functional form to the effective tax rates in the Individual Statistical Tables from the IRS Statistics of Income (SOI) program. The data on these effective tax rates—shown as the dotted line in Figure 9—include exemptions, deductions, filing statuses, and behavioral responses to the tax code.[23] Figure 10 shows the effective tax rate data, fitted effective tax rate function, and more progressive tax rate function in a non-logarithmic scale.

For our baseline income tax rate function, shown as the solid line in Figure 9, we calibrated values of $A = 3.03453 \times 10^{-6}$, $B = 0.222$, $C = 133,261$, and $D = 0.219$. Our policy experiment does not go as high as Piketty's suggested top marginal tax rate of 82 percent. Instead we test the effects of a more modest increase in the progressivity of the tax code. In order to make the wealth and income tax changes

[23]Note that we deliberately do not try to match the decrease in the effective tax rate for incomes over \$5 million. An income tax function with decreasing effective rates for some portion of the support of income would create a nonconvex budget set that would not be feasible to solve in our computational strategy.

**Figure 9: Calibrated effective income tax rate $\tau^{I}\left(\hat{y}_{j,s,t}\right)$, logarithmic income scale**

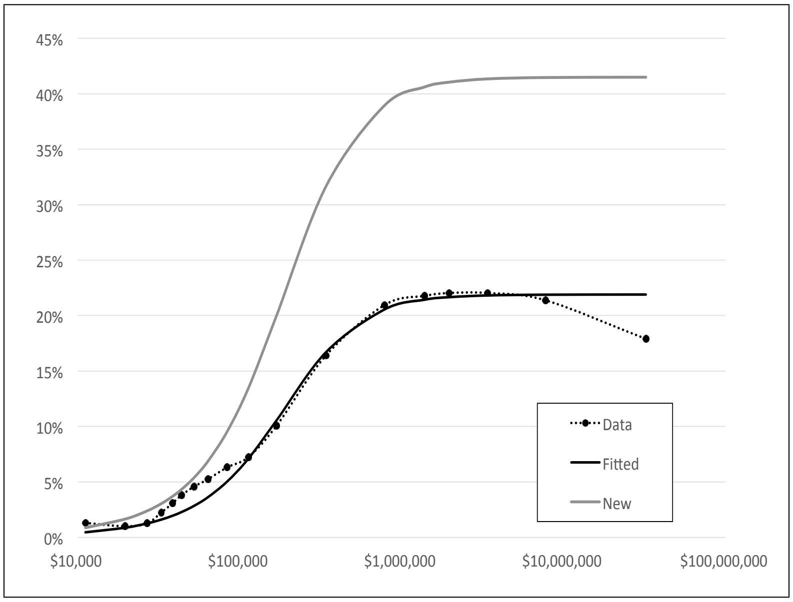
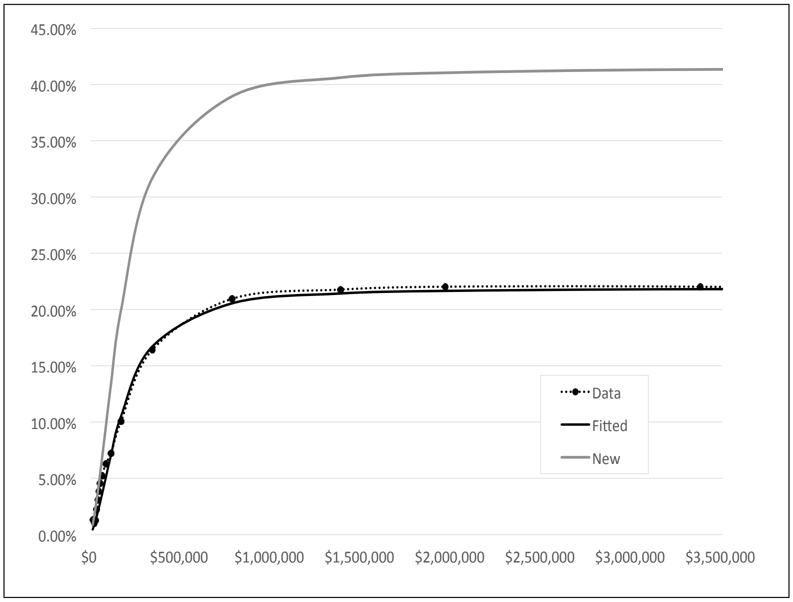



**Figure 10: Calibrated effective income tax rate $\tau^{I}\left(\hat{y}_{j,s,t}\right)$, non-logarithmic income scale**



comparable to one another, we ensure that both taxes generate the same amount of revenue. Increasing each effective income tax rate by 89.5 percent generates the same revenue as instituting the wealth tax ($D = 0.414782$). As shown in Figure 9, this implies that the effective income tax rate on annual income of $70,000 increases from roughly 2.5% to nearly 5.0%.

### 4.3 Effects of Taxes on Steady-state Inequality

In this section, we compare the effects of the wealth tax described in Section 4.1 and of the increased progressive income tax in Section 4.2 on inequality in wealth, income, consumption, and labor supply using Gini coefficients. We also show the respective steady-state effects of these two policies on the distributions of wealth, income, consumption, and labor supply.

Table 3 shows the baseline and treatment Gini coefficients and percent change for the wealth tax and the income tax experiments. Gini coefficients are shown for the steady-state distributions of wealth, income, consumption, and labor supply. Because our model has individuals who are heterogeneous across both ability ($e_{j,s}$, lifetime income profiles) and age $s$, total inequality is generated by variance across both of these dimensions. For this reason, we decompose each total Gini coefficient into the component from inequality across lifetime income groups (averaging across ages) and the component from inequality across the life cycle (averaging across lifetime income groups). Figure 11 shows the percent changes in steady-state wealth, income, consumption, and labor supply for all lifetime income groups and all ages for both the wealth tax and the income tax.

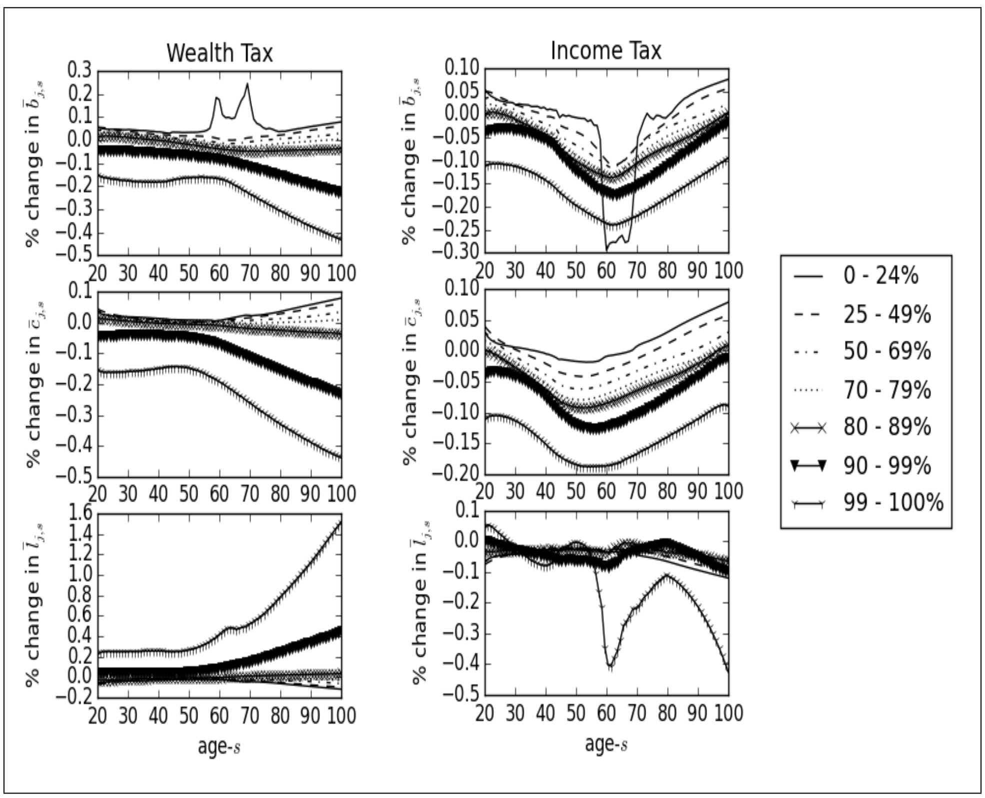
The strongest result from Table 3 is that the wealth tax reduces inequality more than the income tax across all measures of wealth, income, and consumption. The biggest percentage decreases in Gini coefficients for the wealth tax are in the distribution of income. Figure 11 shows that the wealth tax drastically reduces wealth and consumption for the top 10 percent. However, that reduction is much more pro-

## Table 3: Comparison of changes in steady-state Gini coefficients from wealth tax versus income tax

| Steady-state variable | Gini type | Baseline | Wealth tax Treatment | Wealth tax % Chg. | Income tax Treatment | Income tax % Chg. |
|---|---|---|---|---|---|---|
| $\bar{b}_{j,s}$ | Total | 0.943 | 0.929 | -1.48% | 0.939 | -0.42% |
| Wealth | Ability $j$ | 0.954 | 0.942 | -1.26% | 0.950 | -0.42% |
| | Age $s$ | 0.606 | 0.565 | -6.77% | 0.613 | 1.16% |
| $\bar{y}_{j,s}$ | Total | 0.775 | 0.733 | -5.42% | 0.757 | -2.32% |
| Income | Ability $j$ | 0.811 | 0.774 | -4.56% | 0.794 | -2.10% |
| | Age $s$ | 0.425 | 0.377 | -11.29% | 0.423 | -0.47% |
| $\bar{c}_{j,s}$ | Total | 0.664 | 0.621 | -6.48% | 0.644 | -3.01% |
| Cons- | Ability $j$ | 0.716 | 0.679 | -5.17% | 0.697 | -2.65% |
| umption | Age $s$ | 0.305 | 0.272 | -10.82% | 0.305 | 0.00% |
| $\bar{n}_{j,s}$ | Total | 0.240 | 0.258 | 7.50% | 0.236 | -1.67% |
| Labor | Ability $j$ | 0.324 | 0.349 | 7.72% | 0.321 | -0.93% |
| supply | Age $s$ | 0.145 | 0.145 | 0.00% | 0.142 | -2.07% |

Note: Under Gini type, *Total* refers to the Gini coefficient calculated from all the steady-state data, *Ability j* refers to the Gini coefficient calculated by averaging the data over the ages so we are measuring only inequality across lifetime income groups (ability), and *Age s* refers to the Gini coefficient calculated by averaging the data over the life cycle income groups so we are measuring only inequality across ages.

## Figure 11: Differences in steady-state distributions of tax policy versus baseline



nounced for individuals over the age of 60. For this reason, Table 3 shows that the reduction in inequality is large across income groups (Ability $j$) but is largest across the life cycle (Age $s$).

Table 3 also shows that the wealth tax increases inequality in the steady-state distribution of labor supply, and this increase comes primarily from increases in labor supply inequality across lifetime income groups (Ability $j$). The bottom-right corner of Figure 11 shows that the wealth tax causes the top one percent to work 20-percent more from ages 20 to 50, and then labor supply increases dramatically after 55. The increase in labor supply in old age also applies to the other top 20-percent of wage earners.

This suggests that, although the wealth tax is more effective than the income tax at reducing inequality for a given level of steady-state tax revenue, that reduction in inequality comes through a heavy tax on the wealthiest and oldest individuals in the economy. For the top 20-percent of wage earners in the economy, consumption decreases and labor supply increases by a large percentage. And that percentage increases dramatically after age 55.

In contrast, the more progressive income tax reduces inequality by more modest amounts. However, most of the reductions in inequality in wealth, income, and consumption come from reductions in inequality across lifetime earnings groups (Ability $j$) with very little change in inequality across age profiles (Age $s$). In addition, the biggest reductions in wealth and consumption from the income tax fall on individuals age 40 to 70 with less effect on the very young and the very old. This is because the estimated lifetime earnings profiles in Figure 1 are hump shaped. The largest incidence of the income tax will come at the ages associated with the highest wage.[24]

With regard to steady-state labor hours, the income tax does not have a large effect with any of the groups except for the top one percent of wage earners. The income tax causes a drastic reduction for them starting at about age 55 and hitting

[24]Note that the large decrease in wealth of the bottom quartile of wage earners illustrated in the top-right corner of Figure 11 is overstated. That group of agents has very little savings and wealth, but they accumulate a little wealth right around retirement in the baseline case. Because this amount is small, a small reduction from the income tax causes a large percentage decrease in their wealth.

its minimum with a 40-percent reduction in steady-state labor supply at age 60. The increased progressivity of the income tax falls primarily on middle aged agents and on the top 20 percent of wage earners. We test the robustness of these results to different values of the coefficient of relative risk aversion $\sigma$ in Appendix A-8, and the qualitative results remain the same across all specifications.

In summary, the wealth tax is extremely effective at reducing inequality in wealth, income, and consumption relative to an increase in the progressivity of the income tax with the same steady-state tax revenue. While the reductions in inequality across lifetime income groups from the wealth tax are significant, the wealth tax also reduces cross-sectional inequality through life cycle effects. In particular, it reduces the variance in consumption, wealth, and income within income groups over the life cycle. The costs of reducing inequality using the wealth tax are primarily borne by the top 10 percent of wage earners and even more predominantly by individuals over the age of 60.

The income tax has a smaller reduction in inequality, but to the extent it lowers inequality, the effects operate almost entirely by reducing inequality across lifetime income groups. The reductions in wealth and consumption from the income tax are focused primarily among the top 20 percent of wage earners and among middle aged individuals between the ages of 40 and 70. In addition, the income tax only has a small effect on the steady-state distribution of labor supply, with the exception of the top one percent of wage earners who reduce their labor supply significantly after age 55.

## 4.4 Efficiency Costs of Wealth Versus Income Taxation

With distortionary taxes, such as the wealth and income taxes proposed here, the changes in inequality resulting from the tax policies come at a potential cost to economic efficiency. It is therefore important to compare the wealth and income tax policy experiments in terms of their impacts measures of economic efficiency. Table 4 shows total income, capital stock, labor supply, and consumption in the baseline case and for the two policy experiments, as well as percentage changes from each policy.

The table also includes the effects on a utilitarian social welfare function in which total steady-state utility is the population weighted sum of individual utility levels.

**Table 4: Comparison of changes in steady-state aggregate variables from wealth tax versus income tax**

| Steady-state aggregate variable | Baseline | Wealth tax Treatment | Wealth tax % Chg. | Income tax Treatment | Income tax % Chg. |
|---|---|---|---|---|---|
| Income (GDP) $\bar{Y}$ | 0.503 | 0.489 | -2.78% | 0.474 | -5.77% |
| Capital stock $\bar{K}$ | 1.777 | 1.612 | -9.29% | 1.577 | -11.25% |
| Labor $\bar{L}$ | 0.299 | 0.299 | 0.00% | 0.289 | -3.34% |
| Consumption $\bar{C}^*$ | 0.414 | 0.408 | -1.45% | 0.396 | -4.35% |
| Total utility $\bar{U}^*$ | 6185.054 | 6234.131 | 0.79% | 6225.937 | 0.66% |

* Steady-state consumption $\bar{C}$ and total utility $\bar{U}$ are calculated as the population-weighted sum of steady-state individual consumptions and utilities for each individual of type $j$ and age $s$.

The effects on these economic aggregates strengthen the case for the wealth tax. Both taxes have distortionary effects that decrease economic efficiency, as evidenced through the declines in total income, the capital stock, labor supply, and consumption. However, the wealth tax has less of an impact on each of these economic aggregates. Because wealth is so concentrated among the top lifetime income groups, the distortionary effects on consumption and savings are largest for these groups, as Figure 11 shows. The lower lifetime income groups benefit from the increased government transfers resulting from the additional tax revenues in the policy experiments. The increases in consumption and savings from the larger share of the population in the lower income groups offset the sharp decline from the top lifetime income groups.

In terms of utility, both policy experiments have small, but positive, effects on the utilitarian social welfare function. Although the economic aggregates decline under the increases in taxes, the distributional effects help to increase total utility by reallocating consumption from the high income groups, with low marginal utility of consumption, to the lower income groups, who have higher marginal utilities from consumption.

## 4.5 Alternative Inequality Measures

Our baseline analysis above considers the effects of wealth and income taxes on steady-state inequality as measured by the Gini coefficient. While the Gini coefficient is widely used, it is only one particular measure of income inequality and shows a special sensitivity to the discretization of the distribution. To show that our results are robust to alternative measures of inequality, we consider four additional measures of inequality. These measures are the variance in logs, the 90/10 ratio, share of attributable to the top 10%, and the top 1% share. In addition to acting as a robustness test of the baseline results, these alternative measures are informative about what parts of the distribution are most affected by the policy experiments. Table 5 summarizes these results.

**Table 5: Changes in Alternative Inequality Measures, Wealth vs. Income Tax**

| Steady-state variable | Inequality Meaure | Baseline Value | Wealth tax Treatment | Wealth tax % Chg. | Income tax Treatment | Income tax % Chg. |
|---|---|---|---|---|---|---|
| $\bar{b}_{j,s}$ | $var(log(b_{j,s}))$ | 4.293 | 4.076 | -5.05% | 4.103 | -4.43% |
| Wealth | 90/10 ratio | 1395.75 | 1297.65 | -7.03% | 1240.26 | -11.14% |
| | Top 10% share | 0.664 | 0.633 | -4.67% | 0.647 | -2.56% |
| | Top 1% share | 0.189 | 0.160 | -15.34% | 0.176 | -6.88% |
| $\bar{y}_{j,s}$ | $var(log(y_{j,s}))$ | 0.987 | 1.058 | 7.19% | 0.963 | -2.43% |
| Income | 90/10 ratio | 3.790 | 3.615 | -4.62% | 3.438 | -9.28% |
| | Top 10% share | 0.292 | 0.269 | -7.87% | 0.272 | -6.85% |
| | Top 1% share | 0.062 | 0.053 | -14.52% | 0.056 | -9.68% |
| $\bar{c}_{j,s}$ | $var(log(c_{j,s}))$ | 1.210 | 1.348 | 11.40% | 1.210 | 0.00% |
| Cons- | 90/10 ratio | 3.023 | 2.881 | -4.70% | 2.735 | -9.53% |
| umption | Top 10% share | 0.217 | 0.200 | -7.83% | 0.203 | -6.45% |
| | Top 1% share | 0.039 | 0.032 | -17.95% | 0.035 | -10.26% |

Note: The 90/10 ratio for wealth is the ratio of the wealth of the individual at the 90th percentile (high wealth) to the wealth of the individual at the 10th percentile (low wealth). The Top 10% share is the share of total wealth held by the top 10 percent of wealth holders. The Top 1% share is the share of total wealth held by the top 1 percent of wealth holders.

Using alternative measures such as the variance in logs and top shares corroborate the Gini results in showing the wealth tax as a more effective tool of redistribution. In particular, the effects of the policy experiments on the top one percent's share show the wealth tax to have very concentrated effects on the extreme tails of the distributions. The only measure of inequality that declines to a greater degree under the income tax experiment than the wealth tax experiment is the 90/10 ratio. This inequality measure is not only affected by declines in that shares at the upper end of the distribution, but also accounts for increases in shares at the lower end, as it is affected through the 10th percentile in the denominator. Thus, this measure highlights how the income tax does more to not only lower wealth and income at the upper end of the distribution, but also keeps the shares towards the bottom of the distribution from falling to such a degree.

# 5 Conclusion

This paper constructs a large-scale overlapping generations model with heterogeneity across the life cycle and over lifetime income groups. We calibrate the model's parameters to match the behavior of the U.S. economy. With lifetime income groups calibrated from IRS tax returns, we are able to capture behavioral responses and inequality even in the tails of the distribution. Our top income group represents the top one-percent of earners.

We test the effectiveness of a wealth tax against an income tax with equal steady-state revenues at reducing inequality across wealth, income, and consumption. Our robust finding is that the wealth tax is extremely effective at reducing inequality in wealth, income, and consumption relative to an increase in the progressivity of the income tax with the same steady-state tax revenue. The reductions in inequality from the wealth tax are significant across lifetime income groups, and the effects on inequality over the life cycle are even stronger. The costs of reducing inequality using the wealth tax are primarily borne by the top 10 percent of wage earners and even more predominantly by individuals over the age of 60.

The income tax has a smaller reduction in inequality, but comes primarily in reductions in inequality across lifetime income groups. The reductions in wealth and consumption from the income tax are focused primarily among the top 20 percent of wage earners and among middle aged individuals between the ages of 40 and 70. In addition, the income tax only has a small effect on the steady-state distribution of labor supply, with the exception of the top one percent of wage earners who reduce their labor supply significantly after age 55.

It is important to note the limitations of using a closed economy model in this context. The previous analysis of the effects of the wealth tax demonstrate significant reductions in wealth and consumption among the highest wage earners accompanied by significant increases in labor supply in an attempt to smooth out consumption. If capital were mobile and individuals had the option to move to countries with more favorable tax treatments, this would have to dampen the effect of the wealth tax. While there is considerable disagreement about the extent to which capital is mobile, it is certainly not completely immobile as this this model presumes. Incorporating open economy components to this type of analysis will be an important extension of this line of research.

There are two other clear directions in which the analysis here can be extended. The first is to consider stochastic income processes. Individuals are subject to both permanent and transitory income shocks and the persistence of shocks affects savings decisions. Thus these income processes will have significant interactions with wealth and income taxation. The quantitative importance of shocks to total earnings have been highlighted by Guvenen et al. (2015) and DeBacker et al. (2013). DeBacker and Ramnath (2015) show the high degree of volatility in hourly earnings, even for the top one percent of the earning distribution. Extending the model outlined here to incorporate stochastic income processes would be valuable in further understanding the impacts of wealth versus income taxes. Second, one might consider the question of optimal income and wealth taxes in this framework, which benefits from being able incorporate life cycle behavior by those in the tails of the distribution of lifetime income.

# References


**Auerbach, Alan J. and Lawrence J. Kotlikoff**, *Dynamic Fiscal Policy*, Cambridge University Press, 1987.

**Cagetti, Marco and Mariacristina De Nardi**, "Wealth Inequality: Data and Models," *Macroeconomic Dynamics*, September 2008, *12* (Supplement S2), 285–313.

**Census Bureau**, "Annual Estimates of the Resident Population for Selected Age Groups by Sex for the United States, States, Counties, and Puerto Rico Commonwealth and Municipios: April 1, 2010 to July 1, 2013," Technical Report, U.S. Census Bureau, Population Division June 2014.

**Center for Disease Control**, "National Vital Statistics Reports," Technical Report, Center for Disease Control 2010.

**Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri**, "Adjustment Costs, Firm Responses, and Micro vs. Macro Labor Supply Elasticities: Evidence from Danish Tax Records," *The Quarterly Journal of Economics*, 2011, *126* (2), 749–804.

**De Nardi, Mariacristina**, "Wealth Inequality and Intergenerational Links," *Review of Economic Studies*, July 2004, *71* (3), 743–768.

_ , "Quantitative Models of Wealth Inequality: A Survey," NBER Working Paper No. 21106, National Bureau of Economic Research April 2015.

_ **and Fang Yang**, "Bequests and Heterogeneity in Retirement Wealth," *European Economic Review*, November 2014, *72*, 182–196.

**DeBacker, Jason and Shanthi Ramnath**, "Estimating the Hourly Earnings Processes of Top Earners," Technical Report, Mimeo 2015.

_ **, Bradley Heim, Vasia Panousi, Shanthi Ramnath, and Ivan Vidangos**, "Rising Inequality: Transitory or Persistent? New Evidence from a Panel of U.S. Tax Returns," *Brookings Papers on Economic Activity*, 2013, *46* (1 (Spring), 67–142.

**Evans, Richard W. and Kerk L. Phillips**, "OLG Life Cycle Model Transition Paths: Alternate Model Forecast Method," *Computational Economics*, January 2014, *43* (1), 105–131.

_ **and** _ , "Advantages of an Ellipse when Modeling Leisure Utility," *Computational Economics*, March 2018, *51* (3), 513–533.

**Gale, William G. and John Karl Scholz**, "Intergenerational Transfers and the Accumulation of Wealth," *Journal of Economic Perspectives*, Fall 1994, *8* (4), 145–160.

**Gokhale, Jagadeesh, Laurence J. Kotlikoff, James Sefton, and Martin Weale**, "Simulating the Transmission of Wealth Inequality via Bequests," *Journal of Public Economics*, January 2000, *79* (1), 93–128.

**Guvenen, Fatih, Fatih Karahan, Serdar Ozcan, and Jae Song**, "What Do Data on Millions of U.S. Workers Reveal About Life-cycle Earnings Risk?," NBER Working Paper No. 20913, National Bureau of Economic Research January 2015.

**Hurd, Michael D.**, "Mortality Risk and Bequests," *Econometrica*, July 1989, *57* (4), 779–813.

**IRS**, "Individual Statistical Tables, Publication 1304," Technical Report, Internal Revenue Service, Statistics of Income Division (SOI) July 2014.

**Kambourov, Gueorgui, Fatih Guvenen, Burhanettin Kuruscu, and Daphne Chen**, "Efficiency Gains from Wealth Taxation," 2013 Meeting Papers 1112, Society for Economic Dynamics 2013.

**Keane, Michael and Richard Rogerson**, "Micro and Macro Labor Supply Elasticities: A Reassessment of Conventional Wisdom," *Journal of Economic Literature*, 2012, *50.*

**King, Robert G., Charles I. Plosser, and Sergio T. Rebelo**, "Production, Growth and Business Cycles I. The Basic Neoclassical Model," *Journal of Monetary Economics*, 1988, *21.*

**Kocherlakota, Narayana R.**, "Advances in Dynamic Optimal Taxation," in Richard Blundell, Whitney K. Newey, and Torsten Persson, eds., *Advances in Economics and Econometrics: Theory and Applications, Ninth World Congress*, Vol. 1, Cambridge University Press, 2006, chapter 7, pp. 269–297.

**Kopczuk, Wojciech**, "What Do We Know about the Evolution of Top Wealth Shares in the United States?," *Journal of Economic Perspectives*, Winter 2015, *29* (1), 47–66.

**Kotlikoff, Laurence J. and Lawrence H. Summers**, "The Role of Intergenerational Transfers in Aggregate Capital Accumulation," *Journal of Political Economy*, August 1981, *89* (4), 706–732.

**Laitner, John**, "Secular Changes in Wealth Inequality and Inheritance," *Economic Journal*, October 2001, *111* (474), 691–721.

**Lochner, Lance and Youngki Shin**, "Understanding Earnings Dynamics: Identifying and Estimating the Changing Roles of Unobserved Ability, Permanent and Transitory Shocks," NBER Working Papers 20068, National Bureau of Economic Research, Inc April 2014.

**Mertens, Karel and Morten O. Ravn**, "Understanding the aggregate effects of anticipated and unanticipated tax policy shocks," *Review of Economic Dynamics*, 2011, *14*.

**Mirrlees, J. A.**, "An Exploration in the Theory of Optimum Income Taxation," *Review of Economic Studies*, April 1971, *38* (114), 175–208.

**Nishiyama, Shinichi**, "Bequests, Inter Vivos Transfers, and Wealth Distribution," *Review of Economic Dynamics*, October 2002, *5* (4), 892–931.

**Peterman, William B.**, "Reconciling Micro And Macro Estimates Of The Frisch Labor Supply Elasticity," *Economic Inquiry*, January 2016, *54* (1), 100–120.

**Piketty, Thomas**, "On the Long-Run Evolution of Inheritance: France 1820–2050," *The Quarterly Journal of Economics*, 2011, *126* (3), 1071–1131.

_ , *Capital in the Twenty-First Century*, Harvard University Press, 2014.

_ **and Emmanuel Saez**, "Income Inequality In The United States, 1913-1998," *The Quarterly Journal of Economics*, February 2003, *118* (1), 1–39.

_ **, _ , and Stefanie Stantcheva**, "Optimal Taxation of Top Labor Incomes: A Tale of Three Elasticities," *American Economic Journal: Economic Policy*, February 2014, *6* (1), 230–271.

**Saez, Emmanuel and Gabriel Zucman**, "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data," *The Quarterly Journal of Economics*, 2016, *131* (2), 519–578.

**Social Security Administration**, "Actuarial Life Table," Technical Report, U.S. Social Security Administration 2010.

**Stokey, Nancy L. and Robert E. Lucas Jr.**, *Recursive Methods in Economic Dynamics*, Harvard University Press, 1989.

**Suzumura, Kotaro**, "Perron-Frobenius Theorm on Non-Negative Square Matrices: An Elementary Proof," *Hitotsubashi Journal of Economics*, 1983, *24*, 137–141.

**Venti, Steven F. and David A. Wise**, "The Cause of Wealth Dispersion at Retirement: Choice or Chance?," *American Economic Review*, May 1988, *88* (2), 185–191.

**Wolff, Edward N.**, "Changing Inequality of Wealth," *American Economic Review*, May 1992, *82* (2), 552–558.

_ , "Recent Trends in the Size Distribution of Household Wealth," *Journal of Economic Perspectives*, Summer 1998, *12* (3), 131–150.

_ , *Inheriting Wealth in America: Future Boom or Bust?*, Oxford University Press, 2015.

# APPENDIX

## A-1 Characteristics of exogenous population growth assumptions

In this appendix, we describe in detail the exogenous population growth assumptions in the model and their implications. In Section 2.1, we define the laws of motion for the population of each cohort $\omega_{s,t}$ to be the following.

$$\begin{split}\omega_{1,t+1} &= \sum_{s=1}^{E+S} f_s\omega_{s,t} \quad \forall t \\ \omega_{s+1,t+1} &= (1+i_s-\rho_s)\omega_{s,t} \quad \forall t \quad \text{and} \quad 1\leq s\leq E+S-1\end{split} \tag{1}$$

We can transform the nonstationary equations in (1) into stationary laws of motion by dividing both sides by the total populations $N_t$ and $N_{t+1}$ in both periods,

$$\begin{split}\hat{\omega}_{1,t+1} &= \frac{\sum_{s=1}^{E+S} f_s\hat{\omega}_{s,t}}{1+g_{n,t+1}} \quad \forall t \\ \hat{\omega}_{s+1,t+1} &= \frac{(1+\phi_s-\rho_s)\hat{\omega}_{s,t}}{1+g_{n,t+1}} \quad \forall t \quad \text{and} \quad 1\leq s\leq E+S-1\end{split} \tag{A.1.1}$$

where $\hat{\omega}_{s,t}$ is the percent of the total population in age cohort $s$ and the population growth rate $g_{n,t+1}$ between periods $t$ and $t+1$ is defined in (3),

$$\begin{bmatrix}\hat{\omega}_{1,t+1} \\ \hat{\omega}_{2,t+1} \\ \hat{\omega}_{2,t+1} \\ \vdots \\ \hat{\omega}_{E+S-1,t+1} \\ \hat{\omega}_{E+S,t+1}\end{bmatrix} = \frac{1}{1+g_{n,t+1}} \times ...$$

$$\begin{bmatrix} f_1 & f_2 & f_3 & \dots & f_{E+S-1} & f_{E+S} \\ 1+i_1-\rho_1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1+i_2-\rho_2 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1+i_3-\rho_3 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & 0 & \dots & 1+i_{E+S-1}-\rho_{E+S-1} & 0\end{bmatrix}\begin{bmatrix}\hat{\omega}_{1,t} \\ \hat{\omega}_{2,t} \\ \hat{\omega}_{2,t} \\ \vdots \\ \hat{\omega}_{E+S-1,t} \\ \hat{\omega}_{E+S,t}\end{bmatrix} \tag{A.1.2}$$

where we restrict $1+i_s-\rho_s\geq 0$ for all $s$.

We write (A.1.2) in matrix notation as the following.

$$\boldsymbol{\hat{\omega}}_{t+1} = \frac{1}{1+g_{n,t+1}}\boldsymbol{\Omega}\boldsymbol{\hat{\omega}}_t \quad \forall t \tag{A.1.3}$$

The stationary steady state population distribution $\bar{\boldsymbol{\omega}}$ is the eigenvector $\boldsymbol{\omega}$ with eigenvalue $(1+\bar{g}_n)$ of the matrix $\boldsymbol{\Omega}$ that satisfies the following version of (A.1.3).

$$(1+\bar{g}_n)\bar{\boldsymbol{\omega}} = \boldsymbol{\Omega}\bar{\boldsymbol{\omega}} \tag{A.1.4}$$

**Proposition 1.** There exists a unique positive real eigenvector $\bar{\omega}$ of the matrix $\boldsymbol{\Omega}$, and it is a stable equilibrium.

*Proof.* First, note that the matrix $\boldsymbol{\Omega}$ is square and non-negative. This is enough for a general version of the Perron-Frobenius Theorem to state that a positive real eigenvector exists with a positive real eigenvalue. This is not yet enough for uniqueness. For it to be unique by a version of the Perron-Fobenius Theorem, we need to know that the matrix is irreducible. This can be easily shown. The matrix is of the form

$$\boldsymbol{\Omega} = \begin{bmatrix} * & * & * & \dots & * & * & * \\ * & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & * & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & * & 0 & 0 \\ 0 & 0 & 0 & \dots & 0 & * & 0 \end{bmatrix}$$

Where each * is strictly positive. It is clear to see that taking powers of the matrix causes the sub-diagonal positive elements to be moved down a row and another row of positive entries is added at the top. None of these go to zero since the elements were all non-negative to begin with.

$$\boldsymbol{\Omega^2} = \begin{bmatrix} * & * & * & \dots & * & * & * \\ * & * & * & \dots & * & * & * \\ * & 0 & 0 & \dots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & * & 0 & 0 \end{bmatrix} ; \quad \boldsymbol{\Omega^{S+E-1}} = \begin{bmatrix} * & * & * & \dots & * & * & * \\ * & * & * & \dots & * & * & * \\ * & * & * & \dots & * & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ * & * & * & \dots & * & * & * \\ * & 0 & 0 & \dots & 0 & 0 & 0 \end{bmatrix}$$

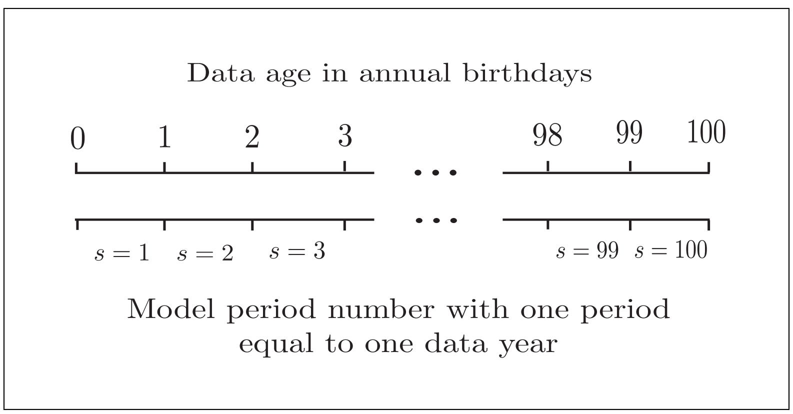
$$\boldsymbol{\Omega^{S+E}} = \begin{bmatrix} * & * & * & \dots & * & * & * \\ * & * & * & \dots & * & * & * \\ * & * & * & \dots & * & * & * \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ * & * & * & \dots & * & * & * \\ * & * & * & \dots & * & * & * \end{bmatrix}$$

Existence of an $m \in \mathbb{N}$ such that $\left(\boldsymbol{\Omega^m}\right)_{ij} \neq 0 \quad (> 0)$ is one of the definitions of an irreducible (primitive) matrix. It is equivalent to saying that the directed graph associated with the matrix is strongly connected. Now the Perron-Frobenius Theorem for irreducible matrices gives us that the equilibrium vector is unique.

We also know from that theorem that the eigenvalue associated with the positive real eigenvector will be real and positive. This eigenvalue, $p$, is the Perron eigenvalue and it is the steady state population growth rate of the model. By the PF Theorem for irreducible matrices, $|\lambda_i| \leq p$ for all eigenvalues $\lambda_i$ and there will be exactly $h$ eigenvalues that are equal, where $h$ is the period of the matrix. Since our matrix $\boldsymbol{\Omega}$ is aperiodic, the steady state growth rate is the unique largest eigenvalue in magnitude. This implies that almost all initial vectors will converge to this eigenvector under iteration. $\square$

For a full treatment and proof of the Perron-Frobenius Theorem, see Suzumura (1983). Because the population growth process is exogenous to the model, we calibrate it to annual age data for age years $s = 1$ to $s = 100$. As is shown in Figure 12, period $s = 1$ corresponds to the first year of life between birth and when an individual turns one year old.

**Figure 12: Correspondence of model timing to data timing for model periods of one year**




Our initial population distribution $\{\omega_{s,1}\}_{s=1}^{100}$ in Figure 13 comes from Census Bureau (2014) population estimates for both sexes for 2013. The fertility rates $\{f_s\}_{s=1}^{100}$ in Figure 14 come from Center for Disease Control (2010, Table 1). The mortality rates $\{\rho_s\}_{s=1}^{99}$ in Figure 15 come from the 2010 death probabilities in Social Security Administration (2010). We enforce a strict maximum age mortality rate of $\rho_{100} = 1$ in our model.

The immigration rates $\{i_s\}_{s=1}^{99}$ in Figure 16 are essentially residuals. We take total population for two consecutive years $N_t$ and $N_{t+1}$ and the population distribution by age in both of those years $\boldsymbol{\omega}_t$ and $\boldsymbol{\omega}_{t+1}$from the Census Bureau (2014) data. We then deduce the immigration rates $\{i_s\}_{s=1}^{99}$ using equation (A.1.1). We do this for three consecutive sets of years, so that our calibrated immigration rates by age are the average of our three years of deduced rates from the data for each age.

**Figure 13: Initial population distribution $\omega_{s,1}$ by year, $1 \leq s \leq 100$**

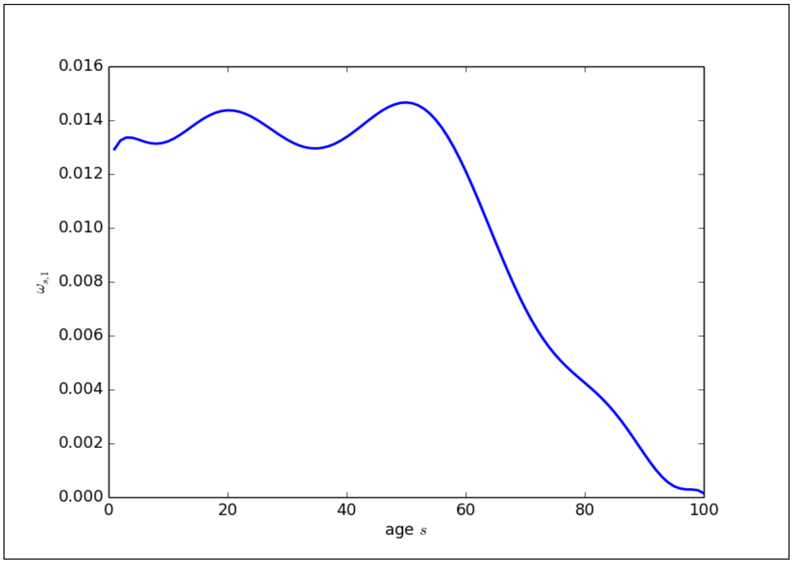



**Figure 14: Fertility rates $f_s$ by year, $1 \leq s \leq 100$**

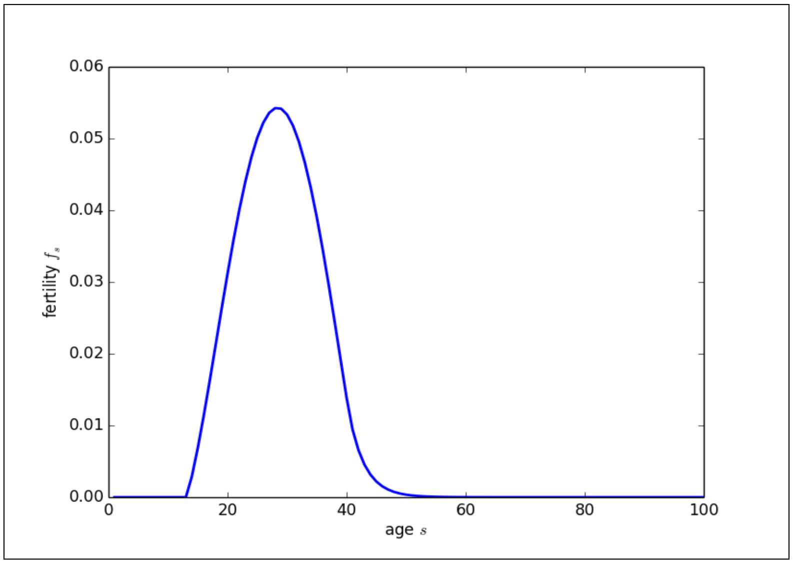


**Figure 15: Mortality rates $\rho_s$ by year, $1 \leq s \leq 100$**

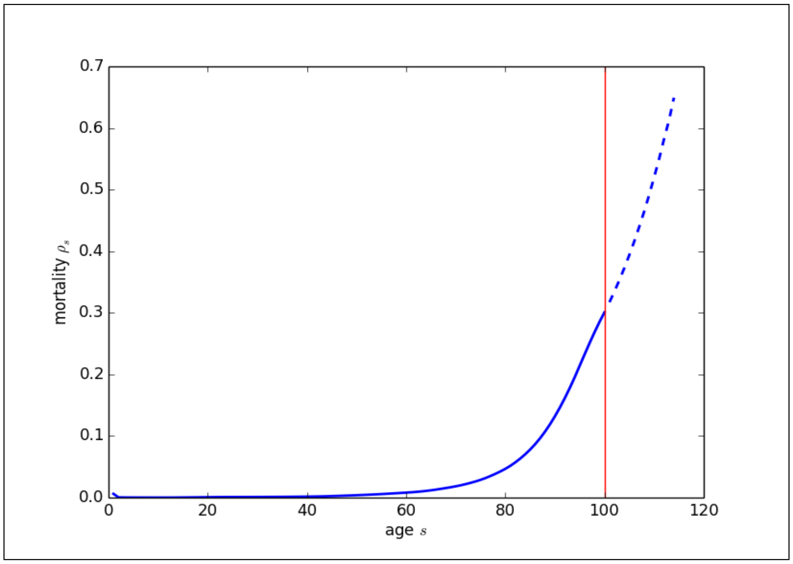



**Figure 16: Immigration rates $i_s$ by year, $1 \leq s \leq 100$**

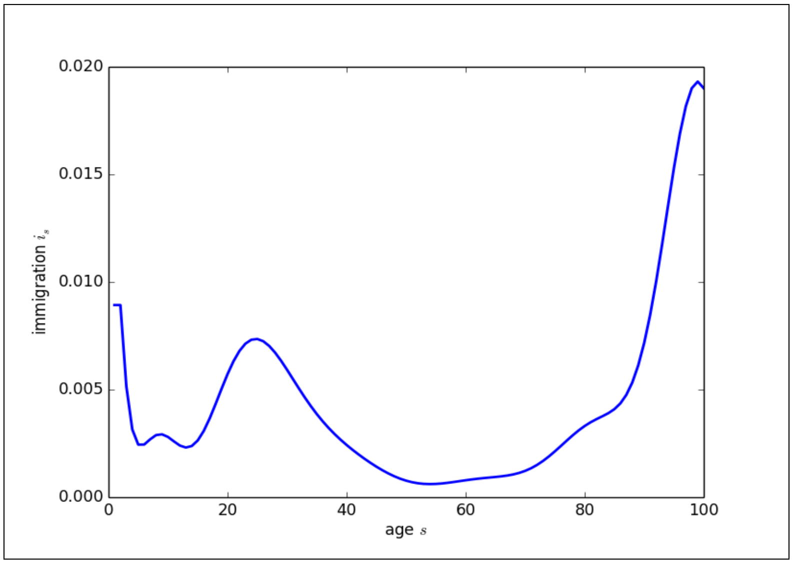


Figure 17 shows the predicted time path of the total population $N_t$ given $\omega_{s,1}$ $f_s$, $i_s$, and $\rho_s$. Notice that the population approaches a constant growth rate. This is a result of the stationary population percent distribution $\bar{\boldsymbol{\omega}}$ eventually being reached. Figure 18 shows the steady-state population percent distribution by age $\bar{\boldsymbol{\omega}}$.

**Figure 17: Forecast time path of population growth rate $g_{n,t}$**

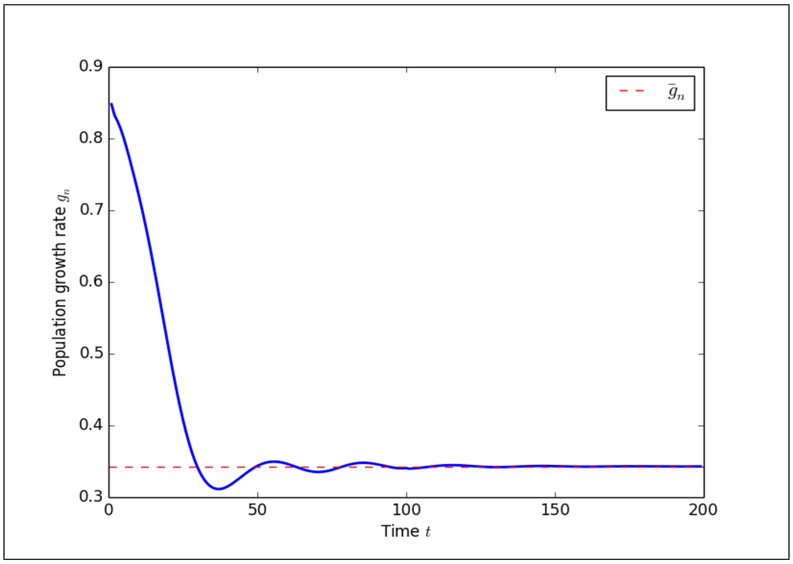
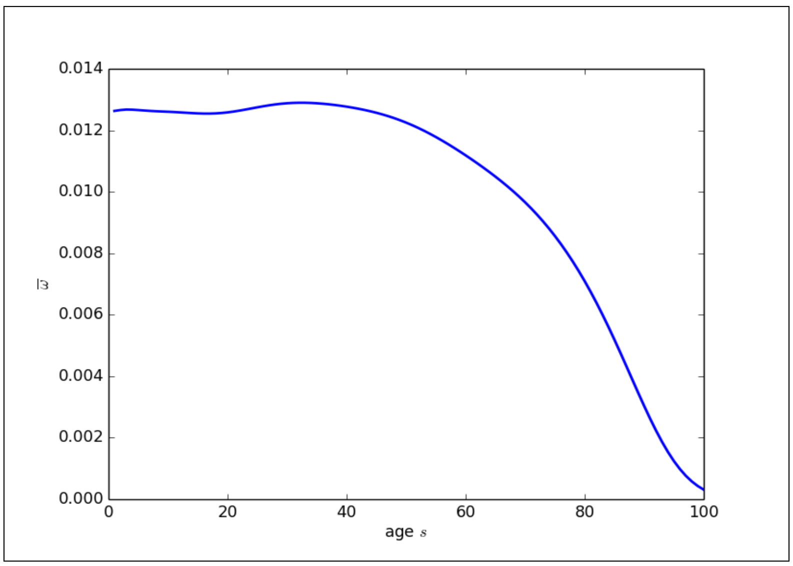


**Figure 18: Steady-state population percent distribution by age $\bar{\omega}$**



# A-2 Calibration of Lifetime Income Group Ability Profiles

We calibrate the model such that each lifetime income group has a different life-cycle profile of earnings. Since the distribution on income and wealth are key aspects of our model, we calibrate these processes so that we can represent earners in the top 1% of the distribution of lifetime income. It is income and wealth attributable to these households that has shown the greatest growth in recent decades (see, for example, Piketty and Saez (2003)). In order to have observations on the earnings of those at very top of the distribution that are not subject to top-coding we use data from the Internal Revenue Services's (IRS) Statistics of Income program (SOI).

## A-2.1 Continuous Work History Sample

The SOI data we draw from are the Continuous Work History Sample (CWHS). From this CWHS, we use a panel that is a 1-in-5000 random sample of tax filers from 1991 to 2009. For each filer-year observation we are able to observe detailed information reported on Form 1040 and the associated forms and schedules. We are also able to merge these tax data with Social Security Administration (SSA) records to get information on the age and gender of the primary and secondary filers. Our model variable of effective labor units maps into wage rates, because the market wage rate in the model, $w_t$, is constant across households. Earnings per hour thus depend upon effective labor units and equal $e_{j,s,t} \times w_t$ for household in lifetime income group $j$, with age $s$, in year $t$. Income tax data, however, do not contain information on hourly earnings or hours works. Rather, we only observe total earned income (wage and salaries plus self-employment income) over the tax year. In order to find hourly earnings for tax filers, we use an imputation procedure. This is described in detail in DeBacker and Ramnath (2015). The methodology applies an imputation for hours worked for a filing unit based on a model of hours worked for a filing unit estimated from the Current Population Survey (CPS) for the years 1992-2010.[25] We then use the imputed hours to calculate hourly earnings rates for tax filing units in the CWHS.

### A-2.1.1 Sample Selection

We exclude from our sample filer-year observations with earned income (wages and salaries plus business income) of less than \$1,250. We further exclude those with positive annual wages, but with hourly wages below \$5.00 (in 2005\$). We also drop one observation where the hourly wage rate exceeds \$25,000.[26] Economic life in the model runs from age 21 to 100. Our data have few observations on filers with ages

[25]The CPS survey asks retrospective questions about income in the last year and average hours worked per week (and weeks worked) in the last year). Therefore, these CPS surveys line up with tax years 1991-2009.

[26]This threshold is equivalent to \$50 million of wage income in one year at full time (40 hours per week) of work.

exceeding 80 years old. Our sample is therefore restricted to those from ages 21 to 80. After these restrictions, our final sample size is 333,381 filer-year observations.

### A-2.2 Lifetime Income

In our model, labor supply and savings, and thus lifetime income, are endogenous. We therefore define lifetime income as the present value of lifetime labor endowments and not the value of lifetime labor earnings. Note that our data are at the tax filing unit. We take this unit to be equivalent to a household. Because of differences in household structure (i.e., singles versus couples), our definition of lifetime labor income will be in per adult terms. In particular, for filing units with a primary and secondary filer, our imputed wage represents the average hourly earnings between the two. When calculating lifetime income we assign single and couple households the same labor endowment. This has the effect of making our lifetime income metric a per adult metric, there is therefore not an over-representation of couple households in the higher lifetime income groups simply because their time endowment is higher than for singles. We use the following approach to measure the lifetime income.

First, since our panel data do not allow us to observe the complete life cycle of earnings for each household (because of sample attrition, death or the finite sample period of the data), we use an imputation to estimate wages in the years of the household's economic life for which they do not appear in the CWHS. To do this, we estimate the following equation, separately by household type (where household types are single male, single female, couple with male head, or couple with female head) :

$$ln(w_{i,t}) = \alpha_i + \beta_1 age_{i,t} + \beta_2 age_{i,t}^2 + \beta_3 * age_{i,t}^3 + \varepsilon_{i,t} \tag{A.2.1}$$

**Table 6:** Initial Log Wage Regressions

| | Single Males | Single Females | Married, Male Head | Married, Female Head |
|---|---|---|---|---|
| *Age* | 0.177*** | 0.143*** | 0.134*** | 0.065** |
| | (0.006) | (0.005) | (0.004) | (0.027) |
| $Age^2$ | -0.003*** | -0.002*** | -0.002*** | -0.000 |
| | (0.000) | (0.000) | (0.000) | (0.001) |
| $Age^3$ | 0.000*** | 0.000*** | 0.000*** | 0.000 |
| | (0.000) | (0.000) | (0.000) | (0.000) |
| Observations | 88,833 | 96,670 | 141,564 | 6,314 |

[a] CWHS data, 1991-2009
[b] * p<0.10, ** p<0.05, *** p<0.01

The parameter estimates, including the household fixed effects, from Equation A.2.1 are shown in Table 6. These estimates are then used to impute values for log wages in years of each households' economic life for which we do not have data.

This creates a balanced panel of log wages of households with heads aged 21 to 80. The actual and imputed wage values are then used to calculate the net present value of lifetime labor endowments per adult for each household. Specifically, we define lifetime income for household $i$ as:

$$LI_i = \sum_{t=21}^{80} \left(\frac{1}{1+r}\right)^{t-21} (w_{i,t} * 4000) \tag{A.2.2}$$

Note that households are all have the same time endowment in each year (4000 hours). Thus the amount of the time endowment scales lifetime income up or down, but does not change the lifetime income of one household relative to another. This is not the case with the interest rate, $r$, which we fix at 4%. Changes in the interest rate differentially impact the lifetime income calculation for different individuals because they may face different earnings profiles. For example, a higher interest rate would reduced the discounted present value of lifetime income for those individuals whose wage profiles peaked later in their economic life by a larger amount than it would reduce the discounted present value of lifetime income for individuals whose wage profiles peaked earlier.

### A-2.3 Profiles by Lifetime Income

With observations of lifetime income for each household, we next sort households and find the percentile of the lifetime income distribution that each household falls in. With these percentiles, we create our lifetime income groupings.

$$\lambda_j = [0.25, 0.25, 0.2, 0.1, 0.1, 0.09, 0.01]$$

That is, lifetime income group one includes those in below the 25th percentile, group two includes those from the 25th to the median, group three includes those from the median to the 70th percentile, group four includes those from the 70th to the 80th percentile, group 5 includes those from the 80th to 90th percentile, group 6 includes those from the 90th to 99th percentile, and group 7 consists of the top one percent in the lifetime income distribution. Table 7 presents descriptive statistics for each of these groups.

To get a life-cycle profile of effective labor units for each group, we estimate the wage profile for each lifetime income group. We do this by estimating the following regression model separately for each lifetime income group using data on actual (not imputed) wages:

$$ln(w_{i,t}) = \alpha + \beta_1 age_{i,t} + \beta_2 age^2_{i,t} + \beta_3 * age^3_{i,t} + \varepsilon_{i,t} \tag{A.2.3}$$

The estimated parameters from A.2.3 are given in Table 8. The life-cycle earnings profiles implied by these parameters are plotted in Figure 19 (duplicate of Figure 1). Note that there are few individuals above age 80 in the data. To extrapolate these estimates for model ages 80-100, we use an arctan function of the following form:

## Table 7: Descriptive Statistics by Lifetime Income Category

| Lifetime Income Category: | 1 | 2 | 3 | 4 | 5 | 6 | 7 | All |
|---|---|---|---|---|---|---|---|---|
| Percentiles | 0-25 | 25-50 | 50-70 | 70-80 | 80-90 | 90-99 | 99-100 | 0-100 |
| Observations | 65,698 | 101,484 | 74,253 | 33,528 | 31,919 | 24,370 | 2,129 | 333,381 |
| Fraction Single | | | | | | | | |
| Females | 0.30 | 0.24 | 0.25 | 0.32 | 0.38 | 0.40 | 0.22 | 0.28 |
| Males | 0.18 | 0.22 | 0.30 | 0.35 | 0.38 | 0.37 | 0.20 | 0.26 |
| Fraction Married | | | | | | | | |
| Female Head | 0.08 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.02 |
| Male Head | 0.45 | 0.53 | 0.45 | 0.32 | 0.23 | 0.23 | 0.57 | 0.39 |
| Mean: | | | | | | | | |
| Age, Primary | 51.72 | 44.15 | 38.05 | 34.09 | 31.53 | 30.79 | 40.17 | 39.10 |
| Hourly Wage | 11.60 | 16.98 | 20.46 | 23.04 | 26.06 | 40.60 | 237.80 | 21.33 |
| Annual Wages | 25,178 | 44,237 | 54,836 | 57,739 | 61,288 | 92,191 | 529,522 | 51,604 |
| Lifetime Income | 666,559 | 1,290,522 | 1,913,029 | 2,535,533 | 3,249,287 | 5,051,753 | 18,080,868 | 2,021,298 |

* CWHS data, 1991-2009, all nominal values in 2005$.

$$y = \left(\frac{-a}{\pi}\right) * arctan(bx + c) + \frac{a}{2} \tag{A.2.4}$$

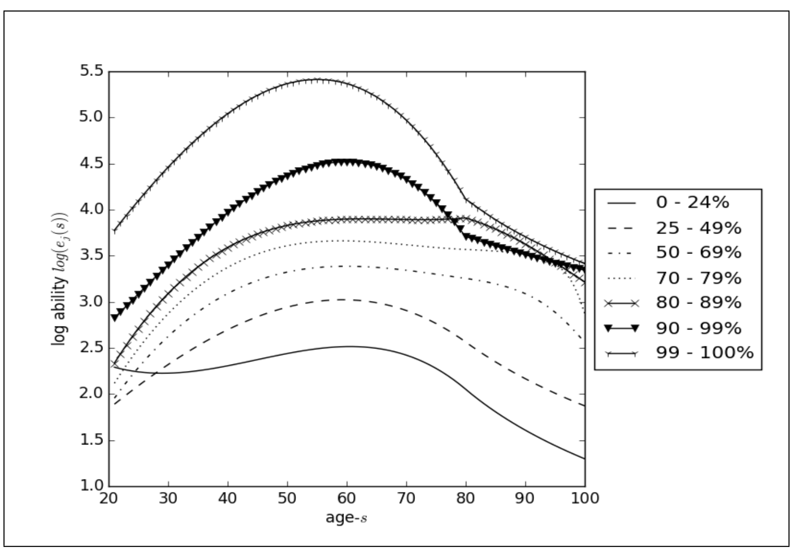
where $x$ is age, and $a$, $b$, and $c$ are the parameters we search over for the best fit of the function to the following three criteria: 1) the value of the function should match the value of the data at age 80 2) the slope of the arctan should match the slope of the data at age 80 and 3) the value of the function should match the value of the data at age 100 times a constant. This constant is .5 for all lifetime income groups, except the 2nd highest ability is .7 (otherwise, the 2nd highest has a lower income than the 3rd highest ability group in the last few years).

**Table 8:** Log Wage Regressions, by Lifetime Income Group

| Lifetime Income Group | Constant | *Age* | *Age*$^2$ | *Age*$^3$ | Observations |
|---|---|---|---|---|---|
| 1 | 3.41000000*** | -0.09720122*** | 0.00247639*** | -0.00001842*** | 65,698 |
| | (0.08718100) | 0.00543339) | 0.00010901) | 0.00000071) | |
| 2 | 0.69689692*** | 0.05995294*** | -0.00004086 | -0.00000521*** | 101,484 |
| | (0.05020758) | (0.00345549) | (0.00007627) | (0.00000054) | |
| 3 | -0.78761958*** | 0.17654618*** | -0.00240656*** | 0.00001039*** | 74,253 |
| | (0.04519637) | (0.00338371) | (0.00008026) | (0.00000061) | |
| 4 | -1.11000000*** | 0.21168263*** | -0.00306555*** | 0.00001438*** | 33,528 |
| | (0.06838352) | (0.00530190) | (0.00012927) | (0.00000099) | |
| 5 | -0.93939272*** | 0.21638731*** | -0.00321041*** | 0.00001579*** | 31,919 |
| | (0.08333727) | (0.00664647) | (0.00016608) | (0.00000130) | |
| 6 | 1.60000000*** | 0.04500235*** | 0.00094253*** | -0.00001470*** | 24,370 |
| | (0.11723131) | (0.00931334) | (0.00022879) | (0.00000176) | |
| 7 | 1.89000000*** | 0.09229392** | 0.00012902 | -0.00001169* | 2,129 |
| | (0.50501510) | (0.03858202) | (0.00090072) | (0.00000657) | |

[a] CWHS data, 1991-2009
[b] * p<0.10, ** p<0.05, *** p<0.01

**Figure 19: Exogenous life cycle income ability paths $\log(e_{j,s})$ with $S = 80$ and $J = 7$**



## A-3 Derivation of elliptical disutility of labor supply

Evans and Phillips (2018) provide an exposition of the value of using elliptical disutility of labor specification as well as its relative properties to such standard disutility of labor functions such as constant relative risk aversion (CRRA) and constant Frisch elasticity (CFE). A standard specification of additively separable period utility in consumption and labor supply first used in King et al. (1988) is the following,

$$u(c,n) = \frac{c^{1-\sigma}-1}{1-\sigma} + \chi^n \frac{\left(\tilde{l}-n\right)^{1+\theta}}{1+\theta} \tag{A.3.1}$$

where $\sigma \geq 1$ is the coefficient of relative risk aversion on consumption, $\theta \geq 0$ is proportional to the inverse of the Frisch elasticity of labor supply, and $\tilde{l}$ is the time endowment or the maximum labor supply possible. The constant $\chi^n$ is a scale parameter influencing the relative disutility of labor to the utility of consumption.

Although labor supply is only defined for $n \in [0, \tilde{l}]$, the marginal utility of leisure at $n = \tilde{l}$ is infinity and is not defined for $n > \tilde{l}$. However, utility of labor in this functional form is defined for $n < 0$. To avoid the well known and significant computational difficulty of computing the solution to the complementary slackness conditions in the Karush, Kuhn, Tucker constrained optimization problem, we impose an approximating utility function that has properties bounding the solution for $n$ away from both $n = \tilde{l}$ and $n = 0$. The upper right quadrant of an ellipse has exactly this property and also has many of the properties of the original utility function. Figure 20 shows how our estimated elliptical utility function compares to the utility of labor from (A.3.1) over the allowed support of $n$.

The general equation for an ellipse in $x$ and $y$ space with centroid at coordinates $(h, k)$, horizontal radius of $a$, vertical radius of $b$, and curvature $v$ is the following.

$$\left(\frac{x-h}{a}\right)^{v} + \left(\frac{y-k}{b}\right)^{v} = 1 \tag{A.3.2}$$

Figure 21 shows an ellipse with the parameterization $[h, k, a, b, v] = [1, -1, 1, 2, 2]$.

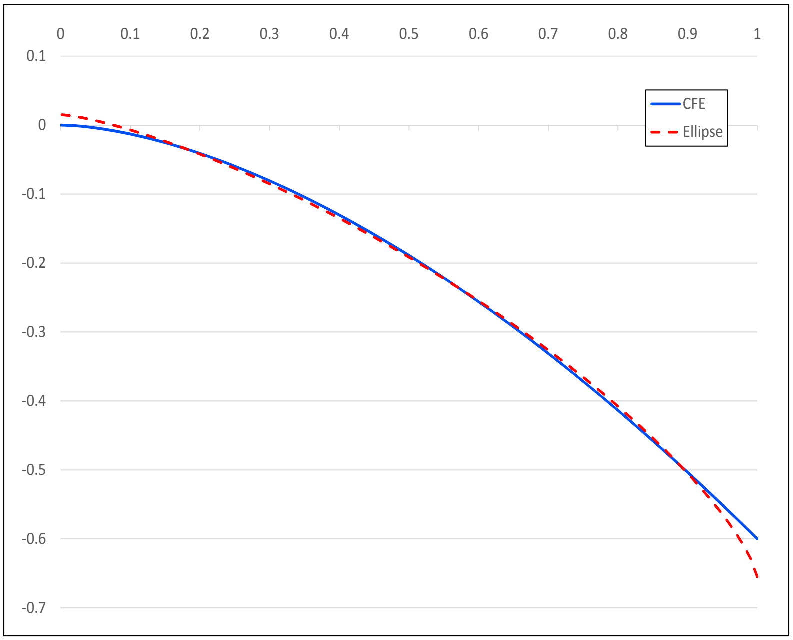
The graph of the ellipse in the upper-right quadrant of Figure 21 ($x \in [1, 2]$ and $y \in [-1, 1]$) has similar properties to the utility of labor term in (A.3.1). If we let the $x$ variable be labor supply $n$, the utility of labor supply be $g(n)$, the $x$-coordinate of the centroid be zero $h = 0$, and the horizontal radius of the ellipse be $a = \tilde{l}$, then the equation for the ellipse corresponding to the standard utility specification is the following.

$$\left(\frac{n}{\tilde{l}}\right)^{v} + \left(\frac{g-k}{b}\right)^{v} = 1 \tag{A.3.3}$$

Solving the equation for $g$ as a function of $n$, we get the following.

$$g(n) = b\left[1 - \left(\frac{n}{\tilde{l}}\right)^{v}\right]^{\frac{1}{v}} + k \tag{A.3.4}$$

**Figure 20: Comparison of standard utility of labor $n$ to elliptical utility**



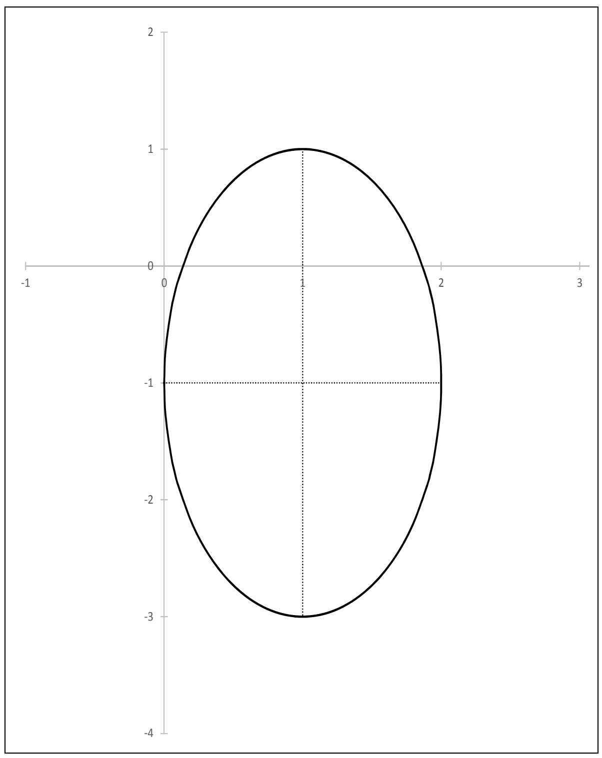

The $\upsilon$ parameter acts like a constant elasticity of substitution, and the parameter $b$ is a shape parameter similar to $\chi^n$ in (A.3.1).

We use the upper-right quadrant of the elliptical utility function because the utility of $n$ is strictly decreasing on $n \in (0, \tilde{l})$, because the slope of the utility function goes to negative infinity as $n$ approaches its maximum of $\tilde{l}$ and because the slope of the utility function goes to zero as $n$ approaches its minimum of 0. This creates interior solutions for all optimal labor supply choices $n^* \in (0, \tilde{l})$. Although it is more realistic to allow optimal labor supply to sometimes be zero, the complexity and dimensionality of our model requires this approximating assumption to render the solution method tractable.

Figure 20 shows how closely the estimated elliptical utility function matches the original utility of labor function in (A.3.1) with a Frish elasticity of 1.5[27] . We choose the ellipse parameters $b$, $k$, and $\upsilon$ to best match the points on the original utility of labor function for $n \in [0,1]$. We minimize the sum of absolute errors for 101 evenly spaced points on this domain. The estimated values of the parameters for the elliptical utility shown in Figure 20 and represented in equation (A.3.4) are $[b, k, \upsilon] = [.6701, -.6548, 1.3499]$.

[27]See Chetty et al. (2011), Keane and Rogerson (2012) and Peterman (2016) for discussion of this choice.

**Figure 21: Ellipse with** $[h, k, a, b, v] = [1, -1, 1, 2, 2]$



## A-4 Calibrating the effective income tax rates

This section shows the calibration of the functional form form the effective income tax rate on total labor and capital income.

We use the Individual Statistical Tables from the IRS (2014, Table 1.2) Statistics of Income (SOI) program, which describes tax payments and reported income for all tax filers in 2012. The tables give values aggregated into 19 different categories by total income and by adjusted gross income (AGI). Table 9 shows the data for broad income categories. This measure of the effective tax rate incorporates the different filing types, different exemptions, and different deductions. The dashed line in Figure 22 shows the effective tax rate by total income.

**Table 9: Effective tax rates by income category for all U.S. tax filers in 2012**

| AGI category | Percent of total returns | Average AGI[a] | Average total income[b] | Effective tax rate (data)[c] | Effective tax rate (model) |
|---|---|---|---|---|---|
| All returns, total | 100.00% | $62,791 | $84,377 | 9.7% | |
| No AGI | 1.47% | -$90,922 | -$85,305 | -0.1% | |
| $\$1 \leq Y < \$5K$ | 7.16% | $2,616 | $11,252 | 0.0% | 0.0% |
| $\$5K \leq Y < \$10K$ | 8.25% | $7,603 | $19,606 | 0.2% | 0.1% |
| $\$10K \leq Y < \$15K$ | 8.72% | $12,505 | $26,843 | 0.5% | 1.3% |
| $\$15K \leq Y < \$20K$ | 8.01% | $17,434 | $33,063 | 1.0% | 1.6% |
| $\$20K \leq Y < \$25K$ | 7.02% | $22,416 | $38,735 | 1.6% | 2.0% |
| $\$25K \leq Y < \$30K$ | 6.03% | $27,437 | $44,341 | 2.2% | 2.3% |
| $\$30K \leq Y < \$40K$ | 9.97% | $34,783 | $52,510 | 3.2% | 2.9% |
| $\$40K \leq Y < \$50K$ | 7.50% | $44,765 | $63,745 | 4.4% | 3.6% |
| $\$50K \leq Y < \$75K$ | 13.10% | $61,553 | $83,573 | 5.9% | 5.0% |
| $\$75K \leq Y < \$100K$ | 8.35% | $86,452 | $112,967 | 7.2% | 7.1% |
| $\$100K \leq Y < \$200K$ | 10.80% | $134,214 | $168,257 | 10.1% | 10.5% |
| $\$200K \leq Y < \$500K$ | 2.87% | $285,681 | $340,508 | 16.4% | 16.7% |
| $\$500K \leq Y < \$1Mil$ | 0.49% | $677,280 | $776,922 | 20.8% | 20.6% |
| $\$1Mil \leq Y < \$1.5Mil$ | 0.12% | $1,208,953 | $1,365,583 | 21.7% | 21.5% |
| $\$1.5Mil \leq Y < \$2Mil$ | 0.05% | $1,720,703 | $1,942,389 | 21.7% | 21.7% |
| $\$2Mil \leq Y < \$5Mil$ | 0.07% | $2,978,821 | $3,342,054 | 21.6% | 21.9% |
| $\$5Mil \leq Y < \$10Mil$ | 0.02% | $6,839,676 | $7,634,948 | 20.9% | 21.9% |
| $Y \geq \$10Mil$ | 0.01% | $30,911,333 | $34,989,584 | 17.5% | 21.9% |

Source: IRS (2014, Table 1.2).
[a] Average AGI is total AGI divided by the number of returns.
[b] Average total income is total AGI plus total exemption amount plus total itemized deductions plus total standard deductions all divided by the total number of returns.
[c] Effective tax rate is the total income tax collected divided by total income, which is total AGI plus total exemption amount plus total itemized deductions plus total standard deductions.

**Figure 22: Effective tax rates by income (logarithmic scale): data versus model**

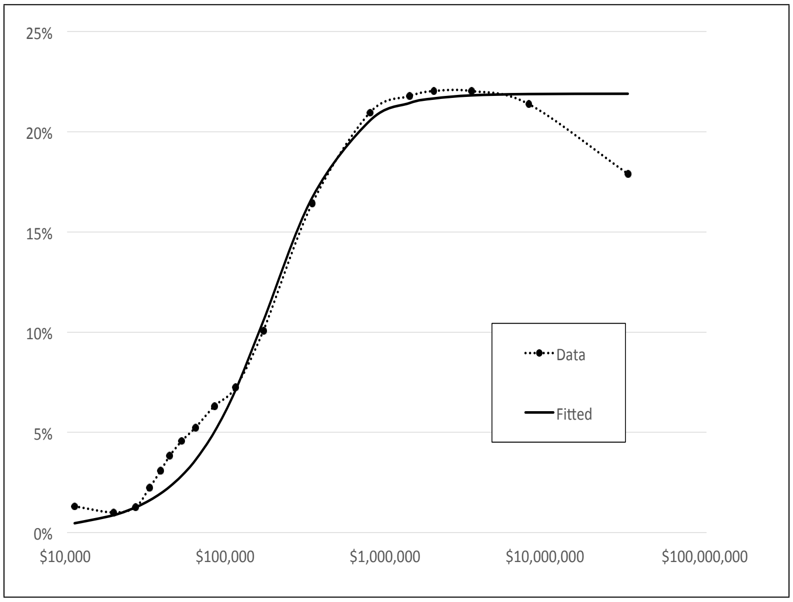
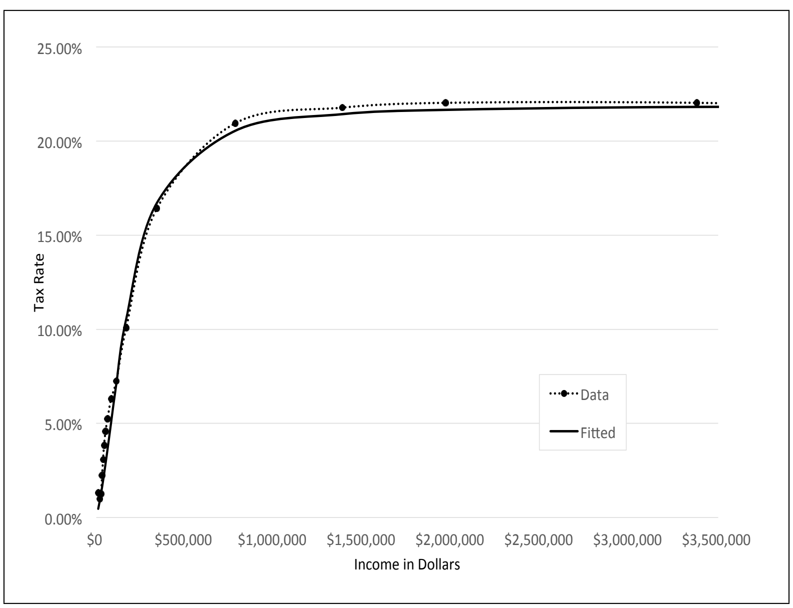



**Figure 23: Effective tax rates by income (no logarithmic scale): data versus model**



It is useful to choose a functional form for the effective tax rate function $\tau^I(Y)$ that is monotonically increasing in income $Y$, is bounded between 0 and 1, matches the general nonlinear shape of the empirical tax rate schedule, and has nice analytical derivatives. The following four-parameter function of income has all these properties.

$$\tau^I(Y) = D\left(\frac{AY^2 + BY}{AY^2 + BY + C}\right) \tag{A.4.1}$$

where $Y$ denotes income. Note that the expression in parenthesis goes to 1 as income $Y$ goes to infinity and goes to zero as $Y$ goes to zero. Then the scale parameter $D \in [0,1]$ acts as a lever to limit or expand the maximum effective tax rate that can be achieved as income $Y$ goes to infinity. The parameters $A$, $B$, and $C$ are unrestricted. The derivative of this function found in the Euler equations from Sections 2.2 and 2.5 takes on the following functional form.

$$\frac{\partial \tau^I(Y)}{\partial Y} = \frac{D(2AY + B)C}{(AY^2 + BY + C)^2} \tag{A.4.2}$$

In a model with growth, the income variable $Y$ must be stationary. We choose parameters $A$, $B$, $C$, and $D$ to minimize the sum of the squared deviations of the observed effective tax rate from the model rates generated from (A.4.1). We estimate $A = 3.0345e-06$, $B = 0.222$, $C = 133,261$, and $D = 0.219$. The solid line in Figures 22 and 23 shows the estimated effective tax rate function compared to the dashed line effective tax rate data.

# A-5 Calibrating the payroll tax replacement rates

The specification of the payroll tax and the corresponding retirement benefit are the following as described in Section 2.2.

$$T^P_{j,s,t} = \begin{cases} \tau^P w_t e_{j,s} n_{j,s,t} & \text{if} \quad s < R \\ \tau^P w_t e_{j,s} n_{j,s,t} - \theta_j w_t & \text{if} \quad s \geq R \end{cases} \tag{11}$$

Because individual lifetime income type (and thus their lifecycle earnings profile) are deterministic from birth, the Social Security replacement rate $\theta_j$ in the payroll tax (11) can be thought of as simply an percent of the age $R-1$ labor earnings. This replacement rate, $\theta_j$, is indexed to current average wage $w_t$, and then the ability $j$-specific $\theta_j$ captures the percent consistent with the average replacement amount of each type. In this way, $e_{j,s}$ is included $\theta_j$. $R$ is the age at which the individual becomes eligible to receive the retirement benefit from the payroll tax.

As mentioned in Section 2.2 and in Table 2, we calibrate the retirement age to be $R = E + s = 65$ and the payroll tax rate to $\tau^P = 0.15$. To calibrate the payroll tax replacement rates $\{\theta_j\}_{j=1}^J$, first we solve for the steady state equilibrium without the retirement benefits. Then, we calculate the monthly level of income for each ability type in dollars in our simulated model. We use the 2014 statutory formula to calculate the monthly retirement benefits or "primary insurance amount" (PIA) using the worker's earnings from the year prior to retirement in place of the average index of monthly earnings (AIME) for each ability type. By multiplying the PIA by the average effective labor participation rate and dividing by the monthly level of income, we generate the replacement rates for each ability type. We cap the replacement rates so that the maximum monthly retirement rate is thirty thousand dollars. In reality the cap is much lower than this, but in our model all wage income is subject to the payroll tax and this cap binds.

With this set of replacement rates in hand, we resolve the model including retirement benefits and repeat the calibration. We do this until the replacement rates assumed when the simulation is performed match those calculated from the statutory formula.

The statutory formula we use for PIA is as follows:

- 90% of AIME for AIME less than \$749.
- 32% of addition AIME up to \$4519.
- 15% of addition AIME up to a maximum payment of \$30,000

Our seven calibrated replacement rate values are $\theta_1 = 0.1332$, $\theta_2 = 0.1368$, $\theta_3 = 0.1368$, $\theta_4 = 0.1368$, $\theta_5 = 0.1368$, $\theta_6 = 0.1368$, and $\theta_7 = 0.1368$.

## A-6 Solving for stationary steady-state equilibrium

This section describes the solution method for the stationary steady-state equilibrium described in Definition 1.

1. Use the techniques in Appendix A-1 to solve for the steady-state population distribution vector $\bar{\boldsymbol{\omega}}$ of the exogenous population process.

2. Choose an initial guess for the stationary steady-state distribution of capital $\bar{b}_{j,s+1}$ for all $j$ and $s = E+2, E+3, ...E+S+1$ and labor supply $\bar{n}_{j,s}$ for all $j$ and $s$.

    - A good first guess is a large positive number for all the $\bar{n}_{j,s}$ that is slightly less than $\tilde{l}$ and to choose some small positive number for $\bar{b}_{j,s+1}$ that is small enough to be less than the minimum income that an individual might have $\bar{w}e_{j,s}\bar{n}_{j,s}$.

3. Perform an unconstrained root finder that chooses $\bar{n}_{j,s}$ and $\bar{b}_{j,s+1}$ that solves the $2JS$ stationary steady-state Euler equations.

4. Make sure none of the implied steady-state consumptions $\bar{c}_{j,s}$ is less-than-or-equal-to zero.

    - If one consumption is less-than-or-equal-to zero $\bar{c}_{j,s} \leq 0$, then try different starting values.

5. Make sure that none of the Euler errors is too large in absolute value for interior stationary steady-state values. A steady-state Euler error is the following, which is supposed to be close to zero for all $j$ and $s$:

$$\frac{\chi^n_s\left(\frac{b}{\tilde{l}}\right)\left(\frac{\bar{n}_{j,s}}{\tilde{l}}\right)^{\upsilon-1}\left[1-\left(\frac{\bar{n}_{j,s}}{\tilde{l}}\right)^{\upsilon}\right]^{\frac{1-\upsilon}{\upsilon}}}{(\bar{c}_{j,s})^{-\sigma}\left(\bar{w}e_{j,s}-\frac{\partial\bar{T}_{j,s}}{\partial\bar{n}_{j,s}}\right)} - 1 \quad \forall j \quad \text{and} \quad E+1\leq s\leq E+S \tag{A.6.1}$$

$$\frac{e^{-g_y\sigma}\left(\rho_s\chi^b_j\left(\bar{b}_{j,s+1}\right)^{-\sigma}+\beta(1-\rho_s)(\bar{c}_{j,s+1})^{-\sigma}\left[(1+\bar{r})-\frac{\partial\bar{T}_{j,s+1}}{\partial\bar{b}_{j,s+1}}\right]\right)}{(\bar{c}_{j,s})^{-\sigma}} - 1 \quad \forall j \quad \text{and} \quad E+1\leq s\leq E+S-1 \tag{A.6.2}$$

$$\frac{\chi^b_j e^{-g_y\sigma}(\bar{b}_{j,E+S+1})^{-\sigma}}{\left(\bar{c}_{j,E+S}\right)^{-\sigma}} - 1 \quad \forall j \tag{A.6.3}$$

## A-7 Solving for stationary non-steady-state equilibrium by time path iteration

This section defines the non-steady-state transition path equilibrium of the model and outlines the benchmark time path iteration (TPI) method of Auerbach and Kotlikoff (1987) for solving the stationary non-steady-state equilibrium transition path of the distribution of savings. The definition of the stationary non-steady-state equilibrium is similar to Definition 1, with the stationary steady-state equilibrium definition being a special case of the stationary non-steady-state equilibrium.

---

**Definition 2** (**Stationary non-steady-state equilibrium**). A non-autarkic stationary non-steady-state equilibrium in the overlapping generations model with $S$-period lived agents and heterogeneous ability $e_{j,s}$ is defined as allocations $n_{j,s,t}$ and $\hat{b}_{j,s+1,t+1}$ and prices $\hat{w}_t$ and $r_t$ for all $j$, $s$, and $t$ such that the following conditions hold:

1. individuals have symmetric beliefs $\Omega(\cdot)$ about the evolution of the distribution of savings, and those beliefs about the future distribution of savings equal the realized outcome (rational expectations),

$$\hat{\mathbf{\Gamma}}_{t+u} = \hat{\mathbf{\Gamma}}^e_{t+u} = \Omega^u\left(\hat{\mathbf{\Gamma}}_t\right) \quad \forall t, \quad u \geq 1$$

2. individuals optimize according to (28), (29), and (30)
3. Firms optimize according to (31) and (23), and
4. Markets clear according to (32) and (33).

---

Taken together, the individual labor-leisure and intended bequest decisions in the last period of life show that the optimal labor supply and optimal intended bequests for age $s = E + S$ are each functions of individual savings, total bequests received, and the prices in that period: $n_{j,E+S,t} = \phi\big(\hat{b}_{j,E+S,t}, \hat{BQ}_{j,t}, \hat{w}_t, r_t\big)$ and $\hat{b}_{j,E+S+1,t+1} = \psi\big(\hat{b}_{j,E+S,t}, \hat{BQ}_{j,t}, \hat{w}_t, r_t\big)$. These two decisions are characterized by final-age version of the static labor supply Euler equation (28) and the static intended bequests Euler equation (30). individuals in their second-to-last period of life in period $t$ have four decisions to make. They must choose how much to work this period, $n_{j,E+S-1,t}$, and next period, $n_{j,E+S,t+1}$, how much to save this period for next period, $\hat{b}_{j,E+S,t+1}$, and how much to bequeath next period, $\hat{b}_{j,E+S+1,t+2}$. The optimal responses for this individual are characterized by the $s = E+S-1$ and $s = E+S$ versions of the static Euler equations (28), the $s = E + S - 1$ version of the intertemporal Euler equation (29), and the $s = E + S$ static bequest Euler equation (30), respectively.

Optimal savings in the second-to-last period of life $s = E+S-1$ is a function of the current savings as well as the total bequests received and prices in the current period and in the next period $\hat{b}_{j,E+S,t+1} = \psi\big(\hat{b}_{j,E+S-1,t}, \hat{BQ}_{j,t}, \hat{w}_t, r_t, \hat{BQ}_{j,t+1}, \hat{w}_{t+1}, r_{t+1}|\Omega\big)$ given beliefs $\Omega$. As before, the optimal labor supply at age $s = E + S$ is a function of the next period's savings, bequests received, and prices.

$$n_{j,E+S,t+1} = \phi\big(\hat{b}_{j,E+S,t+1}, \hat{BQ}_{j,t+1}, \hat{w}_{t+1}, r_{t+1}\big)$$

But the optimal labor supply at age $s = E + S - 1$ is a function of the current savings, current bequests received, and the current prices as well as the future bequests received and future prices because of the dependence on the savings decision in that same period $n_{j,E+S-1,t} = \phi\big(\hat{b}_{j,E+S-1,t}, \hat{BQ}_{j,t}, \hat{w}_t, r_t, \hat{BQ}_{j,t+1}, \hat{w}_{t+1}, r_{t+1}|\Omega\big)$ given beliefs $\Omega$. By induction, we can show that the optimal labor supply, savings, and intended bequests functions for any individual with ability $j$, age $s$, and in period $t$ is a function of current holdings of savings and the lifetime path of total bequests received and prices given beliefs $\Omega$.

$$n_{j,s,t} = \phi\Big(\hat{b}_{j,s,t}, \big(\hat{BQ}_{j,v}, \hat{w}_v, r_v\big)_{v=t}^{t+S-s}|\Omega\Big) \quad \forall j, s, t \tag{A.7.1}$$

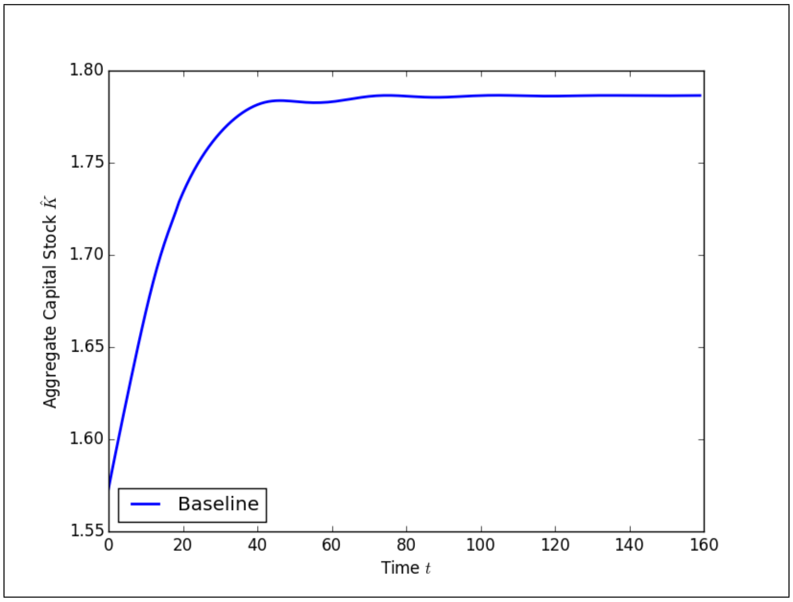
$$\hat{b}_{j,s+1,t+1} = \psi\Big(\hat{b}_{j,s,t}, \big(\hat{BQ}_{j,v}, \hat{w}_v, r_v\big)_{v=t}^{t+S-s}|\Omega\Big) \quad \forall j, t \quad \text{and} \quad E + 1 \leq s \leq E + S \tag{A.7.2}$$

If one knows the current distribution of individuals savings and intended bequests, $\hat{\mathbf{\Gamma}}_t$, and beliefs about $\hat{\mathbf{\Gamma}}_t$, then one can predict time series for total bequests received $\hat{BQ}_{j,t}$, real wages $\hat{w}_t$ and real interest rates $r_t$ necessary for solving each individual's optimal decisions. Characteristic (i) in equilibrium definition 2 implies that individuals be able to forecast prices with perfect foresight over their lifetimes implies that each individual has correct information and beliefs about all the other individuals optimization problems and information. It also implies that the equilibrium allocations and prices are really just functions of the entire distribution of savings at a particular period, as well as a law of motion for that distribution of savings.

In equilibrium, the steady-state individual labor supply, $\bar{n}_{j,s}$, for all $j$ and $s$, the steady-state savings, $\bar{b}_{j,E+S+1}$, the steady-state real wage, $\bar{w}$, and the steady-state real rental rate, $\bar{r}$, are simply functions of the steady-state distribution of savings $\bar{\Gamma}$. This is clear from the steady-state version of the capital market clearing condition (33) and the fact that aggregate labor supply is a function of the sum of exogenous efficiency units of labor in the labor market clearing condition (32). The two firm first order conditions for the real wage $\hat{w}_t$ (31) and real rental rate $r_t$ (23) are only functions of the stationary aggregate capital stock $\hat{K}_t$ and aggregate labor $\hat{L}_t$.

To solve for any stationary non-steady-state equilibrium time path of the economy from an arbitrary current state to the steady state, we follow the time path iteration (TPI) method of Auerbach and Kotlikoff (1987). The approach is to choose an arbitrary time path for the stationary aggregate capital stock $\hat{K}_t$, stationary aggregate labor $\hat{L}_t$, and total bequests received $\hat{BQ}_{j,t}$ for each type $j$. This initial guess of a path implies arbitrary beliefs that violate the rational expectations requirement. We then solve for individuals' optimal decisions given the time paths of those variables,

Figure 24: **Equilibrium time path of $K_t$ for $S = 80$ and $J = 7$ in baseline model**




which decisions imply new time paths of those variables. We then update the time path as a convex combination of the initial guess and the new implied path. Figures 24 and 25 show the equilibrium time paths of the aggregate capital stock and aggregate labor, respectively, for the calibration described in Table 2 for $T = 160$ periods starting from an initial distribution of savings in which $b_{j,s,1} = \bar{\mathbf{\Gamma}}$ for all $j$ and $s$ in the case that no policy experiment takes place. The initial capital stock $\hat{K}_1$ is not at the steady state $\bar{K}$ because the initial population distribution is not at the steady-state.

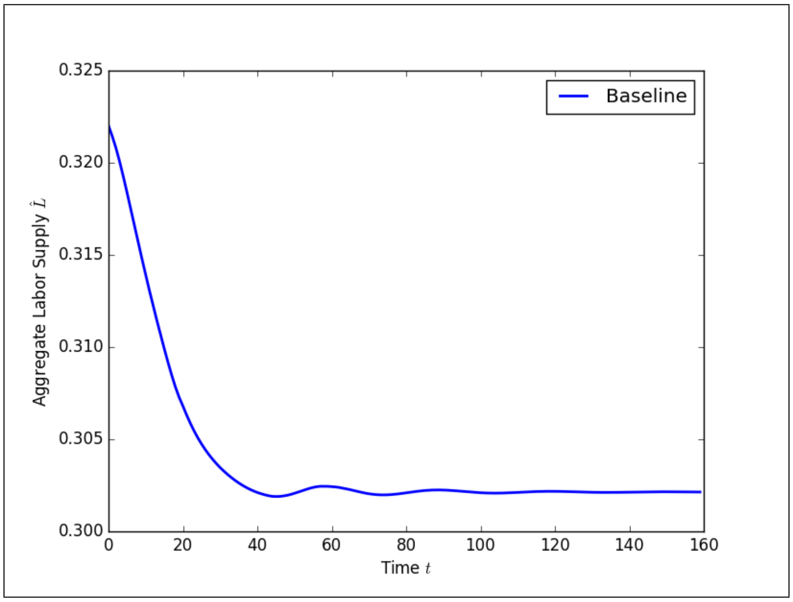
The computational approach to solving for the non-steady-state transition path equilibrium is the time path iteration (TPI) method of Auerbach and Kotlikoff (1987). TPI finds a fixed point for the transition path of the distribution of capital for a given initial state of the distribution of capital. The idea is that the economy is infinitely lived, even though the agents that make up the economy are not. Rather than recursively solving for equilibrium policy functions by iterating on individual value functions, one must recursively solve for the policy functions by iterating on the entire transition path of the endogenous objects in the economy (see Stokey and Lucas (1989, ch. 17)).

The key assumption is that the economy will reach the steady-state equilibrium described in Definition 1 in a finite number of periods $T < \infty$ regardless of the initial state. Let $\hat{\mathbf{\Gamma}}_t$ represent the distribution of stationary savings at time $t$.

$$\hat{\mathbf{\Gamma}}_t \equiv \left\{\left\{\hat{b}_{j,s,t}\right\}_{j=1}^{J}\right\}_{s=E+2}^{E+S+1}, \quad \forall t \tag{18}$$

In Section 2.5, we describe how the stationary non-steady-state equilibrium time path of allocations and price is characterized by functions of the state $\hat{\mathbf{\Gamma}}_t$ and its law of motion. TPI starts the economy at any initial distribution of savings $\hat{\mathbf{\Gamma}}_1$ and solves for its equilibrium time path over $T$ periods to the steady-state distribution $\bar{\mathbf{\Gamma}}_T$.

**Figure 25: Equilibrium time path of $L_t$ for $S = 80$ and $J = 7$ in baseline model**




The first step is to assume an initial transition path for aggregate stationary capital $\hat{\boldsymbol{K}}^i = \left\{\hat{K}_1^i, \hat{K}_2^i, ...\hat{K}_T^i\right\}$, aggregate stationary labor $\hat{\boldsymbol{L}}^i = \left\{\hat{L}_1^i, \hat{L}_2^i, ...\hat{L}_T^i\right\}$, and total bequests received $\hat{\boldsymbol{BQ}}_j^i = \{\hat{BQ}_{j,1}^i, \hat{BQ}_{j,2}^i, ...\hat{BQ}_{j,T}^i\}$ for each ability type $j$ such that $T$ is sufficiently large to ensure that $\hat{\boldsymbol{\Gamma}}_T = \bar{\boldsymbol{\Gamma}}$, $\hat{K}_T^i(\boldsymbol{\Gamma}_T)$, $\hat{L}_T^i(\boldsymbol{\Gamma}_T) = \bar{L}\left(\bar{\boldsymbol{\Gamma}}\right)$, and $\hat{BQ}_{j,T}^i(\boldsymbol{\Gamma}_T) = \bar{BQ}_j\left(\bar{\boldsymbol{\Gamma}}\right)$ for all $t \geq T$. The superscript $i$ is an index for the iteration number. The transition paths for aggregate capital and aggregate labor determine the transition paths for both the real wage $\hat{\boldsymbol{w}}^i = \{\hat{w}_1^i, \hat{w}_2^i, ...\hat{w}_T^i\}$ and the real return on investment $\boldsymbol{r}^i = \{r_1^i, r_2^i, ...r_T^i\}$. The time paths for the total bequests received also figure in each period's budget constraint and are determined by the distribution of savings and intended bequests.

The exact initial distribution of capital in the first period $\hat{\boldsymbol{\Gamma}}_1$ can be arbitrarily chosen as long as it satisfies the stationary capital market clearing condition (33).

$$\hat{K}_1 = \frac{1}{1+\tilde{g}_{n,1}} \sum_{s=E+2}^{E+S+1} \sum_{j=1}^{J} \hat{\omega}_{s-1,0}\lambda_j \hat{b}_{j,s,1} \tag{A.7.3}$$

Simiilarly, each initial value of total bequests received $\hat{BQ}_{j,1}^i$ must be consistent with the initial distribution of capital through the stationary version of (9).

$$\hat{BQ}_{j,1} = \frac{(1+r_1)\lambda_j}{1+\tilde{g}_{n,1}} \sum_{s=E+1}^{E+S} \rho_s \hat{\omega}_{s,0} \hat{b}_{j,s+1,1} \quad \forall j \tag{A.7.4}$$

However, this is not the case with $\hat{L}_1^i$. Its value will be endogenously determined in the same way the $K_2^i$ is. For this reason, a logical initial guess for the time path of aggregate labor is the steady state in every period $L_t^1 = \bar{L}$ for all $1 \leq t \leq T$.

It is easiest to first choose the initial distribution of savings $\hat{\boldsymbol{\Gamma}}_1$ and then choose an initial aggregate capital stock $\hat{K}_1^i$ and initial total bequests received $\hat{BQ}_{j,1}^i$ that correspond to that distribution. As mentioned earlier, the only other restrictions on the initial transition paths for aggregate capital, aggregate labor, and total bequests received is that they equal their steady-state levels $\hat{K}_T^i = \bar{K}\left(\bar{\boldsymbol{\Gamma}}\right)$, $\hat{L}_T^i = \bar{L}\left(\bar{\boldsymbol{\Gamma}}\right)$, and $\hat{BQ}_{j,T}^i = \bar{BQ}_j\left(\bar{\boldsymbol{\Gamma}}\right)$ by period $T$. Evans and Phillips (2014) have shown that the initial guess for the aggregate capital stocks $\hat{K}_t^i$ for periods $1 < t < T$ can take on almost any positive values satisfying the constraints above and still have the time path iteration converge.

Given the initial savings distribution $\hat{\boldsymbol{\Gamma}}_1$ and the transition paths of aggregate capital $\hat{\boldsymbol{K}}^i = \left\{\hat{K}_1^i, \hat{K}_2^i, ...\hat{K}_T^i\right\}$, aggregate labor $\hat{\boldsymbol{L}}^i = \left\{\hat{L}_1^i, \hat{L}_2^i, ...\hat{L}_T^i\right\}$, and total bequests received $\hat{\boldsymbol{BQ}}_j^i = \left\{\hat{BQ}_{j,1}^i, \hat{BQ}_{j,2}^i, ...\hat{BQ}_{j,T}^i\right\}$, as well as the resulting real wage $\hat{\boldsymbol{w}}^i = \{\hat{w}_1^i, \hat{w}_2^i, ...\hat{w}_T^i\}$, and real return to savings $\boldsymbol{r}^i = \{r_1^i, r_2^i, ...r_T^i\}$, one can solve for the period-1 optimal labor supply and intended bequests for each type $j$ of $s = E+S$-aged agents in the last period of their lives $n_{j,E+S,1} = \phi_{j,E+S}(\hat{b}_{j,E+S,1}, \hat{BQ}_{j,E+S,1}, \hat{w}_1, r_1)$ and $\hat{b}_{j,E+S+1,2} = \psi_{j,E+S}(\hat{b}_{j,E+S,1}, \hat{BQ}_{j,E+S,1}, \hat{w}_1, r_1)$ using his two $s = E+S$ static Euler equations (28) and (30).

$$
\begin{aligned}
(\hat{c}_{j,E+S,1})^{-\sigma}\left(\hat{w}_1^i e_{j,E+S} - \frac{\partial \hat{T}_{j,E+S,1}}{\partial n_{j,E+S,1}}\right) = ... \\
\chi_{E+S}^n\left(\frac{b}{\tilde{l}}\right)\left(\frac{n_{j,E+S,1}}{\tilde{l}}\right)^{\upsilon-1}\left[1 - \left(\frac{n_{j,E+S,1}}{\tilde{l}}\right)^{\upsilon}\right]^{\frac{1-\upsilon}{\upsilon}} \quad \forall j \\
\text{where} \quad \hat{c}_{j,E+S,1} = ... \\
\left(1 + r_1^i\right)\hat{b}_{j,E+S,1} + \hat{w}_1^i e_{j,E+S} n_{j,E+S,1} + \frac{\hat{BQ}_{j,1}}{\lambda_j} - e^{g_y}\hat{b}_{j,E+S+1,2} - \hat{T}_{j,E+S,1} \\
\text{and} \quad \frac{\partial \hat{T}_{j,E+S,1}}{\partial n_{j,E+S,1}} = ... \\
\hat{w}_1^i e_{j,E+S}\left[\tau^I\left(F\hat{a}_{j,E+S,1}\right) + \frac{\hat{a}_{j,E+S,1} CDF\left[2A(F\hat{a}_{j,E+S,1}) + B\right]}{\left[A(F\hat{a}_{j,E+S,1})^2 + B(F\hat{a}_{j,E+S,1}) + C\right]^2} + \tau^P\right]
\end{aligned}
\tag{A.7.5}
$$

$$
(\hat{c}_{j,E+S,1})^{-\sigma} = \chi_j^b e^{-g_y\sigma}(\hat{b}_{j,E+S+1,2})^{-\sigma} \quad \forall j \tag{A.7.6}
$$

Note that this is simply two equations (A.7.5) and (A.7.6) and two unknowns $n_{j,E+S,1}$ and $\hat{b}_{j,E+S+1,2}$.

We then solve the problem for all $j$ types of $E+S-1$-aged individuals in period $t = 1$, each of which entails labor supply decisions in the current period $n_{j,E+S-1,1}$ and in the next period $n_{j,E+S,2}$, a savings decision in the current period for the next period $\hat{b}_{j,E+S,2}$ and an intended bequest decision in the last period $\hat{b}_{j,E+S+1,3}$. The labor supply decision in the initial period and the savings period in the initial period for the next period for each type $j$ of $E+S-1$-aged individuals are policy functions of the current savings and the total bequests received and prices in this period and the next $\hat{b}_{j,E+S,2} = \psi_{j,E+S-1}(\hat{b}_{j,E+S-1,1}, \{\hat{BQ}_{j,t}, \hat{w}_t, r_t\}_{t=1}^2)$ and $\hat{n}_{j,E+S-1,1} = \phi_{j,E+S-1}(\hat{b}_{j,E+S-1,1}, \{\hat{BQ}_{j,t}, \hat{w}_t, r_t\}_{t=1}^2)$. The labor supply and intended bequests decisions in the next period are simply functions of the savings, total bequests received, and prices in that period $\hat{n}_{j,E+S,2} = \phi_{j,E+S}(\hat{b}_{j,E+S,2}, \hat{BQ}_{j,2}, \hat{w}_2, r_2)$ and $\hat{b}_{j,E+S+1,3} = \psi_{j,E+S}(\hat{b}_{j,E+S,2}, \hat{BQ}_{j,2}, \hat{w}_2, r_2)$. These four functions are characterized by the following versions of equations (28), (29), and (30).

$$
\begin{split}
(\hat{c}_{j,E+S-1,1})^{-\sigma}\left(\hat{w}_1^i e_{j,E+S-1} - \frac{\partial \hat{T}_{j,E+S-1,1}}{\partial n_{j,E+S-1,1}}\right) = ... \\
\chi^n_{E+S-1}\left(\frac{b}{\tilde{l}}\right)\left(\frac{n_{j,E+S-1,1}}{\tilde{l}}\right)^{\upsilon-1}\left[1-\left(\frac{n_{j,E+S-1,1}}{\tilde{l}}\right)^{\upsilon}\right]^{\frac{1-\upsilon}{\upsilon}} \quad \forall j
\end{split}
\tag{A.7.7}
$$

$$
\begin{split}
&(\hat{c}_{j,E+S-1,1})^{-\sigma} = ... \\
&e^{-g_y\sigma}\left(\rho_{E+S-1}\chi^b_j\left(\hat{b}_{j,E+S,2}\right)^{-\sigma} + \beta(1-\rho_{E+S-1})(\hat{c}_{j,E+S,2})^{-\sigma}\left[(1+r_2^i) - \frac{\partial T_{j,E+S,2}}{\partial b_{j,E+S,2}}\right]\right) \\
&\qquad \forall j \\
&\text{where} \quad \frac{\partial T_{j,E+S,2}}{\partial b_{j,E+S,2}} = ... \\
&\quad r_2^i\left(\tau^I(F\hat{a}_{j,E+S,2}) + \frac{F\hat{a}_{j,E+S,2}CD\left[2A(F\hat{a}_{j,E+S,2})+B\right]}{\left[A(F\hat{a}_{j,E+S,2})^2 + B(F\hat{a}_{j,E+S,2})+C\right]^2}\right)... \\
&\quad \tau^W(\hat{b}_{j,E+S,2}) + \frac{\hat{b}_{j,E+S,2}PHM}{\left(H\hat{b}_{j,E+S,2}+M\right)^2}
\end{split}
\tag{A.7.8}
$$

$$
\begin{split}
(\hat{c}_{j,E+S,2})^{-\sigma}\left(\hat{w}_2^i e_{j,E+S} - \frac{\partial \hat{T}_{j,E+S,2}}{\partial n_{j,E+S,2}}\right) = ... \\
\chi^n_{E+S}\left(\frac{b}{\tilde{l}}\right)\left(\frac{n_{j,E+S,2}}{\tilde{l}}\right)^{\upsilon-1}\left[1-\left(\frac{n_{j,E+S,2}}{\tilde{l}}\right)^{\upsilon}\right]^{\frac{1-\upsilon}{\upsilon}} \quad \forall j
\end{split}
\tag{A.7.9}
$$

$$
(\hat{c}_{j,E+S,2})^{-\sigma} = \chi^b_j e^{-g_y\sigma}(\hat{b}_{j,E+S+1,3})^{-\sigma} \quad \forall j
\tag{A.7.10}
$$

Note that this is four equations (A.7.7), (A.7.8), (A.7.9), and (A.7.10) and four unknowns $n_{j,E+S-1,1}$, $\hat{b}_{j,E+S,2}$, $n_{j,E+S,2}$, and $\hat{b}_{j,E+S+1,3}$.

This process is repeated for every age of individual alive in $t=1$ down to the age $s=E+1$ individual at time $t=1$. Each of these individuals $j$ solves the full set of remaining $S-s+1$ labor supply decisions, $S-s$ savings decisions, and one intended bequest decision at the end of life. After the full set of lifetime decisions has been solved for all the individuals alive at time $t=1$, each ability $j$ individual born in period $t\geq 2$ can be solved for, the solution to which is characterized by the following full set of Euler equations analogous to (28), (29), and (30).

$$(\hat{c}_{j,s,t})^{-\sigma}\left(\hat{w}^i_t e_{j,s} - \frac{\partial \hat{T}_{j,s,t}}{\partial n_{j,s,t}}\right) = \chi^n_s\left(\frac{b}{\tilde{l}}\right)\left(\frac{n_{j,s,t}}{\tilde{l}}\right)^{\upsilon-1}\left[1-\left(\frac{n_{j,s,t}}{\tilde{l}}\right)^{\upsilon}\right]^{\frac{1-\upsilon}{\upsilon}} \quad \forall j \quad \text{and} \quad E+1\leq s\leq E+S \quad \text{and} \quad t\geq 2 \tag{A.7.11}$$

$$\begin{aligned}(\hat{c}_{j,s,t})^{-\sigma} = \ldots \\ e^{-g_y\sigma}\left(\rho_s\chi^b_j\big(\hat{b}_{j,s+1,t+1}\big)^{-\sigma} + \beta(1-\rho_s)(\hat{c}_{j,s+1,t+1})^{-\sigma}\left[(1+r^i_{t+1}) - \frac{\partial T_{j,s+1,t+1}}{\partial b_{j,s+1,t+1}}\right]\right) \\ \forall j \quad \text{and} \quad E+1\leq s\leq E+S-1 \quad \text{and} \quad t\geq 2\end{aligned} \tag{A.7.12}$$

$$(\hat{c}_{j,E+S,t})^{-\sigma} = \chi^b_j e^{-g_y\sigma}\big(\hat{b}_{j,E+S+1,t+1}\big)^{-\sigma} \quad \forall j \quad \text{and} \quad t\geq 2 \tag{A.7.13}$$

For each individual of ability type $j$ entering the economy in period $t\geq 1$, the entire set of $2S$ lifetime decisions is characterized by the $2S$ equations represented in (A.7.11), (A.7.12), and (A.7.13).

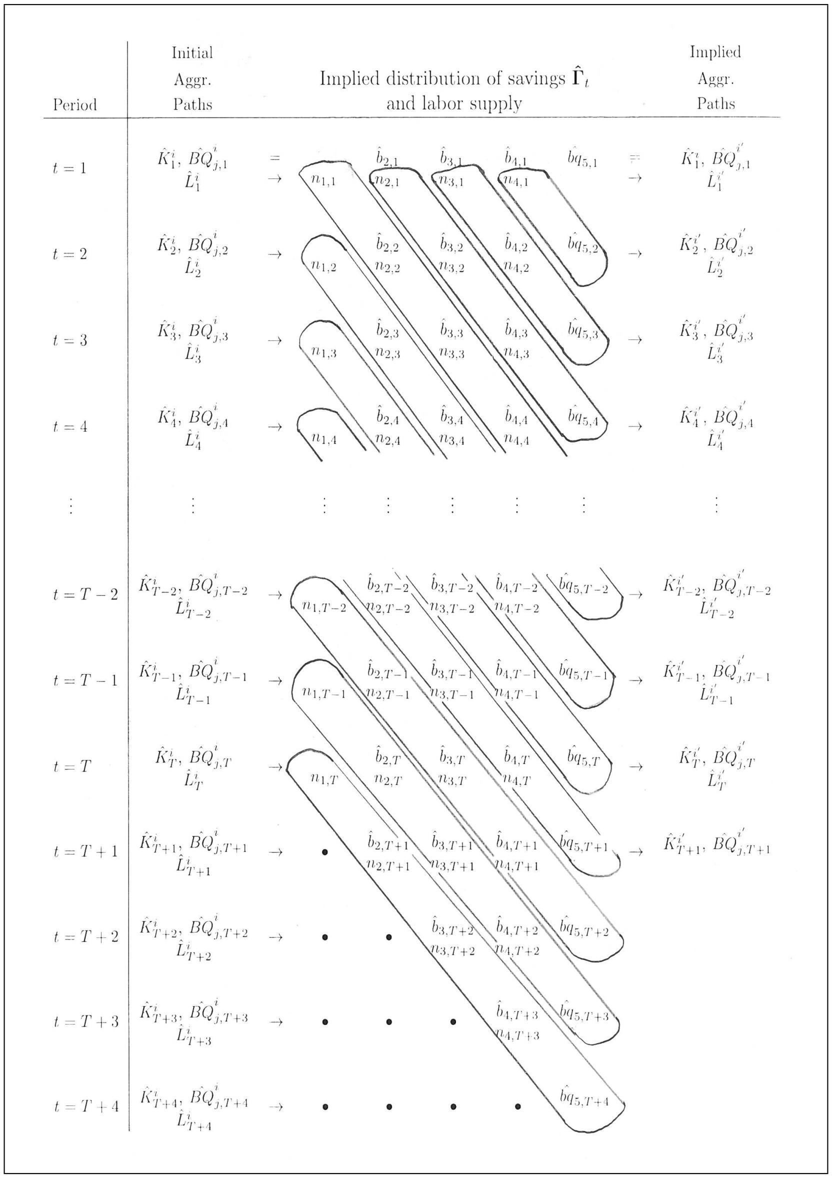
We can then solve for the entire lifetime of savings and labor supply decisions for each age $s=1$ individual in periods $t=2,3,\ldots T$. The central part of the schematic diagram in Figure 26 shows how this process is done in order to solve for the equilibrium time path of the economy from period $t=1$ to $T$. Note that for each full lifetime savings and labor supply path solved for an individual born in period $t\geq 2$, we can solve for the aggregate capital stock and total bequests received implied by those savings decisions $\hat{\boldsymbol{K}}^{i'}$ and $\hat{\boldsymbol{BQ}}^{i'}_j$ and aggregate labor implied by those labor supply decisions $\hat{\boldsymbol{L}}^{i'}$.

Once the set of lifetime saving and labor supply decisions has been computed for all individuals alive in $1\leq t\leq T$, we use the individual decisions to compute a new implied time path of the aggregate capital stock and aggregate labor. The implied paths of the aggregate capital stock $\hat{\boldsymbol{K}}^{i'} = \{\hat{K}^i_1, \hat{K}^{i'}_2, \ldots \hat{K}^{i'}_T\}$, aggregate labor $\hat{\boldsymbol{L}}^{i'} = \{\hat{L}^i_1, \hat{L}^{i'}_2, \ldots \hat{L}^{i'}_T\}$, and total bequests received $\hat{\boldsymbol{BQ}}^{i'}_j = \{\hat{BQ}^i_{j,1}, \hat{BQ}^{i'}_{j,2}, \ldots \hat{BQ}^{i'}_{j,T}\}$ in general do not equal the initial guessed paths $\hat{\boldsymbol{K}}^i = \{\hat{K}^i_1, \hat{K}^i_2, \ldots \hat{K}^i_T\}$, $\hat{\boldsymbol{L}}^i = \{\hat{L}^i_1, \hat{L}^i_2, \ldots \hat{L}^i_T\}$, and $\hat{\boldsymbol{BQ}}^i_j = \{\hat{BQ}^i_{j,1}, \hat{BQ}^i_{j,2}, \ldots \hat{BQ}^i_{j,T}\}$ used to compute the individual savings and labor supply decisions $\hat{\boldsymbol{K}}^{i'} \neq \hat{\boldsymbol{K}}^i$, $\hat{\boldsymbol{L}}^{i'} \neq \hat{\boldsymbol{L}}^i$, and $\hat{\boldsymbol{BQ}}^{i'}_j \neq \hat{\boldsymbol{BQ}}^i_j$.

Let $\|\cdot\|$ be a norm on the space of time paths of the aggregate capital stock $\hat{\boldsymbol{K}} \in \mathcal{K} \subset \mathbb{R}^T_{++}$, aggregate labor supply $\hat{\boldsymbol{L}} \in \mathcal{L} \subset \mathbb{R}^T_{++}$, and $J$ paths of total bequests received $\hat{\boldsymbol{BQ}}_j \in \mathcal{B} \subset \mathbb{R}^T_{++}$. Then the fixed point necessary for the equilibrium transition path from Definition 2 has been found when the distance between these $J+2$ paths is arbitrarily close to zero.

$$\left\|\left[\hat{\boldsymbol{K}}^{i'}, \hat{\boldsymbol{L}}^{i'}, \{\hat{\boldsymbol{BQ}}^{i'}_j\}^J_{j=1}\right] - \left[\hat{\boldsymbol{K}}^i, \hat{\boldsymbol{L}}^i, \{\hat{\boldsymbol{BQ}}^i_j\}^J_{j=1}\right]\right\| \leq \varepsilon \quad \text{for} \quad \varepsilon > 0 \tag{A.7.14}$$

**Figure 26: Diagram of TPI solution method within each iteration for $S = 4$ and $J = 1$**



If the fixed point has not been found $\left\|\left[\boldsymbol{\hat{K}}^{i'},\boldsymbol{\hat{L}}^{i'},\{\boldsymbol{\hat{BQ}}_j^{i'}\}_{j=1}^J\right]-\left[\boldsymbol{\hat{K}}^{i},\boldsymbol{\hat{L}}^{i},\{\boldsymbol{\hat{BQ}}_j^{i}\}_{j=1}^J\right]\right\|>\varepsilon$, then new transition paths for the aggregate capital stock and aggregate labor are generated as a convex combination of $\left[\boldsymbol{\hat{K}}^{i'},\boldsymbol{\hat{L}}^{i'},\{\boldsymbol{\hat{BQ}}_j^{i'}\}_{j=1}^J\right]$ and $\left[\boldsymbol{\hat{K}}^{i},\boldsymbol{\hat{L}}^{i},\{\boldsymbol{\hat{BQ}}_j^{i}\}_{j=1}^J\right]$.

$$\begin{aligned}
\boldsymbol{\hat{K}}^{i+1} &= \nu\boldsymbol{\hat{K}}^{i'} + (1-\nu)\boldsymbol{\hat{K}}^{i} \\
\boldsymbol{\hat{L}}^{i+1} &= \nu\boldsymbol{\hat{L}}^{i'} + (1-\nu)\boldsymbol{\hat{L}}^{i} \\
\boldsymbol{\hat{BQ}}_1^{i+1} &= \nu\boldsymbol{\hat{BQ}}_1^{i'} + (1-\nu)\boldsymbol{\hat{BQ}}_1^{i} \qquad \text{for} \quad \nu\in(0,1] \\
&\ \ \vdots \\
\boldsymbol{\hat{BQ}}_J^{i+1} &= \nu\boldsymbol{\hat{BQ}}_J^{i'} + (1-\nu)\boldsymbol{\hat{BQ}}_J^{i}
\end{aligned} \tag{A.7.15}$$

This process is repeated until the initial transition paths for the aggregate capital stock, aggregate labor, and total bequests received are consistent with the transition paths implied by those beliefs and individual and firm optimization.

In essence, the TPI method iterates on individual beliefs about the time path of prices represented by a time paths for the aggregate capital stock $\boldsymbol{\hat{K}}^i$, aggregate labor $\boldsymbol{\hat{L}}^i$, and total bequests received $\boldsymbol{\hat{BQ}}_j^i$ until a fixed point in beliefs is found that are consistent with the transition paths implied by optimization based on those beliefs.

The following are the steps for computing a stationary non-steady-state equilibrium time path for the economy.

1. Input all initial parameters. See Table 2.

    (a) The value for $T$ at which the non-steady-state transition path should have converged to the steady state should be at least as large as the number of periods it takes the population to reach its steady state $\boldsymbol{\bar{\omega}}$ as described in Appendix A-1.

2. Choose an initial distribution of savings and intended bequests $\boldsymbol{\hat{\Gamma}}_1$ and then calculat the initial state of the stationarized aggregate capital stock $\hat{K}_1$ and total bequests received $\hat{BQ}_{j,1}$ consistent with $\boldsymbol{\hat{\Gamma}}_1$ according to (33) and (A.7.4).

    (a) Note that you must have the population weights from the previous period $\hat{\omega}_{s,0}$ and the growth rate between period 0 and period 1 $\tilde{g}_{n,1}$to calculate $\hat{BQ}_{j,1}$.

3. Conjecture transition paths for the stationarized aggregate capital stock $\boldsymbol{\hat{K}}^1=\{\hat{K}_t^1\}_{t=1}^\infty$, stationarized aggregate labor $\boldsymbol{\hat{L}}^1=\{\hat{L}_t^1\}_{t=1}^\infty$, and total bequests received $\boldsymbol{\hat{BQ}}_j^1=\{\hat{BQ}_{j,t}^1\}_{t=1}^\infty$ where the only requirements are that $\hat{K}_1^i$ and $\hat{BQ}_{j,1}^i$ are functions of the initial distribution of savings $\boldsymbol{\hat{\Gamma}}_1$ for all $i$ is your initial state and that $\hat{K}_t^i=\bar{K}$, $\hat{L}_t^i=\bar{L}$, and $\hat{BQ}_{j,t}^i=\bar{BQ}_j$ for all $t\geq T$. The conjectured transition paths of the aggregate capital stock $\boldsymbol{\hat{K}}^i$ and aggregate labor $\boldsymbol{\hat{L}}^i$ imply specific transition paths for the real wage $\boldsymbol{\hat{w}}^i=\{\hat{w}_t^i\}_{t=1}^\infty$ and the real interest rate $\boldsymbol{r}^i=\{r_t^i\}_{t=1}^\infty$ through expressions (31) and (23).

(a) An intuitive choice for the time path of aggregate labor is the steady-state in every period $\hat{L}^1_t = \bar{L}$ for all $t$.

4. With the conjectured transition paths $\hat{\boldsymbol{w}}^i$, $\boldsymbol{r}^i$, and $\hat{\boldsymbol{BQ}}^i_j$ one can solve for the lifetime policy functions of each individual alive at time $1 \leq t \leq T$ using the systems of Euler equations of the form (28), (29), and (30) and following the diagram in Figure 26.

5. Use the implied distribution of savings and labor supply in each period (each row of $\hat{b}_{j,s,t}$ and $n_{j,s,t}$ in Figure 26) to compute the new implied time paths for the aggregate capital stock $\hat{\boldsymbol{K}}^{i'} = \{\hat{K}^i_1, \hat{K}^{i'}_2, ...\hat{K}^{i'}_T\}$, aggregate labor supply $\hat{\boldsymbol{L}}^{i'} = \{\hat{L}^i_1, \hat{L}^{i'}_2, ...\hat{L}^{i'}_T\}$, and total bequests received $\hat{\boldsymbol{BQ}}^{i'}_j = \{\hat{BQ}^i_{j,1}, \hat{BQ}^{i'}_{j,2}, ...\hat{BQ}^{i'}_{j,T}\}$.

6. Check the distance between the two sets time paths.

$$\left\| \left[\hat{\boldsymbol{K}}^{i'}, \hat{\boldsymbol{L}}^{i'}, \{\hat{\boldsymbol{BQ}}^{i'}_j\}_{j=1}^J\right] - \left[\hat{\boldsymbol{K}}^{i}, \hat{\boldsymbol{L}}^{i}, \{\hat{\boldsymbol{BQ}}^{i}_j\}_{j=1}^J\right] \right\|$$

(a) If the distance between the initial time paths and the implied time paths is less-than-or-equal-to some convergence criterion $\varepsilon > 0$, then the fixed point has been achieved and the equilibrium time path has been found (A.7.14).

(b) If the distance between the initial time paths and the implied time paths is greater than some convergence criterion $\|\cdot\| > \varepsilon$, then update the guess for the time paths according to (A.7.15) and repeat steps (4) through (6) until a fixed point is reached.

# A-8 Sensitivity Analysis of Results

In this section, we perform sensitivity analysis of our main results from Table 3. In particular, we study the results with respect to three different values of the coefficient of relative risk aversion $\sigma = \{1.1, 2.1, 3.2\}$. The baseline value of the coefficient of relative risk aversion in Table 3 is $\sigma = 3.0$. In Table 10 we only look at Gini coefficients for the total model population in the steady state rather than the additional categories from Table 3 in which we average over age and over ability.

**Table 10: Comparison of changes in steady-state Gini coefficients from wealth tax versus income tax for different values of $\sigma$**

| Steady-state variable | Risk aversion value | Baseline | Wealth tax Treatment | Wealth tax % Chg. | Income tax Treatment | Income tax % Chg. |
|---|---|---|---|---|---|---|
| $\bar{b}_{j,s}$ Wealth | $\sigma = 1.1$ | 0.940 | 0.927 | -1.38% | 0.934 | -0.64% |
| | $\sigma = 2.1$ | 0.940 | 0.927 | -1.38% | 0.933 | -0.74% |
| | $\sigma = 3.0$ | 0.943 | 0.929 | -1.48% | 0.939 | -0.42% |
| | $\sigma = 3.2$ | 0.937 | 0.924 | -1.39% | 0.933 | -0.43% |
| $\bar{y}_{j,s}$ Income | $\sigma = 1.1$ | 0.780 | 0.755 | -3.21% | 0.773 | -0.90% |
| | $\sigma = 2.1$ | 0.786 | 0.748 | -4.83% | 0.760 | -3.31% |
| | $\sigma = 3.0$ | 0.775 | 0.733 | -5.42% | 0.757 | -2.32% |
| | $\sigma = 3.2$ | 0.766 | 0.730 | -4.70% | 0.750 | -2.09% |
| $\bar{c}_{j,s}$ Consumption | $\sigma = 1.1$ | 0.602 | 0.567 | -5.81% | 0.562 | -6.64% |
| | $\sigma = 2.1$ | 0.673 | 0.636 | -5.50% | 0.646 | -4.01% |
| | $\sigma = 3.0$ | 0.664 | 0.621 | -6.48% | 0.644 | -3.01% |
| | $\sigma = 3.2$ | 0.671 | 0.634 | -5.51% | 0.654 | -2.53% |
| $\bar{n}_{j,s}$ Labor supply | $\sigma = 1.1$ | 0.430 | 0.447 | 3.95% | 0.426 | -0.93% |
| | $\sigma = 2.1$ | 0.270 | 0.293 | 8.52% | 0.268 | -0.74% |
| | $\sigma = 3.0$ | 0.240 | 0.258 | 7.50% | 0.236 | -1.67% |
| | $\sigma = 3.2$ | 0.218 | 0.236 | 8.26% | 0.219 | -0.46% |

Note: All Gini coefficients are over all steady-state values in the distribution, which corresponds to the "Total" category from Table 3.

The results for each value of the coefficient of relative risk aversion $\sigma$ in Table 10 are qualitatively and quantitatively similar to the results from Table 3. For almost all values of $\sigma$, the wealth tax reduces inequality more than the more progressive income tax that raises the same steady-state revenue. The only exception is for the Gini coefficient on steady-state consumption for $\sigma = 1.1$. In this case, the reduction in inequality is roughly similar, but the reduction from the income tax is slightly greater. All other cases remain as in Table 3.