

# The Redistributive Benefits of Progressive Labor and Capital Income Taxation, or: How to Best Screw the Top 1%

Fabian Kindermann  
University of Würzburg and Netspar

Dirk Krueger  
University of Pennsylvania, CEPR, NBER and Netspar

January 29, 2014

## Abstract

In this paper we argue that a capital income tax is an effective tool for redistribution and insurance even when progressive labor income taxes are available to the policy maker. To make this point we construct a large scale Overlapping Generations Model with uninsurable income risk, show that it has a wealth distribution that matches the data well, and then use it to characterize fiscal policies that achieve a desired degree of redistribution in society. We find that it is suboptimal to rely exclusively on progressive labor income taxes to achieve any given level of redistribution in society, and thus that a positive capital income tax should be part of a government policy aimed at redistributing welfare across ex ante homogeneous, but ex post heterogeneous households. We finally characterize the optimal Rawlsian policy and find that it includes a significantly positive capital income tax, in addition to a redistributive labor income tax.

## 1 Introduction

Large increase in earnings, income and wealth inequality. Cite facts on shares of top 1% by Piketty and Saez. Call for higher marginal tax rate at the top of the distribution (see e.g. Diamond and Saez, XXX). Suppose we take as given that society wants to redistribute, how should it best do it? Capital income taxation? Progressive labor income taxation? Combination of both.

Standard wisdom: redistribute via labor income tax (or if can via initial endowment tax), but do not cause intertemporal distortions. Chamley (1986) and especially Judd (1985) who shows that even if you only care about workers should not tax capital in the long run.

Question in this paper. Can one make a strong case for capital income taxation even if redistributive labor income taxation (and lump-sum taxation) is fully permitted?

## 2 The Model

We study a standard large-scale overlapping generations model in the spirit of Auerbach and Kotlikoff (1987), but augmented by exogenous ex-ante heterogeneity across households by education levels as well as ex-post heterogeneity due to uninsurable idiosyncratic labor productivity and thus wage risk, as in Conesa, Kitao and Krueger (2009). Given the focus of the paper it is especially important that the endogenous earnings and wealth distributions predicted by the model well approximate their empirical counterparts, both at the low and the high end of the distribution.

In order to highlight the key ingredients of the model in its most transparent way for a given government policy we first set out the model using recursive language and define a stationary recursive competitive equilibrium. We then turn to a description of the potential policy reforms and the transition dynamics induced by it.

### 2.1 Technology

The single good in this economy is produced by a continuum of representative, competitive firms that hire capital and labor on competitive spot markets to operate the constant returns to scale technology

$$Y = \Omega K^\epsilon L^{1-\epsilon}, \quad (1)$$

where  $\Omega \geq 0$  parametrizes the level of technology and the parameter  $\epsilon \in [0, 1]$  measures the elasticity of output with respect to capital. Capital depreciates at rate  $\delta_k$  in every period. Given our assumptions of perfect competition in all markets and constant returns to scale production technologies the number of operative firms as well as their size is indeterminate and without loss of generality we can assume the existence of a representative, competitively behaving firm producing according to the aggregate production function (1).

### 2.2 Preferences and Endowments

Households in this economy are finitely lived, with maximal life span given by  $J$  and generic age denoted by  $j$ . In each period a new generation age cohort is born whose size is  $1 + n$  as large as the previous cohort, so that  $n$  is the constant and exogenous population growth rate. We denote by  $\psi_{j+1}$  the conditional probability of survival of each household from age  $j$  to age  $j + 1$ . At age  $j_r < J$  households become unproductive and thus retire after age  $j_r$ .

Households have preferences defined over stochastic streams of consumption and labor  $\{c_j, n_j\}$  determined by the period utility function

$$U(c_j, n_j),$$

and the time discount factor  $\beta$  and are expected utility maximizers (with respect to longevity risk and with respect to idiosyncratic wage risk described below).

Households are ex-ante heterogeneous with respect to the education they have acquired, a process we do not model endogenously. Let  $s \in \{n, c\}$  denote the education level of the household, with  $s = c$  denoting some college education and  $s = n$  representing (less than or equal) high school education. The fraction of college educated households is exogenously given by  $\phi_s$ . In addition, prior to labor market entry households draw a fixed effect<sup>1</sup>  $\alpha$  from an education-specific distribution  $\phi_s(\alpha)$ .

The wage a household faces in the labor market is given by

$$w \cdot e(j, s, \alpha, \eta)$$

where  $w$  is the aggregate wage per labor efficiency unit and  $e(j, s, \alpha, \eta)$  captures idiosyncratic wage variation that is a function of the age, education status and fixed effect of the household as well as a random component  $\eta$  that follows an education specific first order Markov chain with states  $\eta \in \mathcal{E}_s$  and transition matrix  $\pi_s(\eta'|\eta)$ .

Idiosyncratic wage risk (determined by the process for  $\eta$ ) and mortality risk (parameterized by the survival probabilities  $\psi_j$ ) cannot be explicitly insured as markets are incomplete as in Bewley (1986), Huggett (1993) or Aiyagari (1994); however, households can self-insure against these risks by saving at a risk-free after-tax interest rate  $r_n = r(1 - \tau_k)$ . In addition to saving  $a' - a$  the household spends her income, composed of earnings  $we(j, s, \alpha, \eta)n$ , capital income  $r_n a$  and transfers  $b_j(s, \alpha, \eta)$  (including social security for those that are retired as well as accidental bequests for all working households) on consumption  $(1 + \tau_c)c$ , including consumption taxes, and on paying labor income taxes  $T(we(j, s, \alpha, \eta)n)$  as well as payroll taxes  $T_{ss}(we(j, s, \alpha, \eta)n)$ . Implicit in these formulations is that the consumption and capital income tax is assumed to be linear, whereas the labor earnings tax is given by the potentially nonlinear (but continuously differentiable) function  $T(\cdot)$ .

The individual state variables of the household thus include  $(j, s, \alpha, \eta, a)$ , the exogenous age, education and idiosyncratic wage shock, as well as the endogenously chosen asset position. For given (time-invariant) prices, taxes and transfers, the dynamic programming problem of the household then reads as

$$v(j, s, \alpha, \eta, a) = \max_{c, n, a'} U(c, n) + \beta \psi_{j+1} \sum_{\eta'} \pi_s(\eta'|\eta) v(j+1, s, \alpha, \eta', a') \quad (2)$$

---

<sup>1</sup>Both education and the fixed effect will shift life cycle wage profiles in a deterministic fashion in the model, so we could have combined them into a single fixed effect. However, when mapping the model to wage data it is more transparent to distinguish between the two components impacting the deterministic part of wages.

subject to

$$\begin{aligned} & (1 + \tau_c)c + a' + T(we(j, s, \alpha, \eta)n) + T_{ss}(we(j, s, \alpha, \eta)n) \\ = & (1 + r_n)a + b_j(s, \eta) + we(j, s, \alpha, \eta)n \end{aligned} \quad (3)$$

and subject to a tight borrowing limit  $\alpha' \geq 0$ . The result of this dynamic programming problem is a value function  $v$  and policy functions  $c, n, a'$  as functions of the state  $(j, s, \alpha, \eta, a)$  of a household.

### 2.3 Government Policy

The government uses tax revenues from labor earnings, capital income and consumption taxes to finance an exogenously given stream of government expenditures  $G$  and the interest payments on government debt  $B$ . In addition it runs a balanced-budget pay-as you go social security (and medicare program). Finally it collects accidental bequests and redistributes them among the surviving population in a lump-sum fashion. Since the population is growing at a constant rate  $n$  in this economy, it is easier to interpret  $(G, B)$  as per capita variables since these are constant in a stationary recursive competitive equilibrium.

Letting by  $\Phi$  denote the cross-sectional distribution<sup>2</sup> of households (constant in a stationary equilibrium), the budget constraint of the government in a stationary recursive competitive equilibrium with population growth reads as

$$\begin{aligned} & r\tau_k \int a'(j, s, \alpha, \eta, a)d\Phi + \tau_c \int c(j, s, \alpha, \eta, a)d\Phi + \int T(we(j, s, \alpha, \eta)n(j, s, \alpha, \eta, a))d\Phi \\ = & G + (r - n)B \end{aligned} \quad (4)$$

In addition, the PAYGO social security system is characterized by a payroll tax rate  $\tau_{ss}$ , an earnings threshold  $\bar{y}_{ss}$  only below which households pay social security taxes, and benefits  $p(s, \alpha, \eta)$  that depend on the last realization of the persistent wage shock  $\eta$  of working age<sup>3</sup> as well as education  $s$  and the fixed effect  $\alpha$  (which in turn determine expected wages over the life cycle). Thus  $(\tau_{ss}, \bar{y}_{ss})$  completely determine the payroll tax function  $T_{ss}$ . The specific form of the function  $p_j(s, \alpha, \eta)$  is discussed in the calibration section.

The budget constraint of the social security system then reads as

$$\int p(s, \alpha, \eta)d\Phi = \tau_{ss} \int \min\{\bar{y}_{ss}, we(j, s, \alpha, \eta)n(j, s, \alpha, \eta, a)\}d\Phi. \quad (5)$$

---

<sup>2</sup>Formally, and given our notation,  $\Phi$  is a *probability* measure so that the total mass of households is equal to 1.

<sup>3</sup>This formulation has the advantage that we can capture the feature of the actual system that social security benefits are increasing in earnings during working age, without adding an additional continuous state variable (such as average earnings during the working age). Since benefits depend on the exogenous  $\eta$  rather than endogenous labor earnings, under our specification households do not have an incentive to increase labor supply in their last working period to boost pension payments.

Our formulation does require that we keep track of the last realization of  $\eta$  a household had in her working life, though, and thus  $\eta$  remains a state in the retirement period, with public pension payments being denoted as  $p_j = p_j(\eta)$ .

Finally, we assume that accidental bequests are lump-sum redistributed among the surviving working age population, and thus

$$Tr = \frac{\int (1 + r^n)(1 - \psi_{j+1})a'(j, s, \alpha, \eta, a)d\Phi}{\int \mathbf{1}_{\{j \leq j_r\}}d\Phi}. \quad (6)$$

so that transfers received by households are given as

$$b_j(s, \alpha, \eta) = \begin{cases} Tr & \text{if } j \leq j_r \\ p(s, \alpha, \eta) & \text{if } j > j_r \end{cases} \quad (7)$$

## 2.4 Recursive Competitive Equilibrium (RCE)

**Definition 1** *Given government expenditures  $G$ , government debt  $B$ , a tax system characterized by  $(\tau_c, \tau_k, T)$  and a social security system characterized by  $(\tau_{ss}, \bar{y}_{ss})$ , a stationary recursive competitive equilibrium with population growth is a collection of value and policy functions  $(v, c, n, a')$  for the household, optimal input choices  $(K, L)$  of firms, transfers  $Tr$ , prices  $(r, w)$  and an invariant probability measure  $\Phi$*

1. *[Household maximization]: Given prices  $(r, w)$ , transfers  $b_j$  given by (7) and government policies  $(\tau_c, \tau_k, T, \tau_{ss}, \bar{y}_{ss})$ , the value function  $v$  satisfies the Bellman equation (2), and  $(c, n, a')$  are the associated policy functions.*
2. *[Firm maximization]: Given prices  $(r, w)$ , the optimal choices of the representative firm satisfy*

$$\begin{aligned} r &= \Omega\epsilon \cdot \left[ \frac{L}{K} \right]^{1-\epsilon} - \delta_k \\ w &= \Omega(1-\epsilon) \left[ \frac{K}{L} \right]^\epsilon. \end{aligned}$$

3. *[Government Budget Constraints]: Government policies satisfy the government budget constraints (4) and (5).*
4. *[Market clearing]:*

(a) *The labor market clears:*

$$L = \int e(j, s, \alpha, \eta)n(j, s, \alpha, \eta, a)d\Phi$$

(b) *The capital market clears*

$$(1 + n)(K + B) = \int a'(j, s, \alpha, \eta, a)d\Phi$$

(c) *The goods market clears*

$$Y = \int c(j, s, \alpha, \eta, a) d\Phi + \delta K + G$$

5. *[Consistency of Probability Measure  $\Phi$ ]: The invariant probability measure is consistent with the population structure of the economy, with the exogenous processes  $\pi_s$ , and the household policy function  $a'(\cdot)$ . A formal definition is provided in Appendix B.*

## 2.5 Transition Paths

Our thought experiments will involve unexpected changes in government tax policy that will induce the economy to undergo a deterministic transition path from the initial benchmark stationary recursive competitive equilibrium to a final RCE associated with the new long-run policy. At any point of time the aggregate economy is characterized by a cross-sectional probability measure  $\Phi_t$  over household types. The household value functions, policy functions, prices, policies and transfers are now also indexed by time, and the key equilibrium conditions, the government budget constraint and the capital market clearing conditions now read as

$$G + (1 + r_t)B_t = (1 + n)B_{t+1} + r_t \tau_{k,t}(K_t + B_t) + \tau_c \int c_t(j, s, \alpha, \eta, a) d\Phi_t + \int T_t(w_t e(j, s, \alpha, \eta) n_t(j, s, \alpha, \eta, a)) d\Phi_t$$

and

$$(1 + n)(K_{t+1} + B_{t+1}) = \int a'_t(j, s, \alpha, \eta, a) d\Phi_t$$

Note that, in line with the policy experiments conducted below, the capital income tax rate  $\tau_{k,t}$ , the labor earnings tax function  $T_t$  and government debt are now permitted to be functions of time  $t$ . For a complete formal definition of a dynamic equilibrium with time varying policies in an economy very close to ours, see e.g. Conesa, Kitao and Krueger (2009).

## 3 Mapping the Model into Data

Conceptually, we proceed in two steps when we map the initial stationary equilibrium of our model into U.S. data. We first choose a subset of the parameters based on model-exogenous information. Then we calibrate the remaining parameters such that the initial stationary equilibrium is consistent with selected aggregate and distributional statistics of the U.S. economy. Even though it is understood that all model parameters impact all equilibrium entities, the discussion below associates those parameters to specific empirical targets that, in the model, impact the corresponding model statistics most significantly.

Most of the calibration is fairly standard for quantitative OLG models with idiosyncratic risk. However, given the purpose of the paper it is important that the model-generated cross-sectional earnings and wealth distribution is

characterized by the same concentration as in the data, especially at the top of the earnings and wealth distribution. Broadly, we follow Castaneda et al. (2003) and augment fairly standard stochastic wage processes derived from the PSID with labor productivity states that occur with low probability, but induce persistently large earnings when they occur. This allows the model to match the high earnings concentration and the even higher wealth concentration at the top of the distribution. On the other hand, the explicit life cycle structure, including a fully articulated social security system, permits us to generate a distribution of earnings and wealth at the bottom and the middle that matches the data quite well.

### 3.1 Demographics

We set the population growth rate at  $n = 1.1\%$ , the long run average value for the U.S. Data on survival probabilities from the Human Mortality Database for the US in 2010 is used to determine the age-dependent survival probabilities  $\{\psi_j\}$ .

### 3.2 Technology

The production side of the model is characterized by the three parameters  $(\Omega, \epsilon, \delta_k)$ . We set the capital share in production to  $\epsilon = 0.33$  and normalize the level of technology  $\Omega$  such that the equilibrium wage rate per efficiency unit of labor is  $w = 1$ . The depreciation rate on capital  $\delta_k$  is set such that the initial equilibrium interest rate in the economy is  $r = 4\%$ ; this requires an annual depreciation rate of  $\delta_k = 7.6\%$ .

### 3.3 Endowments and Preferences

#### 3.3.1 Labor Productivity

In every period a household is endowed with one unit of time which can be used for leisure and market work. One unit of work time yields a wage  $we(j, s, \alpha, \eta)$ , where  $e(j, s, \alpha, \eta)$  is the idiosyncratic labor productivity (and thus the idiosyncratic component of the wage) of the household which depends on the age  $j$  education  $s$  and the fixed effect  $\alpha$  of the household as well as its idiosyncratic shock  $\eta$ .

We assume that  $\eta \in \mathcal{E}_s$  can take on 7 (education-specific) values; we associate an  $\eta \in \{\eta_{s,1}, \dots, \eta_{s,5}\}$  with “normal” labor earnings observed in household data sets such as the PSID, and reserve  $\{\eta_{s,6}, \eta_{s,7}\}$  for the very high labor productivity and thus earnings realizations observed at the top of the cross-sectional distribution, but not captured by any observations in the PSID. We then specify log-wages as

$$\ln e(j, s, \alpha, \eta) = \begin{cases} \alpha + \varepsilon_{j,s} + \eta & \text{if } \eta \in \{\eta_{s,1}, \dots, \eta_{s,5}\} \\ \eta & \text{if } \eta \in \{\eta_{s,6}, \eta_{s,7}\} \end{cases}$$

|         | $\rho_s$ | $\sigma_{\eta,s}^2$ | $\sigma_{\alpha,s}^2$ | $\phi_s$ |
|---------|----------|---------------------|-----------------------|----------|
| $s = n$ | 0.9850   | 0.0346              | 0.2061                | 0.59     |
| $s = c$ | 0.9850   | 0.0180              | 0.1517                | 0.41     |

Table 1: Labor Productivity Process

That is, as long as the labor productivity shock  $\eta \in \{\eta_{s,1}, \dots, \eta_{s,5}\}$ , idiosyncratic wages are (in logs) the sum of the fixed effect  $\alpha$  that is constant over the life cycle, an education-specific age-wage profile  $\varepsilon_{j,s}$  and the random component  $\eta$ , as is fairly standard in quantitative life cycle models with idiosyncratic risk (see e.g. Conesa et al., 2009). On the other hand, if a household becomes highly productive,  $\eta \in \{\eta_{s,6}, \eta_{s,7}\}$ , wages are independent of education and the fixed effect. We think of these states as representing, in a reduced form, successful entrepreneurial or artistic opportunities that yield very high earnings and that are independent of the education level and fixed effect of the household.<sup>4</sup>

Given these assumptions we need to specify the seven states of Markov chain  $\{\eta_1, \dots, \eta_7\}$  as well as the transition matrices  $\pi_s$ ; in addition we need to determine the education-specific distribution of the fixed effect  $\phi_s(\alpha)$  and the deterministic, education-specific age-wage profile  $\{\varepsilon_{j,s}\}$ . For the latter we use the direct estimates from the PSID by Krueger and Ludwig (2013). Furthermore we assume that for each education group  $s \in \{n, c\}$  the fixed effect  $\alpha$  can take two values  $\alpha \in \{-\sigma_{\alpha,s}, \sigma_{\alpha,s}\}$  with equal probability,  $\phi_s(-\sigma_{\alpha,s}) = \phi_s(\sigma_{\alpha,s}) = 0.5$ . For the "normal" labor productivity states  $\{\eta_{s,1}, \dots, \eta_{s,5}\}$  we use a discretized (by the Rouwenhorst method) Markov chain of a continuous, education-specific AR(1) process with persistence  $\rho_s$  and (conditional) variance  $\sigma_{\eta,s}^2$ . Thus the parameters governing this part of the labor productivity process are the education-specific variances of the fixed effect, the AR(1) process as well as their persistence,  $\{\sigma_{\alpha,s}^2, \sigma_{\eta,s}^2, \rho_s\}$ , together with the share of households  $\phi_s$  with a college education. Table 1 summarizes our choices.

[Need to add a convincing justification]

In order to account for very high earnings realizations we add to the Markov process described above two more states  $\{\eta_{s,6}, \eta_{s,7}\}$ . Recall that for these states wages are simply determined as  $w \exp(\eta)$  and thus do not depend on the fixed effect  $\alpha$  and the deterministic age profile; this formulation leads to a much better fit of the age-earnings and age-asset distributions. We then augment the  $5 \times 5$

---

<sup>4</sup>Conceptually, nothing prevents us to specify

$$e(j, s, \alpha, \eta) = \exp(\alpha + \varepsilon_{j,s} + \eta)$$

for  $\eta \in \{\eta_6, \eta_7\}$ , but it turns out that our chosen specification provides a better fit to the earnings and wealth distributions.

| Parameters                            |                           | Targets                           |
|---------------------------------------|---------------------------|-----------------------------------|
| Prob. to high wage region ( $s = n$ ) | $\pi_{.6,n}$              | 95-99% Earnings                   |
| Prob. to high wage region ( $s = c$ ) | $\pi_{.6,c}$              | 99-100% Earnings                  |
| Persistence high shock ( $s = n$ )    | $\pi_{66,n}$              | Share college in 95-99% Earnings  |
| Persistence high shock ( $s = c$ )    | $\pi_{66,c}$              | Share college in 99-100% Earnings |
| Prob. to highest wage ( $s = n$ )     | $\pi_{67,n}$              | Gini Earnings                     |
| Prob. to highest wage ( $s = n$ )     | $\pi_{67,c}$              | 95-99% Wealth                     |
| Persistence highest shock             | $\pi_{77,n} = \pi_{77,c}$ | 99-100% Wealth                    |
| High wage shock ( $s = n$ )           | $\eta_{n,6}$              | Share college in 95-99% Wealth    |
| High wage shock ( $s = c$ )           | $\eta_{c,6}$              | Share college in 99-100% Wealth   |
| Highest income shock                  | $\eta_{n,7} = \eta_{c,7}$ | Gini Wealth                       |

Table 2: Earnings and Wealth Targets

Markov transition matrices  $\pi_s = (\pi_{ij,s})$  as follows:

$$\pi_s = \begin{bmatrix} \pi_{11,s}(1 - \pi_{16,s}) & \dots & \pi_{13,s}(1 - \pi_{16,s}) & \dots & \pi_{15,s}(1 - \pi_{16,s}) & \pi_{16,s} & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & 0 \\ \pi_{51,s}(1 - \pi_{56,s}) & \dots & \pi_{53,s}(1 - \pi_{16,s}) & \dots & \pi_{55,s}(1 - \pi_{56,s}) & \pi_{56,s} & 0 \\ 0 & \dots & 1 - \pi_{66,s} - \pi_{67,s} & \dots & 0 & \pi_{66,s} & \pi_{67,s} \\ 0 & \dots & 0 & \dots & 0 & 1 - \pi_{77,s} & \pi_{77,s} \end{bmatrix}$$

and assume that  $\pi_{16,s} = \dots = \pi_{56,s} = \pi_{.6,s}$ . Thus from each "normal" state  $\{\eta_{s,1}, \dots, \eta_{s,5}\}$  there is a (small) probability to climb to the high state  $\eta_{s,6}$ . The highest state  $\eta_{s,7}$  can only be reached from state  $\eta_{s,6}$ , and households at the highest state can only fall to state  $\eta_{s,6}$ . If wage productivity falls back to the "normal" range, it falls to  $\eta_{s,3}$  with probability 1. The transition matrix above reflects these assumptions which will permit us to match both the empirical earnings and wealth distribution (including at the top) very accurately. In addition, we assume that  $\eta_{n,7} = \eta_{c,7}$  and  $\pi_{77,n} = \pi_{77,c}$ . This leaves us with ten additional parameters characterizing the labor productivity process which we summarize, including the empirical targets, in table 2. Appendix 10 gives the exact values of the transition probabilities and states of the Markov chains.

### 3.3.2 Preferences

We assume that the period utility function is given by

$$U(c, n) = \frac{c^{1-\gamma}}{1-\gamma} - \lambda \frac{n^{1+\chi}}{1+\chi}$$

We exogenously set risk aversion to  $\gamma = 2$  and  $\chi = 1.67$  in order to obtain a Frisch elasticity of labor supply of  $1/\chi = 0.6$ . The disutility of labor parameter

$\lambda$  is chosen so that households spend, on average, one third of their time endowment on market work. Finally, the time discount factor  $\beta$  is chosen such that the capital-output ratio in the economy is equal to 3.

### 3.4 Government Policies

The two government policies we model explicitly are the tax system and the social security system.<sup>5</sup> The main focus of the paper is on the composition of the labor earnings and the capital income tax schedule, as well as the progressivity of the former, especially at the high end of the earnings distribution.

#### 3.4.1 The Tax System

We assume that the labor earnings tax function is characterized by the marginal tax rate function  $T'(y)$  depicted in figure 1. It is thus characterized by two tax rates  $\tau_l, \tau_h$  and two earnings thresholds  $\bar{y}_l, \bar{y}_h$ . Earnings below  $\bar{y}_l$  are not taxed, earnings above  $\bar{y}_h$  are taxed at the highest marginal rate  $\tau_h$ , and for earnings in the interval  $[\bar{y}_l, \bar{y}_h]$  marginal taxes increase linearly from  $\tau_l$  to  $\tau_h$ . This tax code strikes a balance between approximating the current income tax code in the U.S., being parameterized by few parameters and being continuously differentiable above the initial earnings threshold  $\bar{y}_l$ , which is crucial for our computational algorithm. Varying  $\tau_h$  permits us to control the extent to which labor earnings at the top of the earnings distribution is taxed, and changing  $\bar{y}_h$  controls at what income threshold the highest marginal tax rate sets in. Furthermore, if an increase in  $\tau_h$  is met by a reduction of the lowest positive marginal tax rate  $\tau_l$  (say, to restore government budget balance), the resulting new tax system is more progressive than the original one.

For the initial equilibrium we choose the highest marginal tax rate  $\tau_h = 39.6\%$ , equal to the current highest marginal income tax rate of the federal income tax code.<sup>6</sup> That tax rate applies to labor earnings in excess of 4 times average household *income*, or  $\bar{y}_2 = 4\bar{y}$ . Households below 35% of median income do not pay any taxes,  $\bar{y}_1 = 0.35y^{med}$  and we determine  $\tau_l$  from budget balance in the initial stationary equilibrium, given the other government policies discussed below.<sup>7</sup> This requires  $\tau_l = 12.2\%$ , roughly the midpoint of the two lowest marginal tax rates of the current U.S. federal income tax code (10% and 15%).

<sup>5</sup>In addition the government collects and redistributes accidental bequests. This activity does not require the specification of additional parameters, however.

<sup>6</sup>This value for the highest marginal tax rate is also close to the value assumed by Diamond and Saez (2011) once taxes for Medicare are abstracted from (we interpret Medicare as part of the social security system).

<sup>7</sup>In the data the income thresholds at which the lowest and highest marginal tax rates apply depend on the family structure and filing status of the household. Krueger and Ludwig argue that the value of the tax exemption and standard deduction constitute roughly 35% of median household income, for various household types.

To interpret the upper income threshold  $\bar{y}_h$ , note that in the model  $x\%$  of households in the initial equilibrium have earnings that exceed this threshold.

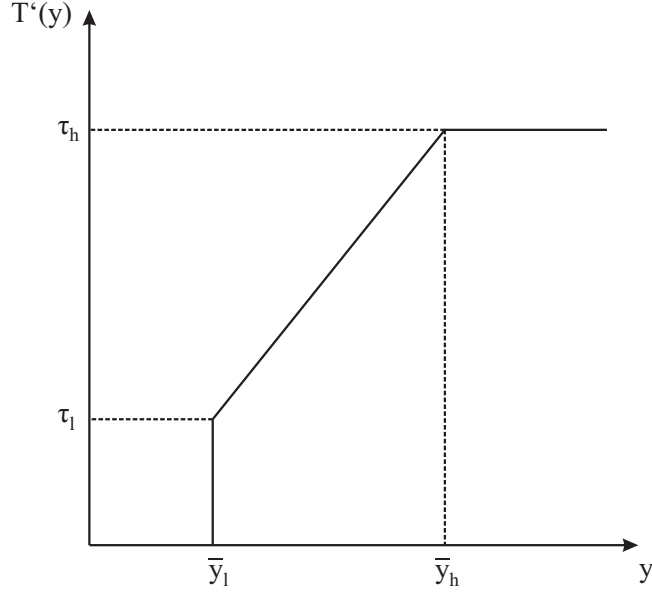


Figure 1: Marginal Labor Income Tax Function

The initial proportional capital income tax rate is set to  $\tau_k = 28.3\%$  and the consumption tax rate to  $\tau_c = 5\%$ . We choose exogenous government spending  $G$  such that it constitutes 17% of GDP; outstanding government debt  $B$  is set such that the debt-to-GDP ratio is 60% in the initial stationary equilibrium. These choices coincide with those in Krueger and Ludwig (2013) who argue that these values reflect well U.S. policy prior to the great recession.

### 3.4.2 The Social Security System

We model the social security system as a flat labor earnings tax  $\tau_{ss}$  up to an earnings threshold  $\bar{y}_{ss}$ , together with a benefit formula that ties benefits to past earnings, but without introducing an additional continuous state variable (such as average indexed monthly earnings). Thus we compute, for every state  $(s, \alpha, \eta)$ , average labor earnings in the population for that state,  $\bar{y}(s, \alpha, \eta)$ , and apply the actual progressive social security benefit formula  $f(y)$  to  $\bar{y}(s, \alpha, \eta)$ . The social security benefit a household of type  $(s, \alpha)$  with shock  $\eta_{65}$  in the last period of her working life receives is then given by

$$p(s, \alpha, \eta) = f(\bar{y}(s, \alpha, \eta = \eta_{65})).$$

We discuss the details of the benefit formula in appendix 10.

| Parameter                               | Value         | Target/Data              |
|-----------------------------------------|---------------|--------------------------|
| Survival probabilities $\{\psi_i\}$     |               | HMD 2010                 |
| Population growth rate $n$              | 1.1%          |                          |
| Capital share in production $\epsilon$  | 33%           |                          |
| Threshold positive taxation $\bar{y}_l$ | 35%           | as fraction of $y^{med}$ |
| Top tax bracket $\bar{y}_h$             | 400%          | as fraction of $y^{med}$ |
| Top marginal tax rate $\tau_h$          | 39.6%         |                          |
| Consumption tax rate $\tau_c$           | 5%            |                          |
| Capital income tax $\tau_k$             | 28.3%         |                          |
| Government debt to GDP $B/Y$            | 60%           |                          |
| Government consumption to GDP $G/Y$     | 17%           |                          |
| Bend points $b_1, b_2$                  | 0.184, 1.114  | SS data                  |
| Replacement rates $r_1, r_2, r_3$       | 90%, 32%, 15% | SS data                  |
| Pension cap $\bar{y}_{ss}$              | 200%          | $\tau_p = 0.124$         |
| Risk aversion $\gamma$                  | 2             |                          |
| Inverse of Frisch elasticity $\chi$     | 1.67          |                          |

Table 3: Exogenously Chosen Parameters

### 3.5 Calibration Summary

The following tables 3 and 4 summarize the choice of the remaining exogenously set parameters as well as those endogenously calibrated within the model. The exogenously chosen parameters include policy parameters describing current U.S. fiscal policy, as well as the capital share in production  $\epsilon$  and the preference parameters  $(\gamma, \chi)$ . The choices for these parameters are standard relative to the literature, with the possible exception of the Frisch labor supply elasticity  $1/\chi = 0.6$ , which is larger than the microeconomic estimates for white prime age males. However, it should be kept in mind that we are modeling household labor supply, including the labor supply of the secondary earner. Note that this choice implies, ceteris paribus, strong disincentive effects on labor supply from higher marginal tax rates at the top of the earnings distribution.

The set of parameters calibrated within the model include the technology parameters  $(\delta_k, \Omega)$ , the preference parameters  $(\beta, \lambda)$  as well as the entry marginal tax rate  $\tau_l$ . The latter is chosen to assure government budget balance in the initial stationary equilibrium. The preference parameters are chosen so that the model equilibrium is consistent with a capital-output ration of 3 and a share of time spent on market work equal to 33% of the total time endowment available to households. The technology parameters are then determined to reproduce a real (pre-tax) return on capital of 4% and a wage rate of 1, the latter being an innocuous normalization of  $\Omega$ . Table 4 summarizes the associated values of the parameters.

| Parameter                          | Value | Target/Data      |
|------------------------------------|-------|------------------|
| Technology level $\Omega$          | 0.922 | $w = 1$          |
| Depreciation rate $\delta_k$       | 7.6%  | $r = 4\%$        |
| Initial marginal tax rate $\tau_l$ | 12.2% | Budget balance   |
| Time discount factor $\beta$       | 0.977 | $K/Y = 3.0$      |
| Disutility from labor $\lambda$    | 36    | $\bar{n} = 33\%$ |

Table 4: Endogenously Calibrated Parameters

| <i>Macroeconomic variables</i> |       |
|--------------------------------|-------|
| Variable                       | Value |
| Capital                        | 288%  |
| Government debt                | 60%   |
| Consumption                    | 58%   |
| Investment                     | 25%   |
| Government Consumption         | 17%   |
| Av. hours worked (in %)        | 33%   |
| Interest rate (in %)           | 4%    |
| Tax revenues                   |       |
| - Consumption                  | 2.9%  |
| - Labor                        | 11.9% |
| - Capital income               | 3.9%  |
| Pension System                 |       |
| Contribution rate (in %)       | 12.4% |
| Total pension payments         | 5.2%  |

Table 5: Endogenously Calibrated Parameters

## 4 Characteristics of the Benchmark Economy

Prior to turning to our tax experiments we first briefly discuss the aggregate and distributional properties of the initial stationary equilibrium. This is perhaps more important than for most applications since a realistic earnings and wealth distribution, especially at the top of the distribution

### 4.1 Macroeconomic Aggregates

In table 5 we summarize the key macroeconomic aggregates implied by the initial stationary equilibrium of our model. It shows that the main source of government tax revenues are taxes on labor earnings.

All variable in % of GDP if not indicated otherwise

| <i>The labor earnings distribution</i> |           |     |      |      |      |         |       |        |       |
|----------------------------------------|-----------|-----|------|------|------|---------|-------|--------|-------|
|                                        | Quintiles |     |      |      |      | Top (%) |       |        | Gini  |
|                                        | 1st       | 2nd | 3rd  | 4th  | 5th  | 90-95   | 95-99 | 99-100 |       |
| Share of total sample (in %)           |           |     |      |      |      |         |       |        |       |
| Model                                  | 0.0       | 5.8 | 11.0 | 17.6 | 65.6 | 11.7    | 18.9  | 21.4   | 0.642 |
| US Data                                | -0.1      | 4.2 | 11.7 | 20.8 | 63.5 | 11.7    | 16.6  | 18.7   | 0.636 |

Table 6: Endogenously Calibrated Parameters

| <i>The wealth distribution</i> |           |     |     |      |      |         |       |        |       |
|--------------------------------|-----------|-----|-----|------|------|---------|-------|--------|-------|
|                                | Quintiles |     |     |      |      | Top (%) |       |        | Gini  |
|                                | 1st       | 2nd | 3rd | 4th  | 5th  | 90-95   | 95-99 | 99-100 |       |
| Share of total sample (in %)   |           |     |     |      |      |         |       |        |       |
| Model                          | 0.0       | 0.8 | 4.1 | 11.6 | 83.6 | 14.6    | 23.3  | 31.8   | 0.810 |
| US Data                        | -0.2      | 1.1 | 4.5 | 11.2 | 83.4 | 11.1    | 26.7  | 33.6   | 0.816 |

Table 7: Endogenously Calibrated Parameters

## 4.2 Earnings and Wealth Distribution

In this section we show that, given our earnings process with small but positive probability of very high earnings realizations, the model is able to reproduce an empirically realistic cross-sectional earnings and wealth distribution.

Table 6 displays the model-implied earnings distribution and table 7 does the same for the wealth distribution. When comparing the model-implied earnings and wealth quintiles to the corresponding statistics from the data we observe that the model fits the data very well, even at the top of the distribution. The same is true for the Gini coefficients of earnings and wealth.

An important aspect of the wealth data is that the bottom 40% of the population has essentially no net worth, something our model reproduces. Since a binding borrowing constraint significantly affects consumption, savings and labor supply choices, from the perspective of the model it is important to understand who are these households. In figure 2 we plot the share of each age cohort, in the initial stationary equilibrium, that has zero wealth. Not unexpected, binding liquidity constraints are mainly prevalent among young households and then are again observed among the very elderly who have outlived their resources and find it optimal to finance their consumption exclusively through social security benefits.

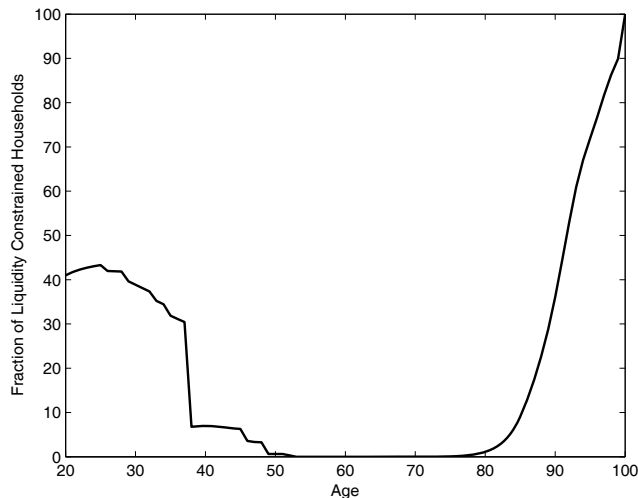


Figure 2: Fraction of Households with  $a = 0$ , by Age

Overall, we do not view the ability of the model to reproduce the earnings and wealth distributions as a success per se, since the stochastic wage process (and especially the two high-wage states) were designed for exactly that purpose. However, that fact that our approach is indeed successful gives us some confidence that ours is an appropriate model to study tax policy experiments that are highly redistributive in nature.

## 5 Quantitative Results

In this section we set out our main results. We first describe the thought experiment we consider, and then turn to the optimal tax analysis. We do so in three steps. First we investigate how large the optimal top marginal labor income tax rate should be, then we assess whether in our model with realistically high wealth inequality capital income taxes are a desired instrument for redistribution and social insurance, and finally we document whether the benevolent government will opt for highly progressive labor income taxes, capital income taxes, or a combination of both instruments when maximizing welfare of the bottom 99% of the population.

### 5.1 The Thought Experiments

We now describe our fiscal policy thought experiments. Starting from the initial steady state fiscal constitution we consider one-time, unexpected (by private households and firms) tax reforms that change either the top marginal labor

earnings tax rate or the capital income tax rate (or the combination of both). The unexpected reform induces a transition of the economy to a new stationary equilibrium, and we model this transition path explicitly. In order to insure that the intertemporal government budget constraint holds, given the initial outstanding debt and given the change in either  $\tau_h$  or  $\tau_k$ , the government in addition (and again permanently) adjusts the entry marginal tax rate  $\tau_l$ . An appropriate sequence of government debt along the transition path insures that the sequential government budget constraints hold for every period  $t$  along the transition.

In the aggregate, a transition path is thus characterized by deterministic sequences of interest rates, wages and government debt  $\{r_t, w_t, B_{t+1}\}_{t=1}^T$  converging to the new stationary equilibrium indexed by a new policy  $(\tau_l, \tau_h, \tau_k, \bar{y}_l, \bar{y}_h)$ . For every period  $t \geq 1$  along the transition path the analysis delivers new lifetime utility  $v_t(j, s, \alpha, \eta, a)$  of households with individual states  $(j, s, \alpha, \eta, a)$ . The optimal tax experiment then consists in maximizing a weighted sum of these lifetime utilities either over  $\tau_h$ , or  $\tau_k$ , or over  $(\tau_h, \tau_k)$ , using adjustments in  $\tau_l$  to insure that the intertemporal government budget constraint is satisfied.

The welfare measure we employ is constructed as follows. After solving for the equilibrium path of a specific tax reform, we calculate the amount of initial wealth transfers needed to compensate every individual back to their initial equilibrium utility level, ex post for the currently living and ex ante for future generations.<sup>8</sup> That is, for each household currently alive we find the transfer  $\Psi_0(j, s, \alpha, \eta, a)$  that satisfies

$$v_1(j, s, \alpha, \eta, a + \Psi_0(j, s, \alpha, \eta, a)) = v_0(j, s, \alpha, \eta, a)$$

where  $v_0$  denotes the value function in the initial steady state and for households born in period  $t \geq 1$  we find the number  $\Psi_t$  such that

$$Ev_t(j = 1, s, \alpha, \bar{\eta}, \Psi_t) = Ev_0(j = 1, s, \alpha, \bar{\eta}, 0)$$

where expectations are taken with respect to initial fixed effect and education. The total present discounted value of all transfers is then given by

$$W = \int \Psi_0(j, s, \alpha, \eta, a) d\Phi_0 + \mu_1 \sum_{t=1}^{\infty} \left( \frac{1+n}{1+r_0} \right)^t \Psi_t$$

When we exclude the top 1% of households from the social welfare function only transfers to the bottom 99% of the current earnings distribution are included in the calculations. In order to turn the welfare measure into easily

---

<sup>8</sup>These wealth transfers induce behavioral responses which we capture when computing the transfers necessary to make a household indifferent. We do however, abstract from the general equilibrium effects these hypothetical transfers would induce.

For future cohort the transfer is one number per cohort, for currently alive households the required transfers differ by characteristics  $(j, s, \alpha, \eta, a)$ . Future transfers are discounted to the present at rate  $\frac{1+n}{1+r_0}$  where  $r_0$  is the interest rate in the initial stationary equilibrium and our aggregate welfare measure is the sum of these transfers, with positive numbers indicating welfare gains.

interpretable numbers we turn the present value of the transfers into an annuity that pays out over the whole transition path and in the new long-run equilibrium and express the size of this annuity as a percent of initial aggregate consumption. That is, we calculate

$$C \sum_{t=0}^{\infty} \left( \frac{1+n}{1+r_0} \right)^t = W$$

and

$$CEV = 100 * \left( \frac{C}{C_0} - 1 \right)$$

This idea of calculating the welfare consequences of policy reforms follows closely that of Huang et al. (1997), and more generally, the hypothetical lump-sum redistribution authority envisioned by Auerbach and Kotlikoff () which would implement the transfer scheme described above.

The advantage of this way of measuring aggregate welfare over alternative approaches [add the usual propaganda...]

## 5.2 Optimal Size of the Top Marginal Earnings Tax Rate

In this section we document the optimal top marginal labor earnings tax rate, which is indeed very high, in excess of 90%. In figure 3 we plot three welfare measures against the top marginal tax rate  $\tau_h$ . The black line plots the aggregate welfare measure, derived by compensating all current and future generations so that their (remaining ) lifetime utility is the same as it would have been under the status quo fiscal policy. The blue line shows exactly the same plot, but ignoring lifetime welfare of the top 1% earnings households in the calculation of aggregate welfare. Finally the red line displays "long run welfare" measured as expected lifetime utility of households born into the steady state associated with a particular top marginal tax rate  $\tau_h$ . In all cases, as  $\tau_h$  is varied, so is the earnings threshold  $\bar{y}_h$  such that the top marginal tax rate consistently applies to the top 1% of the earnings distribution, at least in the initial period of the policy-induced transition path.

The associated entry marginal tax rate  $\tau_l$  required to balance the intertemporal government budget is plotted against  $\tau_h$  in figure 4, together with a plot of  $\tau_h$  against itself, i.e. the 45° line. Not surprisingly, since an increase in the highest marginal tax rate generates additional revenues for the government, the entry tax rate can fall and households pay lower taxes at the bottom of the distribution. Note that we have made the assumption that when  $\tau_l$  hits 0 as  $\tau_h$  increases, the additional tax revenues are used to expand the range of incomes at which this low tax rate applies (rather than to reduce  $\tau_l$  to a negative range).

Figure plots the associated 5 earnings thresholds (as fraction of median earnings for the lower bend point, as fraction of mean earnings for the higher bend point) at which the low and the high marginal tax rates apply, respectively. The upper line shows that as the top marginal tax rate  $\tau_h$  increases, due to endogenous responses in labor supply the earnings distribution compresses and

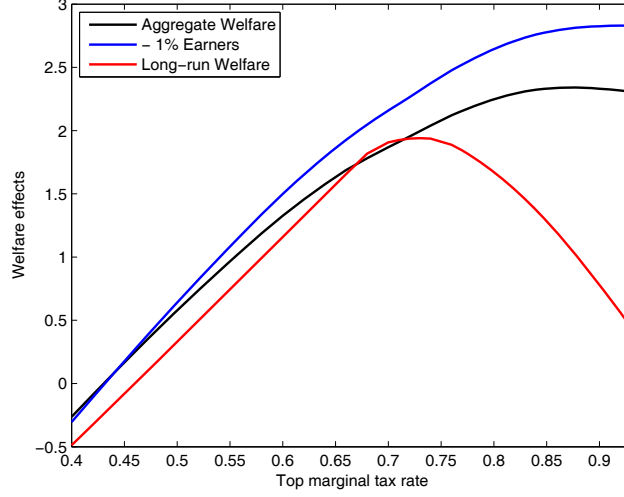


Figure 3: Three Aggregate Welfare Measures, a Functions of  $\tau_h$

the top 1% of earnings are obtained at a lower fraction of mean earnings. Thus the red line in figure 5 falls. The threshold for the entry marginal tax rate  $\tau_l$  by assumption is held constant until  $\tau_l$  hits zero in figure 4, and then expands, thus broadening the range of earnings for which households pay zero marginal taxes.

Figure 3 shows that the optimal marginal tax rate applying to the top 1% of earnings is indeed very high, in excess of 90% for the welfare measure that ignores the top 1% of earners in the economy. To iterate, again note that implementing such a tax rate is *not costless* for the bottom 99% of the population, as it triggers a severe decline in aggregate wages along the transition. **[Add GDP plot against  $\tau_h$ , either for  $t = 1$  or  $T$ , or preferably for both  $t$  in the same plot. Normalize GDP in initial steady state to 1].** Interestingly, even when we include the top 1% of households in our welfare measure, the optimal marginal tax rate at the top is still very high, at 87% of earnings. And even if we restrict attention to steady state welfare (and thus on the long-run consequences of a lower capital stock and wages induced by higher tax progressivity) we find an optimal marginal tax rate of  $\tau_h = 73\%$ . Not surprisingly, ignoring the top 1% of earnings households in the social welfare function leads to higher optimal tax progressivity, and ignoring the benefits of using part of the capital stock for consumption during the transition lowers optimal tax progressivity. In our view the most salient feature of our result thus far is, however, that independent of the welfare metric applied the optimal top marginal tax rate  $\tau_h$  is high, and certainly significantly higher than in the current U.S. status quo. Therefore we study, in the next subsections, the reasons for and robustness of

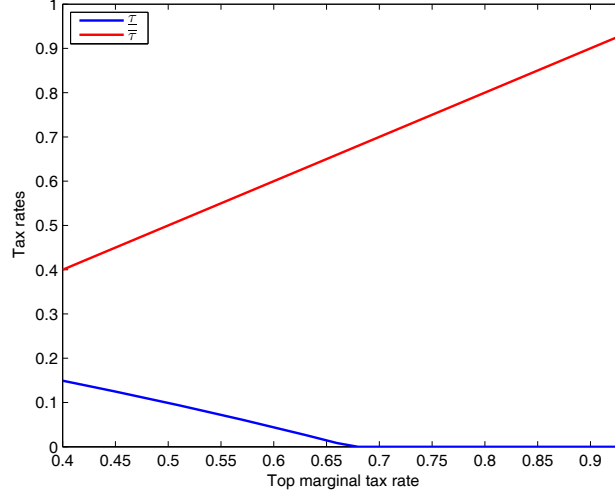


Figure 4: Marginal Tax Rates Associated with  $\tau_h$

this very stark result.

**Transition plots showing what happens to economy over time.**

### 5.2.1 Where do the Welfare Gains Come From?

To start investigating the source of the welfare gains from the tax reform, figure 6 plots the marginal tax rates faced by households in the status quo and in the optimal system, where optimal is defined by the aggregate welfare measure that *includes* the top 1% of the population. Figure 7 does the same for average tax rates. Both figures display clearly that the large majority of the population is paying lower tax rates on their earnings: all households below about five times median earnings faces lower tax rates after the tax reform.

Distribution of after-tax labor earnings

Welfare consequences by cohort. The welfare gains are declining in the horizon of the cohort, reflecting that over time the economy shrinks and the transition gains are eaten up.

### 5.2.2 Realistic Income Inequality is Key for the Results

Suppose instead households face a labor productivity process that does *not* contain the small chance of very high wage and thus earnings realizations.<sup>9</sup>

<sup>9</sup>One interpretation of this version of the model is that it describes the 1960's and early 1970's, the period prior to the large increase in the income share of the top 1% of the distribution.

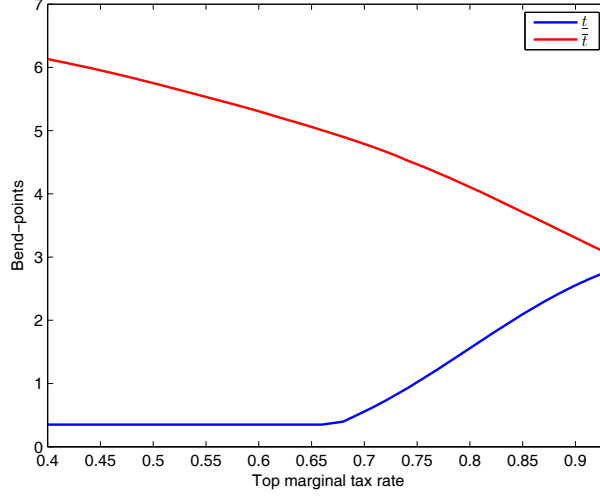


Figure 5: Income Thresholds for Marginal Tax Rates  $\tau_l$  and  $\tau_h$ .

By implication, in this version of the model the earnings, income and wealth distributions will not display the degree of concentration observed in U.S. data, and thus it won't paint an accurate picture of who the top 1% are and what are their economic circumstances. This economy serves, however, a useful role for understanding what drives our results of desirable high marginal income tax rates for the top earners in society.

### 5.3 Progressive Labor Income Taxes or Capital Income Taxation

## 6 Sensitivity Analysis

### 6.1 The Form of the Social Welfare Function

### 6.2 Other Mechanisms to Generate Large Earnings and Wealth Inequality

## 7 Conclusion

## References

- [1] Aiyagari (1995)

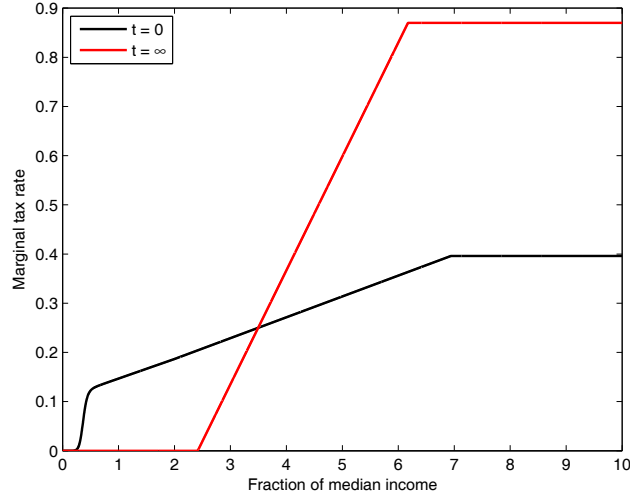


Figure 6: Marginal Taxes in the Status Quo and the Optimal System

- [2] Badel, A. and Mark Huggett (2013) Taxing Top Earners: A Human Capital Perspective
- [3] Bassetto, M. (1999), Optimal Fiscal Policy with Heterogeneous Agents, Working paper.
- [4] Cagetti on (after tax) interest elasticity of savings
- [5] Cagetti and De Nardi on entrepreneurs and wealth taxation, and on bequests
- [6] Castaneda, A., J. Diaz-Gimenez and V. Rios-Rull (2003), "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 111, 818-857
- [7] Corak, M. (2013) *Journal of Economic Perspectives*
- [8] Davila, Hong, Krusell and Rios Rull *Econometrica* forthcoming
- [9] Diamond, P. and E. Saez (2011), The Case for a Progressive Tax: From Basic Research to Policy Recommendations, *Journal of Economic Perspectives*
- [10] Diaz-Gimenez, J., A. Glover and V. Rios-Rull Facts on the Distributions of Earnings, Income, and Wealth in the United States: 2007 Update
- [11] Domeij, D. and J. Heathcote (2004), On the Distributional Effects of Reducing Capital Taxes, *International Economic Review*, 45, 523-554.

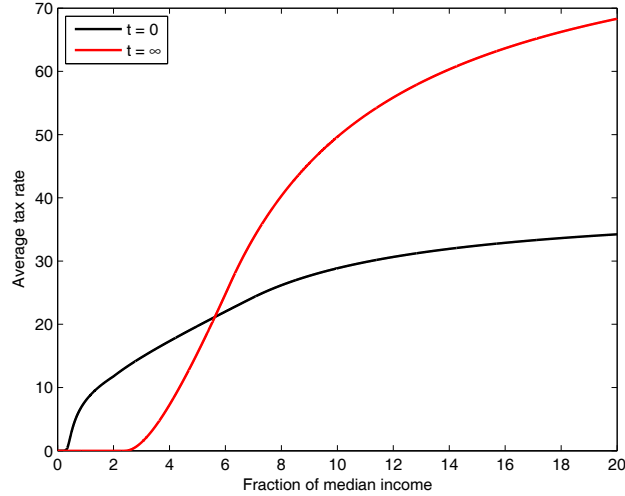


Figure 7: Average Taxes as a Function of Median Earnings

- [12] De Nardi, French and Jones on health shocks and the wealth distribution
- [13] Erosa and Gervais
- [14] Jacobs, B. and D. Schindler (2012), On the desirability of taxing capital income in optimal social insurance, *Journal of Public Economics*, 96, 853–868
- [15] Krueger and Ludwig (2013) on
- [16] Krusell and Smith JPE paper
- [17] Mankiw, G. (2013), *Journal of Economic Perspectives*
- [18] OECD (2010) Education at a Glance
- [19] Piketty and Saez (2013) on concentration of earnings and the taxation of the top 1% *Journal of Economic Perspectives*
- [20] Rios-Rull and friends generating appropriate wealth distribution with high income shocks
- [21] Storesletten, K., C. Telmer and A. Yaron (2004) *Journal of Monetary Economics*
- [22] Trabandt, M. and H. Uhlig (2011), The Laffer Curve Revisited, *Journal of Monetary Economics*, 58, 305-327

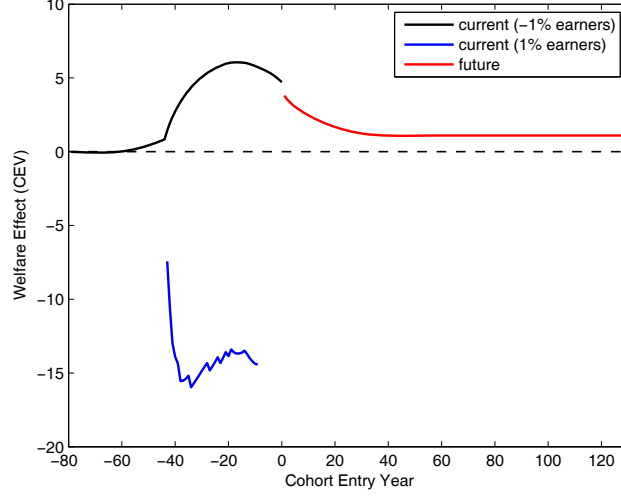


Figure 8: Average CEV for Different Age Cohorts

- [23] Vogelgesang, U. (2000), Optimal Capital Income Taxation and Redistribution, Finanzarchiv, 57, 412-434.

## 8 Appendix A: Details of the Computational Approach

## 9 Appendix B: Definition of Invariant Probability Measure

First we construct the share of the population in each age group. Let  $\tilde{\mu}_1 = 1$ , and for each  $j \in \{2, \dots, J\}$  define recursively

$$\tilde{\mu}_j = \frac{\psi_j \tilde{\mu}_{j-1}}{1 + n}.$$

Then the share of the population in each age group is given by

$$\mu_j = \frac{\tilde{\mu}_j}{\sum_t \tilde{\mu}_t}.$$

Next, we construct the measure of households of age 1 across characteristics  $(s, \alpha, \eta, a)$ . By assumption (see the calibration section, section ?? of the paper) newborn households enter the economy with zero assets,  $a = 0$  and at the mean idiosyncratic productivity shock  $\bar{\eta}$ . The share of college-educated households is

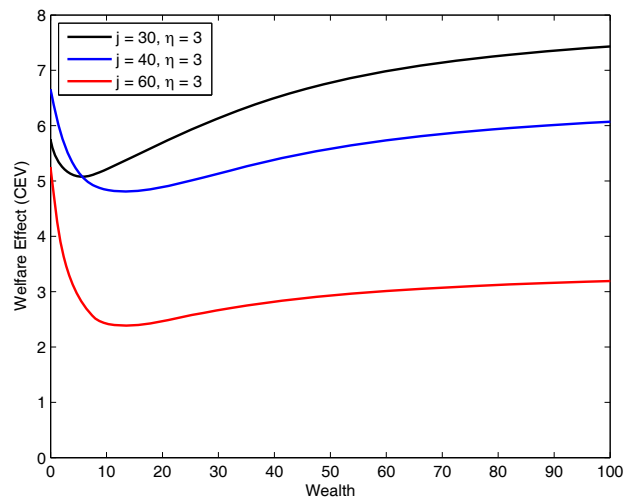


Figure 9:

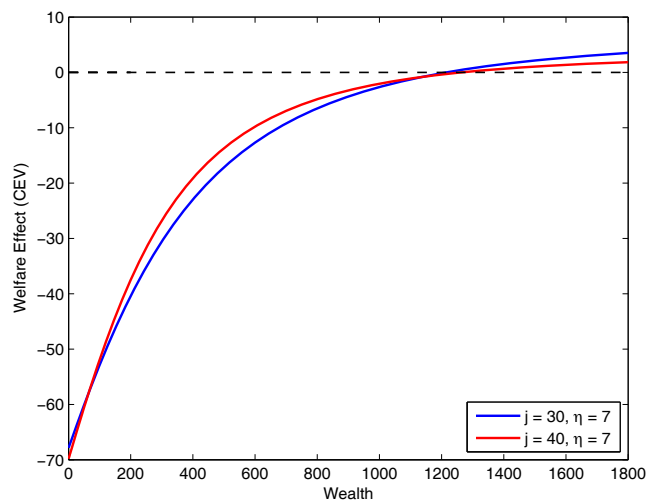


Figure 10:

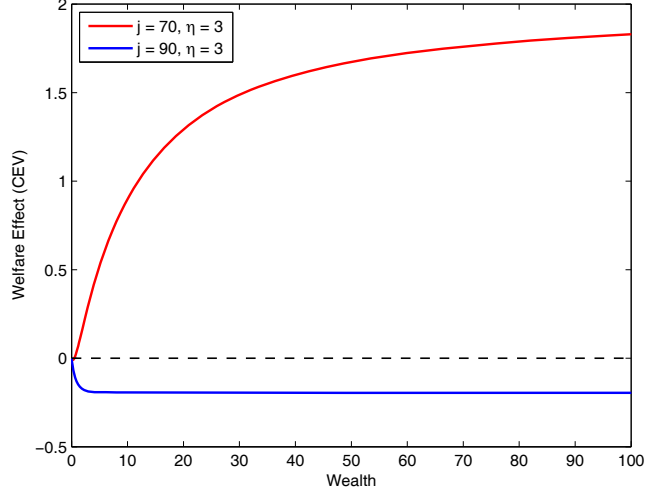


Figure 11:

exogenously given by  $\phi_c$  and  $\phi_n = 1 - \phi_c$ , and the fixed effect is drawn from a discrete pdf  $\phi_s(\alpha)$ . Thus

$$\Phi(\{j = 1\}, \{\alpha\}, \{s\}, \{\bar{\eta}\}, \{0\}) = \mu_1 \phi_s \phi_s(\alpha)$$

for  $s = \{n, c\}$  and zero else.

Finally we construct the probability measure for all ages  $j > 1$ . For all Borel sets of assets  $\mathcal{A}$  we have

$$\Phi(\{j+1\}, \{\alpha\}, \{s\}, \{\eta'\}, \mathcal{A}) = \frac{\psi_{j+1} \pi_s(\eta'|\eta)}{1+n} \int \mathbf{1}_{\{a'(j,s,\alpha,\eta,a) \in \mathcal{A}\}} \Phi(\{j\}, \{\alpha\}, \{s\}, \{\eta\}, da)$$

where

$$\int \mathbf{1}_{\{a'(j,s,\alpha,\eta,a) \in \mathcal{A}\}} \Phi(\{j\}, \{\alpha\}, \{s\}, \{\eta\}, da)$$

is the measure of assets  $a$  today such that, for fixed  $(j, s, \alpha, \eta)$ , the optimal choice today of assets for tomorrow,  $a'(j, s, \alpha, \eta, a)$  lies in  $\mathcal{A}$ .

## 10 Appendix C: Details of the Calibration

### 10.1 Markov Chain for Labor Productivity

The Markov chain governing idiosyncratic labor productivity for both education groups is given by

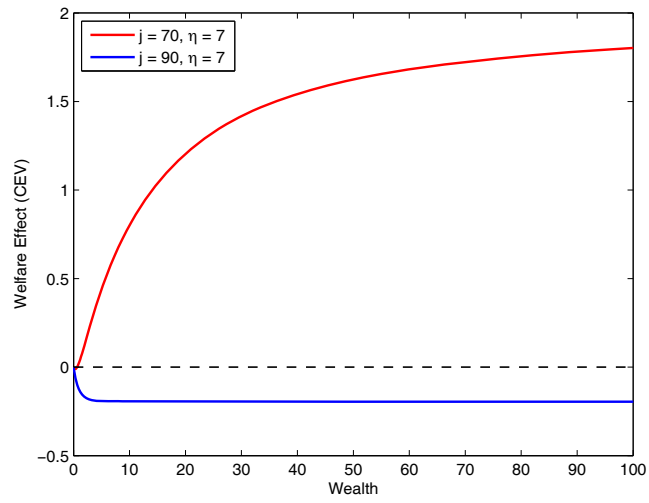


Figure 12:

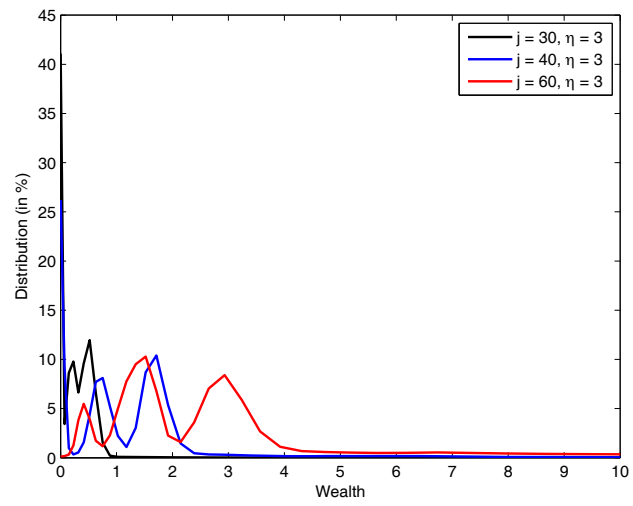


Figure 13:

| $s = n$            |          |          |          |          |          |          |           |
|--------------------|----------|----------|----------|----------|----------|----------|-----------|
| $i, j$             | 1        | 2        | 3        | 4        | 5        | 6        | 7         |
| 1                  | 0.969680 | 0.029310 | 0.000332 | 0.000002 | 0.000000 | 0.000676 | 0.000000  |
| 2                  | 0.007328 | 0.969846 | 0.021984 | 0.000166 | 0.000000 | 0.000676 | 0.000000  |
| 3                  | 0.000055 | 0.014656 | 0.969901 | 0.014656 | 0.000055 | 0.000676 | 0.000000  |
| 4                  | 0.000000 | 0.000166 | 0.021984 | 0.969846 | 0.007328 | 0.000676 | 0.000000  |
| 5                  | 0.000000 | 0.000002 | 0.000332 | 0.029310 | 0.969680 | 0.000676 | 0.000000  |
| 6                  | 0.000000 | 0.000000 | 0.024837 | 0.000000 | 0.000000 | 0.949980 | 0.025183  |
| 7                  | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.267852 | 0.732148  |
| $\exp(\eta_{n,i})$ | 0.1159   | 0.3405   | 1.0000   | 2.9369   | 8.6255   | 15.8180  | 1284.3139 |

and

| $s = c$            |          |          |          |          |          |          |           |
|--------------------|----------|----------|----------|----------|----------|----------|-----------|
| $i, j$             | 1        | 2        | 3        | 4        | 5        | 6        | 7         |
| 1                  | 0.957899 | 0.028954 | 0.000328 | 0.000002 | 0.000000 | 0.012817 | 0.000000  |
| 2                  | 0.007239 | 0.958063 | 0.021717 | 0.000164 | 0.000000 | 0.012817 | 0.000000  |
| 3                  | 0.000055 | 0.014478 | 0.958118 | 0.014478 | 0.000055 | 0.012817 | 0.000000  |
| 4                  | 0.000000 | 0.000164 | 0.021717 | 0.958063 | 0.007239 | 0.012817 | 0.000000  |
| 5                  | 0.000000 | 0.000002 | 0.000328 | 0.028954 | 0.957899 | 0.012817 | 0.000000  |
| 6                  | 0.000000 | 0.000000 | 0.028087 | 0.000000 | 0.000000 | 0.969688 | 0.002225  |
| 7                  | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.000000 | 0.267852 | 0.732148  |
| $\exp(\eta_{c,i})$ | 0.2112   | 0.4595   | 1.0000   | 2.1761   | 4.7353   | 7.3949   | 1284.3139 |

## 10.2 The Social Security System

We use the US pension formula to calculate pension payments. Specifically, for a given average labor earnings  $\bar{y} = \bar{y}(s, \alpha, \eta = \eta_{65})$  we set

$$p(s, \alpha, \eta) = f(\bar{y}(s, \alpha, \eta = \eta_{65})) = \begin{cases} r_1 \bar{y} & \text{if } \bar{y} < b_1 y^{med} \\ r_1 b_1 y^{med} + r_2 (\bar{y} - b_1 y^{med}) & \text{if } \bar{y} < b_2 y^{med} \\ r_1 b_1 y^{med} + r_2 (b_2 - b_1) y^{med} + r_3 (\bar{y} - b_2 y^{med}) & \text{otherwise} \end{cases}$$

Here  $r_1, r_2, r_3$  are the respective replacement rates and  $b_1$  and  $b_2$  the bend points. We express these points in terms of median household income  $y^{med}$  which is the median of income from labor and assets (including bequests and pension payments). We use  $y^{med} = 50,000$  as a reference value for this (see US Census Bureau for 2009). Consequently, the bend points are  $b_1 = 0.184$  and  $b_2 = 1.144$  and the respective replacement rates are  $r_1 = 0.90$ ,  $r_2 = 0.32$  and  $r_3 = 0.15$ . The maximum amount of pension benefit a household can receive is therefore 30, 396, or 0.608 times the median income. All data is taken from the information site of the social security system for 2012. Finally, we calibrate the contribution cap of the pension system  $\bar{y}_{ss}$  in order to obtain a contribution rate of 12.4 percent.