Revisiting Tax on Top Income

by

Ayse Imrohoroglu

and

Cagri Kumru

and

Arm Nakornthab

ANU Working Papers in Economics and Econometrics
660

February 2018

JEL: D31, E21, H2

ISBN: 0 86831 660 1

Revisiting Tax on Top Income*

Ayşe İmrohoroğlu
University of Southern California †

Çağrı S. Kumru
Australian National University ‡

Arm Nakornthab
Australian National University §

February 19, 2018

Abstract

In this paper, we study optimal income taxation in a model with entrepreneurial activity. We conduct two types of changes in tax policy: changing the overall progressivity of taxes versus changing the tax rate of the richest 1% of the population. We study the implications of these tax policies on welfare, inequality, and government revenues. Our results indicate that increasing the overall progressivity of taxes results in lower wealth inequality and higher welfare relative to increasing the tax rate on the richest 1% of the population.

JEL Classification: D31, E21, H2
Keywords: Entrepreneurship, taxation, progressivity, labor supply

*We thank the seminar participants at the European University Institute, the University of Queensland and the University of Tasmania for their comments.

†Marshall School of Business, University of Southern California, Los Angeles, CA 90089-0804, USA Email: ayse@marshall.usc.edu

‡Research School of Economics, Australian National University, Canberra, ACT 2601, Australia. Email: cagri.kumru@anu.edu.au

§Research School of Economics, Australian National University, Canberra, ACT 2601, Australia. Email: arm.nakornthab@anu.edu.au

1 Introduction

Optimal taxation of top income has gained a lot of attention as a natural consequence of the increasing inequalities in income and wealth in the U.S. in the last 40 years.[1] This was also a period when marginal income tax rates decreased substantially.[2] Diamond and Saez (2011) survey the static models of optimal taxation of labor income and argue that the marginal tax rate on top earners in the U.S. should be about 73%. A number of studies from the quantitative dynamic tax literature reach significantly different conclusions about optimal tax rates on top income. For example, while Badel and Huggett (2015) and Guner et al. (2016) find optimal marginal tax rates for top income earners to be around 49% and 42% respectively, Kindermann and Krueger (2017) report an optimal tax rate on the top 1% of earners that is more than 90%.[3]

In this paper, we examine the optimal taxation of top income earners in a model with an entrepreneurial decision making process in the spirit of Cagetti and De Nardi (2009) and Quadrini (2000). There are two major reasons why modeling entrepreneurial activities may be important for this question. First, although entrepreneurs represent only 7% of the population, they earn 17% of the total income. Second, 40% of the top 1% of income earners are entrepreneurs (see Survey of Consumer Finances (2010)). Hence, taking entrepreneurs into account could be potentially important for understanding the implications of top income taxation on welfare, inequality, and government revenues.[4]

We develop a simplified life-cycle model with stochastic aging that includes entrepreneurial activity, individual heterogeneity, and endogenous labor supply. Individual heterogeneity stems from differences in entrepreneurial abilities and uninsurable productivity shocks. A young individual can choose to be either a worker or an entrepreneur while an old individual can choose to stay as an entrepreneur provided that she/he was previously an entrepreneur or become a retiree. The existence of entrepreneurs helps the model generate a wealth inequality that mimics the data well.[5] We use a parametric tax function proposed by Benabou

[1]Piketty and Saez (2003) and Alvaredo et al. (2013) provide empirical evidence for this trend.

[2]Alvaredo et al. (2013) state that the United States experienced a reduction of 47 percentage points in its top income (federal and local income) tax rate between 1960 and 2009.

[3]See also Ales et al. (2017); Ales and Sleet (2016).

[4]Smith et al. (2017) show that private business owners who actively manage their firms are key for top income inequality. They also show that private business income accounts for most of the rise of top incomes since 2000. In our model the entrepreneurship sector corresponds to the private business owners in Smith et al. (2017) who actively manage their firms.

[5]As a result, models with entrepreneurs are used to explain the implications of tax policies and financial

(2002) to model the relationship between income and income taxes paid at the federal level as well as a flat rate income tax that captures state and local taxes, a flat corporate income tax, and a flat consumption tax. The distribution of income, wealth, and the share of tax payments by different income quantiles that is generated by the model mimic the data reasonably well. In this framework we conduct two types of changes in tax policy: changing the overall progressivity of taxes versus changing the tax rate of the richest 1% of the population. We examine the impact of these changes on output, government revenues, wealth inequality, and welfare. First, we shift the tax burden toward high income earners by increasing the value of the progressivity parameter and calculate the welfare and revenue-maximizing marginal and average tax rates. Second, we calculate the welfare and revenue-maximizing tax rate by changing the tax rate for the top 1% of income earners only. Our results indicate substantial differences in output, government revenues, wealth inequality, and welfare across these experiments.

We find that if the purpose of the government is to maximize government revenues, then increasing the tax rate on the richest 1% of the population to 55% is more effective then increasing the overall progressivity of taxes. Changing the overall progressivity of taxes yields the revenue-maximizing effective marginal tax rate for the richest 1% of households to be 33.1%. In this case, tax revenues from Federal income tax increase by 5.33%. When we search for the revenue-maximizing marginal tax rate that targets the richest 1% of the households, we find the optimal marginal tax rate to be 55%. In this case, revenues from federal income increase by 16.3%. Note that both of these tax rates are higher than the 22.9% used in the benchmark economy that is calibrated to the U.S. economy. The stark differences in revenues that we find are due to the impact of the changes in taxes on labor supply, capital stock, and output. When the tax rate on the richest 1% of the population is increased to 55%, labor supply, capital, and output in the economy decline by 0.7%, 8.2%, and 3.9%, respectively. If overall progressivity is changed, however, labor supply, capital, and output decline by 1.1%, 15.1%, and 6%, respectively. As we noted before, about 40% of the top earners are entrepreneurs. Therefore, an increase in the overall progressivity of taxes affects a larger number of entrepreneurs compared to an increase in the tax rate facing the richest 1% of the individuals. The capital stock and hours worked by the entrepreneurs

frictions. See Buera et al. (2015) for a detailed review of the literature on entrepreneurship. In models without entrepreneurial activity, an appropriate calibration of the income process incorporating a luck factor, as in Castaneda et al. (2003), is needed to achieve a meaningful distribution of earnings and wealth.

takes a larger hit in this case. Consequently, output declines more and government revenues increase less when the overall progressivity of taxes is increased. If, instead, only the tax rate for the richest 1% is increased, the capital stock and output decline less, resulting in larger tax collections.

If the purpose of the government is welfare maximization, however, we find an increase in the overall progressivity of taxes to be more effective. Welfare-maximizing progressivity parameter results in a marginal tax rate of 42.2% for the richest 1% of households. This is lower than the optimal marginal tax rate found for this group (55%) if the government only targets the tax rate of the richest 1% of the population. Households in the lower income distribution benefit more when the overall progressivity is changed. Wealth inequality goes down and lower income households are able to smooth their consumption better. In fact, increasing the tax rate on the richest 1% of the households has a negligible effect on the wealth distribution whereas increasing the overall progressivity of taxes reduces the wealth Gini from 0.84 to 0.79.

We are also able to clarify some of the differences in the findings of seemingly similar papers on this topic. Badel and Huggett (2015) analyze the change in the top tax rate on general income (capital and labor) without changing the tax rate schedule below the bracket. They search for both revenue and welfare-maximizing rates. Guner et al. (2016), on the other hand, alter the overall progressivity of income (capital+labor) tax function to find revenue maximizing tax schemes. Kindermann and Krueger (2017) alter the top tax rate on labor earnings and calculate both revenue and welfare-maximizing rates. Guner et al. (2016) report that the revenue-maximizing progressivity parameter implies an effective federal income marginal tax rate of 36.6% for the richest 5% of households. They also report that the revenue-maximizing federal income marginal tax rate that applies to the top 5% is 42%. Badel and Huggett (2015) report that the peak of Laffer curve happens at a tax rate of 49% for the top 1%. In our framework, the revenue-maximizing progressivity parameter implies 27.8% and 33.1% effective marginal tax rates for the top 5% and 1% of the individuals, respectively. All of these findings are significantly different from the over the 90% optimal revenue and welfare-maximizing tax rate for the richest 1% of the population found in Kindermann and Krueger (2017). In our experiments where only the marginal tax rate on the top 1% of the population is changed, we find the revenue and welfare-maximizing tax rate to be 55%.[6] In our counterfactual experiments, we show that the source of the income

[6]Using a similar model, Bruggemann (2017) finds the welfare maximizing tax rate for the top 1% of the

that is taxed plays an important role in the findings in Kindermann and Krueger (2017). In their experiments, they only alter the tax rate that is applicable to labor income. In our paper as well as in Guner et al. (2016), both labor and capital income are subject to the same tax rate as it is in the current tax code. In a counterfactual experiment, we find the welfare-maximizing tax rate on the top 1% of the population to be 80% when only the labor income is subject to this tax.

This paper is organized as follows. Section 2 explains the model. Section 3 describes a calibration procedure. Section 4 discusses the features of the benchmark economy. Section 5 evaluates the model experiments. Section 6 discusses our sensitivity analysis. Section 7 concludes.

2 Model

2.1 Demographics

The model is a simplified life-cycle model with intergenerational altruism. The model period is one year. We assume that there are young and old cohorts in the economy, and aging is stochastic. The young stay young with a constant probability π_y and get old with a probability $1 - \pi_y$ in the next period. The old continue to live with a constant probability π_o and die with a probability $1 - \pi_o$ in the next period. These probabilities are calibrated to match the proportion of young and old households in the economy. When the old die, their offsprings receive after tax bequests and enter the economy in the next period. For simplicity, we assume that each household has only one offspring. We do not consider cases that differentiate between single households, married households, or households with no offspring. The measure of households is normalized to 1.

2.2 Preferences

We assume that preferences are time-separable with a discount factor β. The instantaneous utility function is given by:

$$u(c_t, 1 - l_t) = \frac{c_t^{1-\sigma_1}}{1 - \sigma_1} + \chi \frac{(1 - l_t)^{1-\sigma_2}}{1 - \sigma_2}$$

population to be 52.5%.

where c_t is consumption and l_t is labor supply. We assume l_t of the retirees equals zero. The total time endowment is 1. σ_1 is the coefficient of risk aversion and σ_2 is the inverse of Frisch elasticity of labor supply. χ is the disutility from working.

2.3 Technology

Following Quadrini (2000), we assume that there are two production sectors: the corporate sector and the entrepreneurial sector. Each person has two types of ability: working and entrepreneurial. Types are stochastic, positively correlated over time, and uncorrelated with each other. Productivity of an individual as an entrepreneur is given by θ_t, while y_t represents the capacity to produce income out of labor by working in the corporate sector.

At the beginning of the period, current ability levels are revealed to individuals. Next period ability levels remain unknown. A young individual with an an asset level a_t, an entrepreneurial ability θ_t, and a worker productivity y_t makes a decision regarding whether she will be a worker or an entrepreneur in the current period.

Entrepreneurs can borrow subject to a borrowing limit, invest capital, hire labor, and run a technology. The return from the production technology is dependent on an entrepreneurial ability. When the entrepreneur invests k_t, the output is given by:

$$f(k_t, n_t) \quad = \quad \theta_t \left(k_t^\gamma \left(l_t + n_t\right)^{1-\gamma}\right)^\nu, \tag{1}$$

where $0 \leq \gamma \leq 1$ is the share of entrepreneurial capital. $\nu < 1$ indicates the decreasing returns to scale from investing in capital and labor as in Lucas (1978). Capital depreciates at a rate of δ. Entrepreneurs provide their own labor l_t and also hire labor at the amount of $n_t \geq 0$.

The corporate sector is represented by a Cobb-Douglas functional form:

$$F(K_t^c, L_t^c) \quad = \quad A(K_t^c)^\alpha (L_t^c)^{1-\alpha}, \tag{2}$$

where K_t^C and L_t^C are total capital and labor inputs used in the corporate sector. A represents the level of technology and is constant, whereas α is the capital share in the corporate sector.

2.4 Credit Markets

We assume that young workers and old retirees cannot borrow, i.e., $a_{t+1} \geq 0$. The size of capital that entrepreneurs can borrow depends on their current net worth. Default is not an option in this setting. We assume there are no financial intermediaries. Hence, there is no difference between the saving and borrowing interest rates.

2.5 Government

The government is assumed to live forever. It collects taxes, pays a pension benefit p to each retiree, provides goods and services g, and pays interest on the debt, $(1+r_t)D_t$. Households do not derive utility from consumption of government goods and services. During every period, tax revenues are equal to government purchases, pension payments, and interest payments on the debt.

We use Benabou's (2002) functional form to model the progressive income tax schedule. The total amount of federal income tax $T_t(Y_t)$ for total taxable income income Y_t is given by:

$$T_t(Y_t) \quad = \quad (1-\lambda Y_t^{-\tau})Y_t + \tau_t^{bal}Y_t + \tau_t^k r_t a_t \tag{3}$$

where τ_t^{bal} is a proportional income tax rate (other than the federal income tax rate) that captures state and local income taxes. The parameter λ captures the revenue requirement and τ governs the curvature of the tax function. In addition to the income tax, the government collects corporate income and consumption taxes. The corporate income and consumption tax rates are denoted by τ_t^k and τ_t^c, respectively.

Denoting y_H and τ_H as the income threshold for those having income in the top 1% and the marginal tax rate at top 1%, respectively, the tax function takes the following form:

$$T_t(Y_t) \quad = \quad \begin{cases} (1-\lambda Y_t^{-\tau})Y_t + \tau_t^{bal}Y_t + \tau_t^k r_t a_t & \text{if } Y_t < Y_H \\ (1-\lambda Y_H^{-\tau})Y_H + \tau_t^{bal}Y_H + \tau_H * (Y_t - Y_H) + \tau_t^k r_t a_t & \text{if } Y_t > Y_H \end{cases} \tag{4}$$

Our tax function and tax base are the same as those of Guner et al. (2016). Kindermann and Krueger (2017) use a different functional form (two bracket tax function) and choose labor earnings as the income tax base.

2.6 Household's Problem

Households are divided into two groups: young and old. A young individual can choose to be either a worker or an entrepreneur. An old individual can choose to stay as an entrepreneur provided that she/he was an entrepreneur before getting old or become a retiree. If an old individual was a worker before retirement, he cannot become an entrepreneur when he is old. Notice that the bequest the young receive in the next period is what the old decide to save in the current period given by $a_t + 1$. With probability π_o, the old stay alive. With probability $1 - \pi_o$, the old reincarnate as the young and start making economic decisions as the young with initial assets equal to $a_t + 1$.

The value function of a young individual is given by:

$$V_t^Y(a_t, y_t, \theta_t) \quad = \quad \max\left\{V_t^{Y,e}(a_t, y_t, \theta_t), V_t^{Y,w}(a_t, y_t, \theta_t)\right\}. \tag{5}$$

The young individual decides whether to become a worker or an entrepreneur at the beginning of the period. $V_t^{Y,e}(\cdot)$ is the value function of a young individual who becomes an entrepreneur and $V_t^{Y,w}(\cdot)$ is the value function of a young individual who becomes a worker.

The young worker's problem can be written as:

$$V_t^{Y,w}(a_t, y_t, \theta_t) = \max_{c_t, l_t, a_{t+1}} \{u(c_t, 1 - l_t) + \beta\pi_y E_t\left[V_{t+1}^Y(a_{t+1}, y_{t+1}, \theta_{t+1})\right] + \beta(1 - \pi_y)V_{t+1}^{O,r}(a_{t+1})\}, \tag{6}$$

subject to

$$\begin{aligned} Y_t^w &= w_t l_t y_t + r_t a_t, & (7)\\ (1 + \tau_t^c)c_t + a_{t+1} &= w_t l_t y_t + (1 + r_t)a_t - T_t(Y_t^w), & (8)\\ 0 &\leq l_t \leq 1, & (9)\\ 0 &\leq a_{t+1}, & (10) \end{aligned}$$

where w_t is the equilibrium wage rate and r_t is the equilibrium interest rate. The term $V_{t+1}^{O,r}(a_{t+1})$ is the value function of the retirees. The expected value of $V_{t+1}^Y(a_{t+1}, y_{t+1}, \theta_{t+1})$ is conditional on the joint distribution of y_t and θ_t. The young entrepreneur's problem can be written as:

$$V_t^{Y,e}(a_t, y_t, \theta_t) = \max_{c_t, l_t, k_t, n_t, a_{t+1}} \{u(c_t, 1-l_t)+\beta\pi_y E_t\left[V_{t+1}^Y(a_{t+1}, y_{t+1}, \theta_{t+1})\right]+\beta(1-\pi_y)E_t V_{t+1}^O(a_{t+1}, \theta_{t+1})\}, \tag{11}$$

subject to

$$
\begin{aligned}
Y_t^e &= \theta_t \left(k_t^{\gamma}(l_t+n_t)^{1-\gamma}\right)^{\nu} - \delta k_t - r_t(k_t-a_t) - w_t n_t, & (12)\\
(1+\tau_t^c)c_t + a_{t+1} &= Y_t^e - T_t(Y_t^e) + a_t, & (13)\\
0 &\leq a_{t+1}, & (14)\\
0 &\leq n_t, & (15)\\
0 &\leq l_t \leq 1, & (16)\\
0 &\leq k_t \leq (1+d)a_t. & (17)
\end{aligned}
$$

Working capital k_t includes own and borrowed assets. Y_t^e is the entrepreneur's total profit.[7] Following Kitao (2008), we set d as an exogenous borrowing limit. The term $V_{t+1}^{O,e}(a_{t+1},\theta_{t+1})$ is the value function of the old entrepreneur at the beginning of the next period before deciding whether to stay as an entrepreneur or retire. This is different from the young worker's decision since he has no choice but to retire. The expected value of $V_{t+1}^{Y}(a_{t+1},y_{t+1},\theta_{t+1})$ is taken similarly as the young workers. The expected value of $V_{t+1}^{O}(a_{t+1},\theta_{t+1})$ is conditional only on θ_t.

The old individual's problem is given as follows:

$$V_t^O(a_t,\theta_t) = \max\left\{V_t^{O,e}(a_t,\theta_t), V_t^{O,r}(a_t)\right\} \tag{18}$$

V_t^O is the value function of the old individual in the current period before deciding whether to stay as an entrepreneur or retire. $V_t^{O,e}$ is the value function for the old entrepreneur who stays as an entrepreneur and $V_t^{O,r}$ is the value function for the retirees.

The old retiree's problem is given by:

$$V_t^{O,r}(a_t) = \max_{c_t,a_{t+1}} \{u(c_t,1) + \beta\pi_o V_{t+1}^{O,r}(a_{t+1}) + \beta(1-\pi_o)E_t[V_{t+1}^{Y}(a_{t+1},y_{t+1},\theta_{t+1})]\} \tag{19}$$

subject to

$$
\begin{aligned}
(1+\tau_t^c)c_t + a_{t+1} &= (1+r_t)a_t + p - T_t\left(r_t a_t + p\right) & (20)\\
0 &\leq a_{t+1}. & (21)
\end{aligned}
$$

[7] In a recent paper, Fagereng et al. (2016) highlight the importance of heterogeneity in returns on wealth distributions by using Norway's administrative tax records. In our model there is return heteregonity since the returns to capital in the entrepreneurial sector are different from the equilibrium interest rate.

The old retired individual receives a social security transfer payment p in every period. Since the old retiree in this case in not an entrepreneur, the probability of a retiree's offspring being an entrepreneur depends on the joint invariant distribution of y_t and θ_t. The expected value of the offspring's value function is given by: $V^Y_{t+1}(a_{t+1}, y_{t+1}, \theta_{t+1})$.

The old entrepreneur's problem is given by:

$$V^{O,e}_t(a_t, \theta_t) = \max_{c_t, l_t, k_t, n_t, a_{t+1}} \{u(c_t, 1-l_t) + \beta\pi_o E_t[V^O_{t+1}(a_{t+1}, \theta_{t+1})] + \beta(1-\pi_o) E_t\left[V^Y_{t+1}(a_{t+1}, y_{t+1}, \theta_{t+1})\right]\}, \tag{22}$$

subject to

$$\begin{aligned} Y^e_t &= \theta_t \left(k^\gamma_t (l_t + n_t)^{1-\gamma}\right)^\nu - \delta k_t - r_t(k_t - a_t) - w_t n_t, && (23)\\ (1+\tau^c_t)c_t + a_{t+1} &= Y^e_t - T_t(Y^e_t) + a_t, && (24)\\ 0 &\leq a_{t+1}, && (25)\\ 0 &\leq n_t, && (26)\\ 0 &\leq {}_t \leq 1, && (27)\\ 0 &\leq k_t \leq (1+d)a_t. && (28) \end{aligned}$$

An entrepreneur's offspring is born with ability levels (y_{t+1}, θ_{t+1}). The expected value of the offspring's value function with respect to y_{t+1} is computed using the invariant distribution of y_t. However, its expected value with respect to θ_{t+1} is conditional on the parent's θ_t and evolves according to the same Markov process governing the entrepreneurial abilities (see subsection Endowments). This reflects the fact that the offspring inherits her parent's business.

2.7 Equilibrium Definition

Each individual's state vector is given by $\mathbf{s}_t = (a_t, y_t, \theta_t, \xi_t)$. a_t stands for the current asset holdings. Idiosyncratic productivity shocks $y_t \in \mathbb{Y}$. $\theta_t \in \Theta = \{0, \theta_1, \theta_2\}$ is an entrepreneurial ability. We can think of θ_t as an idea to start or maintain a business. An individual with no idea or ability to maintain a business has $\theta_t = 0$. $\xi_t \in \Xi = \{YW, YE, OE, OW\}$ stands for an occupational status: young workers, young entrepreneurs, old entrepreneurs, and old retirees, respectively. The entire state space is given by $\mathbb{S} = \mathbb{R}_+ \times \mathbb{Y} \times \Theta \times \Xi$. We can generate the transition matrix, $\Gamma_t(\mathbf{s}_t, \mathbf{s}_{t+1})$ by using the decision rules that solve the maximization problems and the exogenous Markov process for income and entrepreneurial ability. The transition function provides the probability distribution of the next period's state conditional on the current state.

A stationary equilibrium is given by a risk-free interest rate r_t; wage rate w_t; tax functions $T_t(\cdot)$; tax rates τ_t^c, τ_t^{bal}, and τ_t^k; social security payment p; allocations of consumption $c_t(\mathbf{s}_t)$; labor supply $l_t(\mathbf{s}_t)$; savings $a_t(\mathbf{s}_t)$; investment $k_t(\mathbf{s}_t)$; labor hired by the entrepreneurs $n_t(\mathbf{s}_t)$; and a constant distribution of households over the state variables Φ^* such that given r_t, w_t, and taxes:

- The allocations c_t, a_t, l_t, k_t and n_t solve the individual's optimization problem for each state $\mathbf{s}_t \in \mathbb{S}$.
- $r_t = \frac{\partial F(K_t^c, L_t^c)}{\partial K_t^c} - \delta$: the marginal product of capital net of depreciation in the corporate sector is equal to the risk-free interest rate.
- $w_t = \frac{\partial F(K_t^c, L_t^c)}{\partial L_t^c}$: the marginal product of labor employed in the corporate sector is equal to the wage rate.
- The capital markets clear, i.e.,

$$\int_{\mathbf{s}_t} k_t(\mathbf{s}_t) d\Phi_t(\mathbf{s}_t) + K_t^c + D_t \quad = \quad \int_{\mathbf{s}_t} a_t(\mathbf{s}_t) d\Phi_t(\mathbf{s}_t). \tag{29}$$

- Total assets, $\int_{\mathbf{s}_t} a_t(\mathbf{s}_t) d\Phi_t(\mathbf{s}_t)$, are equal to the sum of the total capital in entrepreneurial sector, $\int_{\mathbf{s}_t} k(\mathbf{s}_t) d\Phi_t(\mathbf{s}_t)$, the total capital in the corporate sector, K_t^c, and the total government debt, D_t. Tightening
- Labor markets clear i.e.

$$\int_{\mathbf{s}_t} n_t(\mathbf{s}_t) d\Phi_t(\mathbf{s}_t) + L_t^c \quad = \quad \int_{\mathbf{s}_t} l(\mathbf{s}_t) d\Phi_t(\mathbf{s}_t). \tag{30}$$

- Total efficient labor, $\int_{\mathbf{s}_t} l(\mathbf{s}_t) d\Phi_t(\mathbf{s}_t)$, is equal to the sum of the total hired labor in the entrepreneurial sector, $\int_{\mathbf{s}_t} n_t(\mathbf{s}_t) d\Phi_t(\mathbf{s}_t)$, and the total labor employed in the corporate sector, L_t^c.
- The sum of income, consumption, and corporate income tax revenues, and net borrowing is equal to the sum of government purchases, total transfers, and interest payments

on debt. π_r is the fraction of retirees in the population and it is determined endogenously. In the steady state, we must have $D_t = \bar{D}$.

$$\int_{\mathbf{s}_t} [T_t(Y^s) + \tau_t^c c_t(\mathbf{s}_t)]\, d\Phi_t(\mathbf{s}_t) + D_{t+1} = g_t + p\pi_r + (1 + r_t)D_t. \tag{31}$$

- The invariant distribution of individuals is given by Φ, where:

$$\Phi'_{t+1} = \Gamma_t(\mathbf{s}_t, \mathbf{s}_{t+1})' \Phi'_t. \tag{32}$$

In the steady state, $\Phi_t = \Phi^*$.

3 Calibration

In this section, we explain how we map the model initial steady state to the data. Table 1 shows the model parameters that we choose exogenously so that they are not used to match the moments in the data. Parameters in Table 2 are chosen such that the model-generated moments from the initial steady state are matched with their corresponding moments in the data. In modeling high productivity workers, we follow Kindermann and Krueger (2017), which helps generating right income distributions. In order to match the percentage of entrepreneurs at the 1% of income distribution successfully, we need to have superstar entrepreneurs. Hence, we extend Cagetti and De Nardi's (2009) entrepreneurial ability transition matrix by incorporating superstar entrepreneurs.

Preferences

The coefficient of relative risk aversion, σ_1, is set to 1.5, which is a common value in the literature.[8] The parameter σ_2 is set to 1.67 to make the Frisch elasticity equal to 0.6. The disutility from working parameter, χ, is chosen to match work hours of 1/3 of time endowment. Discount factor β is chosen to match the capital to output ratio of 2.9.

Demographics

Since the model period is set to be one year, the probability of aging and death are given as $1-\pi_y$ and $1-\pi_o$. These probabilities are chosen in such a way that the average working and retirement periods are 45 and 11 years, respectively. This implies that in the equilibrium, 80% of the population are young individuals.

[8] See, for example, Attanasio et al. (1999).

Technology

The capital share in corporate sector, α, is set to 0.33 as in Kindermann and Krueger (2017). Level of technology, A, is normalized to one. The depreciation rate, δ, is set to 0.06 as in Stokey and Rebelo (1995). The entrepreneurial exogenous borrowing constraint, d, is set to 0.5 as in Kitao (2008), which implies that entrepreneurs cannot borrow more than 1.5 times their current assets. The degree of decreasing returns to scale, ν, is set to 0.88 as in Bassetto et al. (2015). The entrepreneurial capital share, γ, is chosen to equal 0.45.

Table 1: Fixed Parameters

Parameter		Value
Preferences, technology, and demographics		
Risk aversion	σ_1	1.5
Inverse of Frisch elasticity	σ_2	1.67
Capital share	α	0.33
Technology	A	1
Probability of staying young	π_y	0.978
Probability of staying old	π_o	0.911
Depreciation	δ	0.06
Entr. return to scale	ν	0.88
Entr. borrowing constraint	d	0.5
Labor income process and social security payments		
Autocorrelation	ρ	0.958
Pension/average annual income	p	40%
Public purchases, government debt, and taxes		
Fraction of government spending to output	g	0.035
Fraction of government debt to total capital	D	0.27
Consumption tax	τ_c	5%
Capital tax	τ_k	7.4%
State and local tax	τ^{bal}	5%
Revenue requirement	λ	0.911
Tax progressivity	τ	0.053

Endowments

In order to generate income and wealth distributions and the share of entrepreneurs at the top 1% of income realistically, we introduce highly productive workers and highly successful entrepreneurs to the model. In every period, a worker is endowed with one unit of time to be used as a leisure and work time. One unit of work time yields a wage earning wy, where y is the idiosyncratic labor productivity.

We assume that y can take 6 values. The first five $\{y_1, ..., y_5\}$ are associated with normal labor earnings and y_6 represents very high earnings observed in household data sets such as the SCF. For the normal labor productivity states $\{y_1, ..., y_5\}$ we use a discretized Markov chain of a continuous AR (1) process with persistence ρ and standard deviation, σ_y. We use Rouwenhorst's method in discretization, and set $\rho = 0.958$ as in Kaplan (2012). We assume that the income process and the entrepreneurial ability processes evolve independently. The following 6 x 6 transition matrix captures the very high earning realizations where from any lower state there is a small probability, π_6, to jump to a high productivity state that has earnings realization of y_6. Thus, each individual has the same probability of reaching this high productivity. At the same time, we assume that with probability $1 - \pi_{66}$ workers in a high productivity state can fall below to the median earnings state, y_3.

Overall, there are four parameters - $\sigma_y, y_6, \pi_6, \pi_{66}$ - to be calibrated. We choose σ_y, y_6, π_6, and π_{66} to be 0.18, 11.5, 0.002, and 0.931 to match income and wealth distributions and the Gini coefficient of labor earnings of 0.51 in SCF(2010) data. The exact numerical values for the transition matrix is shown in the appendix.[9]

$$\mathbf{P}_y = \begin{bmatrix} \pi_{11}(1-\pi_6) & \pi_{12}(1-\pi_6) & \pi_{13}(1-\pi_6) & \pi_{14}(1-\pi_6) & \pi_{15}(1-\pi_6) & \pi_6 \\ \pi_{21}(1-\pi_6) & \pi_{22}(1-\pi_6) & \pi_{23}(1-\pi_6) & \pi_{24}(1-\pi_6) & \pi_{25}(1-\pi_6) & \pi_6 \\ \pi_{31}(1-\pi_6) & \pi_{32}(1-\pi_6) & \pi_{33}(1-\pi_6) & \pi_{34}(1-\pi_6) & \pi_{35}(1-\pi_6) & \pi_6 \\ \pi_{41}(1-\pi_6) & \pi_{42}(1-\pi_6) & \pi_{43}(1-\pi_6) & \pi_{44}(1-\pi_6) & \pi_{45}(1-\pi_6) & \pi_6 \\ \pi_{51}(1-\pi_6) & \pi_{52}(1-\pi_6) & \pi_{53}(1-\pi_6) & \pi_{54}(1-\pi_6) & \pi_{55}(1-\pi_6) & \pi_6 \\ 0 & 0 & 1-\pi_{66} & 0 & 0 & \pi_{66} \end{bmatrix}$$

We also introduce highly successful entrepreneurs, whose entrepreneurial ability, θ_2, is much higher than that of standard entrepreneurs, θ_1. Notice that $\theta_0 = 0$ captures the no entrepreneurial ability. We choose $\theta_1 = 1.8$ and $\theta_2 = 2.75$. $\pi(\theta_i|\theta_j)$ is the probability of having the ability, θ_i, conditional on having the ability, θ_j, in the previous period. This implies no worker can be a highly successful entrepreneur before becoming a standard entrepreneur. Similarly, no highly successful entrepreneur can be a worker without being a standard entrepreneur first. Finally, we impose $\pi(\theta_0|\theta_0) = \pi(\theta_2|\theta_2)$ to reduce the number of calibrated parameters. Hence, we choose $\pi(\theta_0|\theta_0) = 0.9775, \pi(\theta_1|\theta_1) = 0.759925$, and

[9]Kindermann and Krueger (2017)'s top productivity is over 1000 and it is quite transitory while in our model the top productivity level is much smaller but more persistent. We would have less persistent top income by choosing much higher productivity. In contrast to Kindermann and Krueger, in our model, the existence of entrepreneurship generates the right wealth and income distributions.

$\pi_(\theta_2|\theta_1) = 0.000075$. In total, there are 5 parameters, including θ_1 and θ_2, to be calibrated for the entrepreneurial ability process. The entrepreneurial transition matrix is given by:

$$\begin{bmatrix} \pi(\theta_0|\theta_0) & \pi(\theta_1|\theta_0) & \pi(\theta_2|\theta_0) \\ \pi(\theta_0|\theta_1) & \pi(\theta_1|\theta_1) & \pi(\theta_2|\theta_1) \\ \pi(\theta_0|\theta_2) & \pi(\theta_1|\theta_2) & \pi(\theta_2|\theta_2) \end{bmatrix} = \begin{bmatrix} 0.98 & 0.023 & 0 \\ 0.24 & 0.76 & 0.000075 \\ 0 & 0.025 & 0.9775 \end{bmatrix} \tag{33}$$

Government Policies

The social security replacement rate, p, is set to 40% of average gross income as in Kotlikoff et al. (1999). The fraction of government debt to total capital, D, is set to equal 0.27 as in Bassetto et al. (2015). The fraction of government spending to output, g, is chosen to satisfy the budget and τ^{bal} and τ_k are fixed at 5% and 7.4%, respectively, as in Guner et al. (2016). The tax rate on consumption, τ_c, is set to equal 5% as in Kindermann and Krueger (2017). The Benabou's tax function parameters, λ, which represents the revenue requirement, and τ, which represents the overall progressivity of taxes, are set to equal 0.911 and 0.053 as in Guner et al. (2016). These estimates imply an average federal tax rate of 8.9% and marginal federal tax rate of 13.7% for households with mean income.

Table 2 summarizes the parameters calibrated to match the seventeen targets in the data that are presented in the next section.

Table 2: Calibrated Parameters

Calibrated parameter		Value
Discount factor	β	0.9396
Entrepreneurial ability	$\{\theta_0, \theta_1, \theta_2\}$	$\{0, 1.8, 2.75\}$
Entr. transition probabilities	see eq. 33	
Entr. capital share	γ	0.45
Disutility from working	χ	1.9
Standard deviation of productivity shock	σ_y	0.18
Value of highest productivity	y_6	11.5
Probability of having highest productivity	π_6	0.002
Probability of staying highest productivity	π_{66}	0.9307

4 Features of the Benchmark Economy

In this section, we discuss the aggregate and distributional properties of the benchmark economy. In order to conduct meaningful policy experiments regarding changes in the progressivity and the top tax rate, we need to make sure that the model delivers realistic income and wealth distributions. Table 3 compares the model-generated moments with those in the data.[10]

Table 3: Target Moments

Targets	Data	Model
Capital to output ratio	2.9	2.9
% Entrepreneurs	7.5-7.6	7.2
% Exiting entrepreneurs	22-24	24
% Workers to entrepreneurs	2-3	2.34
% Hiring entrepreneurs	57.4-64.6	65
% Average worked hours	33	33.4
Income distribution		
Income Gini	0.55	0.56
Entr. income Gini	0.66	0.62
Worker earnings Gini	0.51	0.51
99-100% income	17.2	21.2
95-99% income	16.6	18.9
% entr. in top 1%	40	35.3
Wealth distribution		
Wealth Gini	0.85	0.84
99-100% wealth	34.1	34.5
95-99% wealth	26.8	28.7
% People at zero wealth	7-13	13.8
Ratio of median net worth entr. to workers	5.3-6.5	5.2

Table 4 summarizes the key macroeconomic aggregates in the benchmark economy. In the table, the labor tax rate represents the tax burden that workers face in percentage terms. The low interest rate corresponds to the federal funds rate during 2011-2016.

[10] The percentage of entrepreneurs at the top 1% of income is taken from Malm and Sanandaji (2015), Table 7.

Table 4: Macroeconomic Aggregates

Variable	Value
Capital	289.5%
Government debt	78.2%
Consumption	79.2%
Investment	17.4%
Government consumption	3.5%
Average hours worked	33%
Interest rate	0.27%
Tax revenues	
- Consumption tax	4.0%
- Labor tax	8.9%
- Proportional capital tax	7.9%
Pension system	
- Total pension payment	11.8%

Tables 5 and 6 summarize the model-generated income and wealth distributions together with their counterparts in the data.[11] The standard life-cycle models often fail to generate income and wealth distributions correctly at the upper end.[12] Our model with workers and entrepreneurs is able to generate a realistic wealth and income distribution.

Table 5: Income Distribution in the Benchmark Economy

	Share of income (in %)								
	Income quintiles					Top			
	0-20%	20-40%	40-60%	60-80%	80-100%	90-95%	95-99%	99-100%	Gini
Data	3	6.5	10.9	18.1	61.4	10.7	16.6	17.2	0.58
Model	4.1	7.7	11.5	16.9	59.8	8.5	18.9	22.2	0.56

[11]Both income and wealth distribution data are taken from Khun and Rios-Rull (2016).

[12]Guner et al. (2016) introduce superstar individuals who are extremely productive but have a small share in the population. This leads to a labor income distribution that is in line with the data. Yet the model does not generate the wealth distribution well. Guner et al. (2016) report that the top 1%, 5%, and 10% own 15.2%, 35.1%, and 49.1% of the total wealth respectively, which is less than what we observe in data. According to SCF (2010), the top 1%, 5%, and 10% own 34.1%, 60.9%, and 74.4% of the total wealth. Kindermann and Krueger (2017) follow Castaneda et al. (2003), and their model generates earnings and wealth distributions quite realistically.

Table 6: Wealth Distribution in the Benchmark Economy

	Share of wealth (in %)								
	Wealth quintiles					Top			
	0-20%	20-40%	40-60%	60-80%	80-100%	90-95%	95-99%	99-100%	Gini
Data	-0.7	0.7	3.3	9.9	86.7	13.5	26.8	34.1	0.85
Model	0.2	0.8	3.8	7.9	87.2	13.1	28.7	34.5	0.84

Table 7 shows the distribution of income taxes paid in the data and the model generated distribution.[13] The distribution of tax payments is more concentrated than the income distribution but is less concentrated than the wealth distribution. In the data, first and second income quantiles are responsible for 2.5% of income tax payments. In the model this equal to 4.6%. Also in the data, fifth income quantile is responsible for 74.6% of income tax payments. The corresponding value in the model is 77.5%. The concentration in income tax payments is the natural consequence of the concentration in income and wealth distribution.

Table 7: Share of Tax Payments in the Benchmark Economy

			Share of tax (in %)		
	Income quintiles				
	0-20%	20-40%	40-60%	60-80%	80-100%
Data	0.3	2.2	6.9	15.9	74.6
Model	1.2	3.4	6.6	11.4	77.5

Overall our model matches income and wealth distributions quite well.

5 Experiments

In this section, we present results from two tax experiments. In the first experiment, we examine the impact of changes in the overall progressivity of taxes on government revenues and welfare. For both of these experiments, we search for the revenue and welfare-maximizing tax rates. In searching for the revenue-maximizing tax rate, we keep the value of all other tax parameters constant except for the tax progressivity parameter (in the first experiment) and the top marginal tax rate (in the second experiment). To satisfy the government budget condition, we vary the ratio of government expenditures to GDP, g as in Guner et al. (2016).

[13]The share of tax payments are taken from Guner et al. (2016), which is based on Internal Revenue Service (IRS) data.

In searching for the welfare-maximizing tax rate, we keep the government revenues constant and maximize the ex-ante expected utility of the agent under the two experiments.[14] We follow Heer and Trede (2003) and compute the consumption equivalent variation (Δ_c) as:

$$\Delta_c = \left[\frac{W(\Omega') - W(\Omega)}{\int_{\mathscr{X}} E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{c_t^{1-\sigma}}{1-\sigma} \right] d\Psi(\Omega)} + 1 \right]^{\frac{1}{1-\sigma}} - 1 \tag{34}$$

where $W(\Omega')$ and $W(\Omega)$ are the value functions after and before (benchmark) the policy changes. Here, $\mathscr{X}$ refers to the state space; Ω refers to the initial steady state, and Ω' refers to the final steady state.[15]

If the utility function is separable in consumption and leisure, Heer and Trede (2003) propose the following CEV measure:[16]

$$W(\Omega') = \int_{\mathscr{X}} E_0 \left[\sum_{t=0}^{\infty} \beta^t \frac{((1+\Delta_c)c_t)^{1-\sigma_1}}{1-\sigma_1} + \chi \frac{(1-l_t)^{1-\sigma_2}}{1-\sigma_2} \right] d\Psi(\Omega). \tag{35}$$

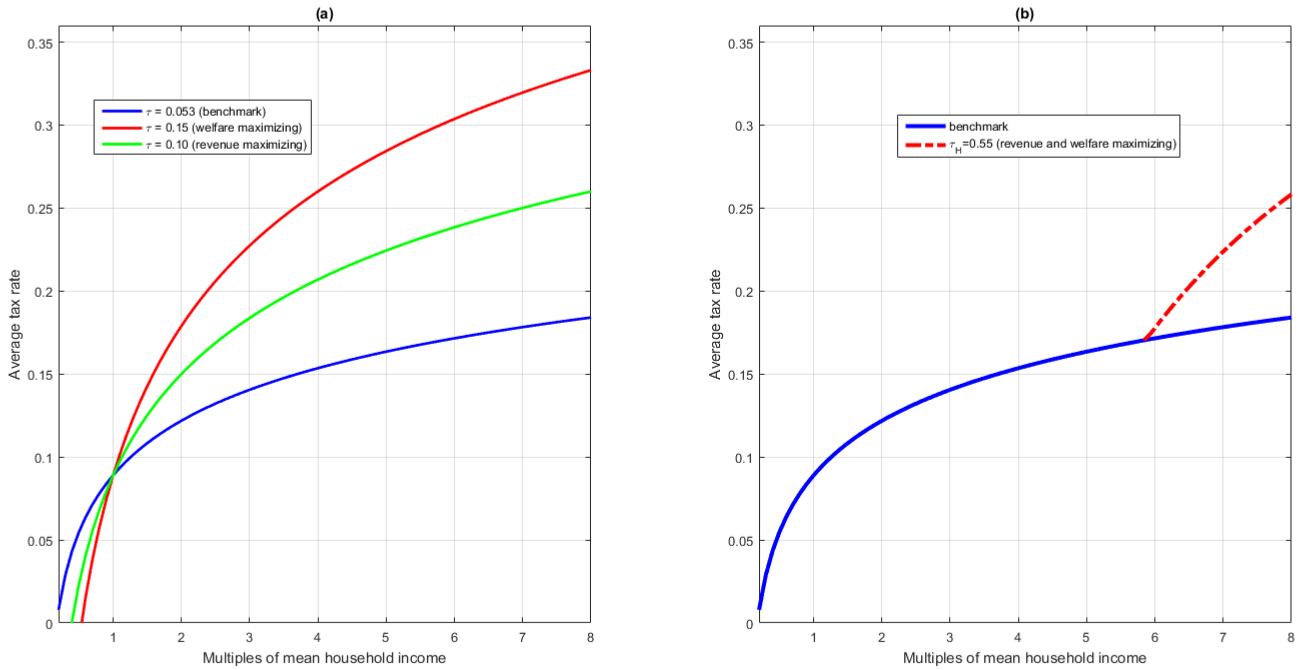
Figure 1 presents the optimal tax rates generated by our experiments under these alternative tax experiments that are discussed in detail in the rest of this section. Panel (a) in Figure 1 displays the average federal income tax rates implied by the first experiment where τ, the overall progressivity of taxes is altered. Panel (b) displays the average federal income tax rates implied by the second experiment where the only tool the government has is the marginal tax rate (τ_H) that the richest 1% of the households face.

[14] We keep the revenues constant by varying the proportional income tax rate that captures all other (state and local) in the economy (τ^{bal}). This is similar to Kindermann and Krueger (2017) who use a lower tax bracket, τ_l, as an adjustment parameter.

[15] Note that if there is no labor-leisure choice, the welfare measure in Heer and Trede (2003) becomes equivalent to the standard welfare measure $\Delta_c = \left(\frac{W(\Omega')}{W(\Omega)}\right)^{\frac{1}{1-\sigma}} - 1$.

[16] In the non-separable Cobb-Douglas specification, there is no need for the generalized formula as the labor in the numerator and the denominator get canceled.

Figure 1: Tax Rates



5.1 Revenue-Maximization

In this section, we present the impact of the two experiments on government revenues. In the first experiment, we fix the level parameter, λ, of the tax function at its benchmark value and search for the revenue-maximizing progressivity of taxes by varying the parameter, τ. In the second experiment, we calculate the revenue-maximizing marginal tax rate that applies to the top 1% of income.

Revenue-Maximizing Progressivity of Taxes

Table 8 displays the implications of the changes in the tax progressivity parameter, τ, on a number of economic outcomes. All variables, except for the interest rate, are normalized to 100 at the benchmark level of tax progressivity ($\tau = 0.053$). Thus, $\tau > 0.05$ displays the properties of economies where progressivity of taxes is higher than the benchmark and $\tau < 0.05$ summarizes the findings where progressivity is reduced.

Table 8: Changes in Progressivity-Revenue Maximizing

Progressivity	τ=0.035	τ=0.05	τ=0.07	τ=0.09	τ=0.10	τ=0.12	τ=0.15
Output	104.4	100.3	99.0	94.9	94.0	91.8	88.4
Labor supply	104.8	100.0	99.9	99.0	98.9	98.4	98.0
Capital	109.6	101.3	97.3	86.3	84.9	80.9	74.7
Revenues							
Federal income tax	96.0	99.0	102.7	105.27	**105.33**	104.0	97.7
State and local taxes	102.9	100.1	98.2	96.9	96.2	94.6	92.0
Corporate income tax	23.0	80.4	196.6	275.8	296.3	350.3	415.9
All taxes	98.9	99.5	101.0	**102.0**	101.8	100.5	96.2
Additional targets							
Interest rate	0.06	0.22	0.58	0.87	0.95	1.18	1.52
Worker avg. hours worked	104.8	100	99.4	99	98.9	98.4	98.1
Entr. avg. hours worked	100.7	100	95.2	94	91.5	87.7	86.2
Labor supply in corp sector	106	100.3	97.8	96.7	98.2	100.1	102.4
Labor supply in entr. sector	101.5	99.7	100.4	100.6	99.6	98.1	95
Capital in corp sector	111.9	101.5	91.1	84.5	84.3	81.9	78.2
Capital in entr. sector	107.1	100.7	93.7	88.2	85.5	79.9	71.2
Δ%entr. in overall economy	97.7	100	100.2	101.5	101.6	100.1	101.8

We find that revenues from the federal income tax schedule is maximized when $\tau = 0.10$ and tax revenues from all sources are maximized when $\tau = 0.09$. Both values are much larger than the benchmark value of 0.053. When $\tau = 0.10$, the federal income tax revenues increase by 5.33% and tax collected from all sources increase by 1.8% relative to the benchmark. The significant rise in marginal federal income tax rates in comparison to the average tax rates leads to standard disincentives in labor supply and saving decisions.[17] At this revenue-maximizing rate, capital, labor supply, and output decrease by 15.1%, 1.1%, and 6%, respectively. Local and state taxes and corporate income taxes are proportional to output and capital and hence, reductions in capital stock and output affect them negatively.

Higher tax progressivity reduces the capital stocks in both corporate and entrepreneurial sectors by 15.7% and 14.5%, respectively. The decrease in average hours worked is more pronounced in the entrepreneur sector, 8.5% compared 1.1% in the corporate sector. The more progressive federal income tax leads to a moderate increase (1.6%) in the population share of entrepreneurs through its effect on the interest rate. In the benchmark case, the interest rate is equal to 0.27%. It increases to 0.95% when $\tau = 0.10$ due to the decrease in the capital

[17] See Table 10 for a summary of the average and marginal tax rates at $\tau = 0.10$.

stock. When the interest rate is higher, workers who have high abilities as entrepreneurs can earn higher returns from their savings. Hence, they can become entrepreneurs more quickly. Although the capital stock decreases, the corporate income tax revenue increases substantially due to the large increase in the interest rate.

Our finding that a more progressive federal income tax schedule, relative to the benchmark (that is calibrated to the U.S) maximizes revenue is in line with the findings of Guner et al. (2016). They show that the federal income tax revenue is maximized when $\tau = 0.13$ and revenues from all income sources are maximized when $\tau = 0.10$. Similar to our results, they observe significant reductions in economic aggregates as well.

Revenue-Maximizing Top Tax Rate

In the second experiment, we vary the marginal tax rate for the richest 1% of the population only. We normalize the values of all the economic aggregates, except for the interest rate, in Table 9 to 100 at the benchmark value of $\tau_H = 0.229$ and report changes from this benchmark. Our findings indicate that both federal income tax revenue and overall tax revenue are maximized when the marginal income tax rate for the top 1% is 55%. At this rate, Federal income tax revenue increases by 16.3% and the tax revenue from all sources increases by 5.4%. Note that tax revenues from federal income and all sources increase substantially more in this case. In other words, a targeted increase in the marginal tax rate of 1% generates much more tax revenue than the experiment that affects larger income groups. Imposing a 55% marginal tax rate reduces capital stock, labor supply, and output by 8.2%, 0.7%, and 3.9%, respectively. These reductions are substantially lower than what we observe in the earlier case where overall progressivity was altered. Lower reductions lead to relatively higher tax revenues in this case.[18]

Higher tax progressivity, reduces the capital stocks in corporate and entrepreneurial sectors by 5.2% and 11.2% respectively. Average hours worked decline by 0.7% for the workers and 2.2% for the entrepreneurs.

[18] This result is slightly different than that of Guner et al. (2016) who claim that there is not much revenue available from shifting the tax burden towards top earners. Our results show that more revenue can be extracted imposing a higher marginal tax to 1% instead of increasing progressivity in a way that affects relatively larger income group. The revenue maximizing marginal income tax rate we find is closer to the 49% found by Badel and Huggett (2015). In contrast, Kindermann and Krueger (2017) report revenue maximizing tax rates for the top 1% to be 86% in the short run and 98% in the long run. In calculating this rate, Kindermann and Krueger (2017) maximize revenues focusing on labor earnings only. We discuss this case in detail in Section 6.

Table 9: Changes in Tax for Top 1% - Revenue Maximizing

Marginal tax for top 1%	$\tau_H = 0.2$	τ_H=0.4	τ_H=0.55	τ_H=0.6	τ_H=0.8
Output	101.1	98.2	96.1	92.4	88.7
Labor supply	100.2	99.7	99.3	98.7	97.7
Capital	104.6	95.8	91.8	87.9	84.4
Revenues					
Federal income tax	88.7	107.3	**116.3**	109.8	95.7
State and local taxes	86	86.4	86.5	86.9	86.6
Corporate income tax	49.6	141.1	195.8	248.8	314.9
All taxes	90.6	100.7	**105.4**	101.5	93.3
Additional targets					
Interest rate	0.13	0.40	0.58	0.63	1.02
Worker avg. hours worked	100.2	99.7	99.3	98.7	97.7
Entr. avg. hours worked	100.5	98.8	97.8	99.6	98.6
Labor supply in corp sector	102.4	98.6	101.7	114	125.9
Labor supply in entr. sector	99	99.3	97	88	79.6
Capital in corp sector	106	95.8	94.8	101.7	106.7
Capital in entr. sector	103.1	95.8	88.8	73.9	61.6
Δ%entr. in overall economy	97.9	100.1	100.1	101.6	101.7

Table 10 summarizes the average and marginal income tax rates and share of tax payments for various income quantiles for three economies: 1) the benchmark, 2) the economy where revenues from the federal income tax schedule is maximized ($\tau = 0.10$), and 3) the economy where the revenue-maximizing marginal income tax for the top 1% is equal to 55%. In the benchmark economy, average tax rates are 12.3%, 15%, and 18.6% and marginal tax rates are 16.9%, 19.5%, and 22.9% for the richest 10%, 5%, and 1% of households, respectively. In the second economy, average tax rates are 14.9%, 19.7%, and 25.6% and marginal tax rates are 23.4%, 27.8%, and 33.1%, respectively. In the third economy, average tax rates are 14.1%, 15.8%, and 28.4% and marginal tax rates are 20.3%, 22.7%, and 55%, respectively. In both tax experiments, summarized by the second and third cases, average and marginal tax rates increases substantially, which explain the large decreases found in economic aggregates. When the overall progressivity is altered, ($\tau = .10$), tax rates faced by all groups increase somewhat more uniformly. As discussed earlier, this results in more entrepreneurs being affected by the changes in taxes and leads to larger decreases in economic aggregates relative to the case where the tax rate of the richest 1% of the population is targeted.

Table 10: Average and Marginal Tax Rates and Share of Tax Payments

Percentiles of income	Benchmark	τ=0.10	τ_H=0.55
	Average tax rate		
Top 10%	12.3	14.9	14.1
Top 5%	15.0	19.7	15.8
Top 1%	18.6	25.6	28.4
	Marginal tax rate		
Top 10%	16.9	23.4	20.3
Top 5%	19.5	27.8	22.7
Top 1%	22.9	33.1	55.0
	Share of tax payments		
Income quintiles			
0-20%	1.2	-1.0	1.1
20-40%	3.4	0.4	3.1
40-60%	6.6	3.5	6.0
60-80%	11.4	8.0	10.3
80-100%	77.5	89.2	79.5

The contribution to income tax payments by households at different income levels shift from lower income quintiles to the higher ones at a substantial degree in the second experiment. For instance, the poorest 20% of households now make negative contributions while the richest 20% of household's share of tax payments increase from 77.5% to 89.2%. This change, on the other hand, is somewhat limited in the third case. The poorest 20% of household's contribution decreases from 1.2% to 1.1% and the richest 20% of household's contribution increases from 77.5% to 79.5%.

Table 11 displays the changes in wealth and income distributions for various income quantiles for these three economies. The wealth share of the top 1% decreases the most in the third economy. The wealth Gini falls in both cases.

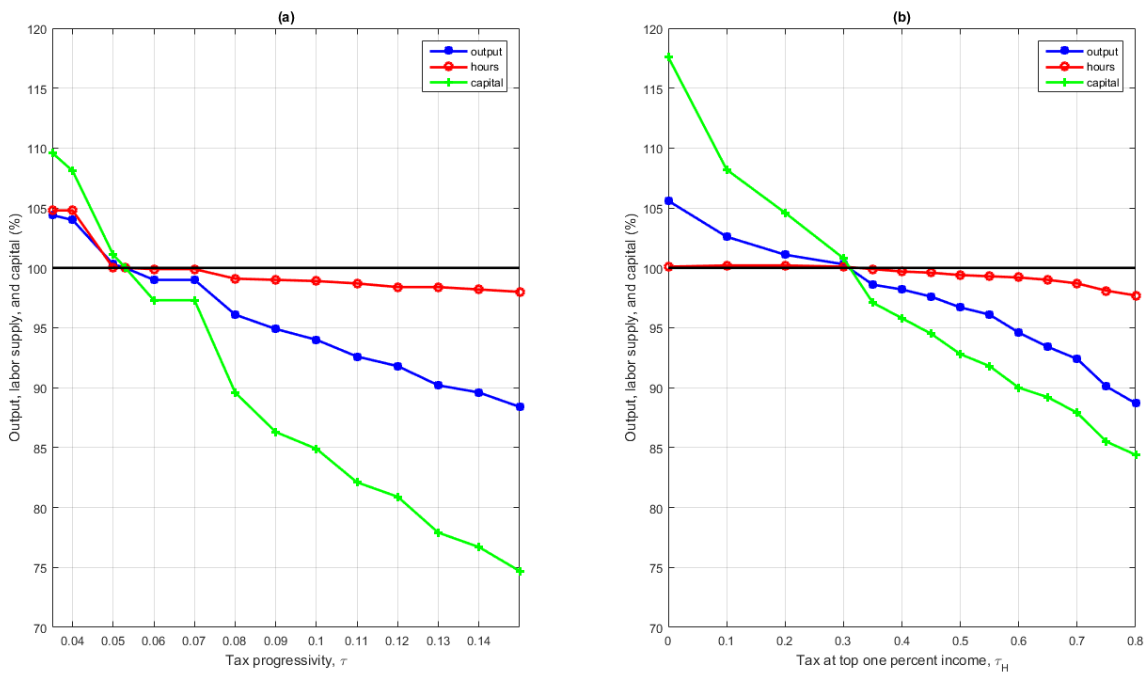
Table 11: Changes in Wealth and Income Distribution - Revenue Maximizing

	Benchmark	τ=0.10	τ_H=0.55
	Wealth distribution		
Wealth quintiles			
0-20%	0.2	0.2	0.2
20-40%	0.8	1.4	1.0
40-60%	3.8	4.7	4.4
60-80%	7.9	9.6	9.4
80-100%	87.2	84.1	85.1
Top			
10%	76.3	71.9	72.8
5%	63.2	58.6	58.6
1%	34.5	31.0	28.8
Wealth Gini	0.84	0.81	0.82
	Income distribution (all)		
Income quintiles			
0-20%	4.1	4.3	4.3
20-40%	7.7	7.7	7.8
40-60%	11.5	11.6	11.6
60-80%	16.9	16.8	17.0
80-100%	59.8	59.6	59.3
Top			
10%	49.7	49.2	48.9
5%	41.2	40.7	40.1
1%	22.2	21.3	19.8
Income Gini	0.56	0.55	0.55

Comparison of the Two Tax Experiments

Figure 2 summarizes the impact of the two experiments on economic outcomes. As discussed earlier, increases in the overall progressivity of taxes (panel a) imply a larger decline, especially in capital and output compared to the increase in the tax rate that targets the richest 1% of the population (panel b). This is partly due to the fact that an increase in the overall progressivity of taxes affects a larger number of entrepreneurs compared to an increase in the tax rate facing the richest 1% of the individuals. In the benchmark economy, the percent of entrepreneurs in the top 1% of income is 35%. Thus, an increase in the tax rate that the richest 1% face does not impact all the entrepreneurs in the economy. On the contrary, an increase in the overall progressivity of taxes impacts all the entrepreneurs. Therefore, the capital stock and the hours worked by the entrepreneurs react more negatively to an increase in the overall progressivity of taxes. Thus, revenue maximizing through imposing a higher marginal tax rate to the richest 1% of households creates fewer distortions on economic aggregates, resulting in relatively more revenues.

Figure 2: Changes in Output, Labor Supply, and Capital



5.2 Welfare Maximization

In this section, we search for the welfare maximizing progressivity of taxes by varying the parameter, τ in the first experiment. In the second experiment, we calculate the welfare-maximizing marginal tax rate that applies to the top 1 % of income.

Welfare-Maximizing Progressivity of Taxes

Table 12 summarizes the changes in economic aggregates and welfare as the progressivity parameter τ is changed. All variables, except for the interest rate, are normalized to 100 at the benchmark level of tax progressivity ($\tau = 0.053$). Welfare increases with increases in τ and peaks at $\tau = 0.15$. This value is higher than the value found in the first experiment focusing on revenue maximization ($\tau = .10$). At this level of progressivity, capital, labor supply, and output decrease by 25.9%, 8.4%, and 12.9%, respectively. Both capital and labor in the entrepreneurial sector decline more relative to the corporate sector. Compared to the benchmark, this level of progressivity leads to a decrease in the federal income tax revenue by 3.2% and the total tax revenue by 5.8%.

Table 12: Changes in Progressivity - Welfare Maximizing

Progressivity	τ=0.035	τ=0.06	τ=0.09	τ=0.12	τ=0.15	τ=0.18	τ=0.21
Output	104.3	99.2	95.1	92.1	87.1	80.3	75.1
Labor supply	104.8	99.9	99.0	98.4	91.6	90.8	90.3
Capital	109.0	97.8	87.5	81.4	74.1	64.0	56.3
Revenues							
Federal income tax	95.9	101.5	105.3	104.6	96.8	74.1	53.1
State and local taxes	113.5	94.6	77.9	73.9	87.2	129.7	168.9
Corporate income tax	34.4	134.9	249.4	336.2	385.8	501.3	593.4
All taxes	101.3	99.6	97.6	96.1	94.2	90.9	87.9
Local tax rate, τ_{bal}	5.5	4.8	4.0	3.9	4.8	7.6	10.4
Average CEV							
CEV (All)	-1.06	0.38	2.02	3.48	**4.25**	2.39	1.03
CEV (Work)	-1.07	0.37	1.99	3.45	4.28	2.38	1.01
CEV (Entr.)	-0.98	0.51	2.46	3.79	3.93	2.59	1.19
Additional targets							
Interest rate	0.09	0.38	0.78	1.13	1.43	2.14	2.89
Worker avg. hours worked	104.8	99.9	99.0	98.4	91.6	90.8	90.3
Entr. avg. hours worked	102.1	98.5	93.1	88.5	86.8	77.9	71.0
Labor supply in corp sector	106.2	100.7	96.0	99.1	99.2	110.1	120.2
Labor supply in entr. sector	101.4	99.8	100.3	98.3	93.2	84.1	77.4
Capital in corp sector	111.2	98.1	85.7	81.9	77.2	74.7	70.5
Capital in entr. sector	106.6	97.5	89.4	81	70.8	53.1	41.7
Δ%entr. in overall economy	97.7	100.1	101.5	100.1	101.7	102.2	102.3

Table 13, explores the forces behind the welfare gains despite the fact that there are large drops in economic aggregates. In our model, there are four distinct groups: young workers (YW), young entrepreneurs (YE), old workers (OW), and old entrepreneurs (OE). YW make up the largest share of the population, 73%, followed by 19.5% OW, 6.7% YE, and 0.5% OE.[19] Panel A of Table 13 documents average consumption and hours worked for the whole economy as well as for different income groups. All the information provided

[19]These shares do not vary much across experiments.

is relative to the benchmark, which is normalized to 100. Panel B of Table 13 documents variances of consumption and average hours worked, again relative to the benchmark. In the overall economy, average consumption of the young entrepreneurs decrease by 28.5% while young workers experience a more moderate drop of 6.5%. The economy-wide changes, however, mask the rich heterogeneity in the responses of different groups. For example, young entrepreneurs who are at the top 1% of incomes experience a 31.5% decline in their consumption while young workers in the same income group experience only a 4.1% decline in their consumption. Young workers in the lowest 33% of income experience a 20.8% decline in consumption while the middle income YW experience a 45.7% increase in consumption. Since the middle 33% of YW make up 36.6% of the population, an increase in their consumption contributes to the welfare improvement we observe. Our results indicate a decline in hours worked for most of these groups and in particular for the poorest YW, which contributes to the decline in their average consumption. For all groups, the variance of consumption declines. These large declines in consumption variances, substantial increases in leisure time, and the increase in the average consumption of a large group of workers all contribute to the overall welfare gains we observe. Although young and old entrepreneurs are affected quite negatively, their small share in the population reduces their impact on the overall welfare results.

Table 13: Consumption and Hours - Welfare Maximizing Progressivity

Panel A	Average consumption				Average hours worked		
Experiment $\tau = 0.15$	YW	YE	OW	OE	YW	YE	OE
whole economy	93.5	71.5	95.4	57.4	87.2	87.2	72.7
top 1%	95.9	68.5	N/A	55.7	100.0	100.0	100.0
bottom 99%	95.2	98.0	95.1	92.3	85.8	85.8	71.0
67-100%	99.4	70.0	95.4	57.1	96.2	85.5	73.0
34-66%	145.7	95.6	N/A	112.2	93.2	101.5	100.0
0-33%	79.3	N/A	92.9	N/A	89.8	N/A	N/A
Panel B	Variance consumption				Variance hours worked		
	YW	YE	OW	OE	YW	YE	OE
whole economy	54.2	19.0	49.0	18.7	58.3	94.4	65.8
bottom 99%	41.1	81.0	30.0	65.4	57.7	96.3	66.6

Welfare-Maximizing Top Tax Rate

Results of the second experiment, where we vary the tax rate at the top, are presented in Table 14. The welfare maximizing marginal tax rate for the top 1% is found to be 55%, the same rate as in the revenue-maximizing tax rate for the top 1%. At this tax rate, capital, labor supply, and output decrease by 7%, 0.8%, and 3.8%, respectively. Targeting the top 1% generates a moderate welfare gain (CEV increases by 0.72%) compared to the experiment where the overall progressivity is increased (CEV increases by 4.25%). At the welfare-maximizing rate, workers and entrepreneurs' average hours worked decrease slightly, by 0.8% and 2.4%, respectively. Capital stock in the corporate sector decreases by 4.3%, and capital stock in the entrepreneur sector decreases by 9.7%. As discussed earlier, changing the tax rate for the richest 1% creates smaller distortions than changing the overall progressivity of taxes. This fact also contributes to the smaller welfare gains found in this case.

Table 14: Changes in Tax for Top 1% - Welfare Maximizing

Marginal tax for top 1%	$\tau_H = 0$	τ_H=0.2	τ_H=0.4	τ_H=0.55	τ_H=0.7	τ_H=0.8
Output	104.4	100.7	98.5	96.2	92.7	88.7
Labor supply	105.7	100.4	99.6	99.2	98.9	97.7
Capital	108.9	102.7	96.6	93	89	83.7
Revenues						
Federal income tax	62.9	88.5	107.6	114.9	110.1	95.9
State and local taxes	189	127.9	80.5	61.6	69	96.3
Corporate income tax	85	92	127.4	155.6	236.8	334.3
All tax	101.1	100.3	99.5	98.8	97.5	95.7
Local tax rate, τ_{bal}	11	7.5	4.7	3.5	4	5.6
Average CEV						
All	-5.97	-2.48	-0.04	**0.72**	-0.81	-3.79
Workers	-5.98	-2.48	-0.07	0.66	-0.97	-4.07
Entr.	-5.89	-2.52	0.35	1.58	1.29	-0.18
Additional targets						
Worker avg. hours worked	105.7	100.4	99.6	99.2	99	97.7
Entr. avg. hours worked	104.8	103	98.8	97.6	97.5	98.4
Labor supply in corp sector	109.4	103.3	98.2	100.4	104	125.8
Labor supply in entr. sector	101.6	99.1	99.5	96.4	93.9	80.2
Capital in corp sector	111.2	104.2	96.5	95.7	97.8	105.6
Capital in entr. sector	106.6	101.1	96.8	90.3	85.5	61.3
Δ% entr. in overall economy	97.3	99.8	100.1	100	100.1	101.7

Table 15 presents the changes in the level and the variance of consumption and hours worked for different types of individuals in the economy at the welfare maximizing level of τ_H relative to the benchmark. Average consumption of the entrepreneurs at the top 1% falls by 38.6%. Average consumption by the richest 1% young workers' decrease by 23.8%. Hours worked declines for most groups, except for those on the top 1% of incomes. Variance of consumption declines for all except the old workers.

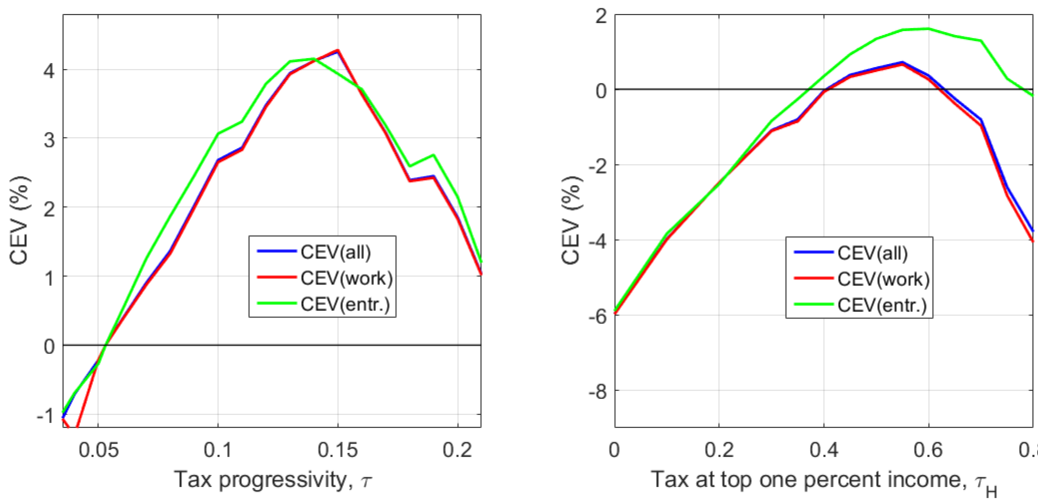
Table 15: Consumption and Hours - Welfare-Maximizing Tax at the Top

Panel A	Average consumption				Average hours worked		
Experiment $\tau_H = 0.55$	YW	YE	OW	OE	YW	YE	OE
whole economy	100.2	82.2	100.1	74.3	99.2	97.8	89.2
top 1%	76.2	61.4	N/A	54.6	115.2	100.0	100.0
bottom 99%	102.9	92.4	100.1	99.6	99.5	98.7	90.1
67-100%	109.8	80.2	100.0	73.6	97.2	99.1	90.8
34-66%	139.4	88.1	N/A	N/A	98.4	103.8	N/A
0-33%	89.2	N/A	102.7	N/A	99.0	N/A	N/A
Panel B	Variance consumption				Variance hours worked		
	YW	YE	OW	OE	YW	YE	OE
whole economy	75.3	40.7	67.1	40.4	99.3	108.1	89.1
bottom 99%	79.2	66.9	67.1	112.0	99.7	108.7	91.2

Comparison of the Two Tax Experiments

Our results indicate that the optimal tax rate that targets the richest 1% of the population generates a moderate welfare gain (CEV increases by 0.72%) compared to the experiment where the overall progressivity is increased (CEV increases by 4.25%). Figure 3 summarizes the welfare results from these experiments. Panel A displays the welfare gain/loss as we increase the progressivity of taxes by varying τ. Panel B, summarizes changes in welfare as we increase the tax rate that applies to the top 1% of income levels.

Figure 3: Welfare Maximizing



To delve deeper into the reasons behind the welfare results, in Table 16, we display the income and wealth distributions in the three economies considered: the benchmark economy, the economy where welfare is maximized by changing the overall progressivity of taxes ($\tau = 0.15$), and the economy where the welfare-maximizing marginal income tax rate for the top 1% is found to be equal to 55%. While the income distribution is not very different across these three economies, the wealth distribution displays important differences. While the Wealth Gini in the benchmark economy is 0.84, the second economy where the overall progressivity of taxes is altered produces a Wealth Gini of 0.79. Increasing only the tax rate of the richest 1% results in a small change in the Wealth Gini (0.82) relative to the benchmark. While the wealth share of the top 1% does not change much across two experiments, the wealth share of the top 10% decreases and the wealth share of most of the lower quantiles increases in the progressivity maximization case. Overall, we find that an improvement of the overall progressivity of taxes generates a much larger welfare gain than imposing a very high marginal tax at the top 1%.

Table 16: Changes in Wealth and Income Distribution - Welfare Maximizing

	Benchmark	τ=0.15	τ_H=0.55
	Wealth distribution		
Wealth quintiles			
0-20%	0.2	0.1	0.2
20-40%	0.8	1.6	1.0
40-60%	3.8	5.7	4.2
60-80%	7.9	11.2	9.2
80-100%	87.2	81.4	85.4
Top			
10%	76.3	68.2	73.2
5%	63.2	54.8	58.8
1%	34.5	28.1	28.6
Wealth Gini	0.84	0.79	0.82
	Income distribution (all)		
Income quintiles			
0-20%	4.1	4.0	4.2
20-40%	7.7	7.4	7.9
40-60%	11.5	11.8	11.7
60-80%	16.9	17.4	17.2
80-100%	59.8	59.4	59.1
Top			
10%	49.7	48.7	48.7
5%	41.2	39.8	39.9
1%	22.2	19.7	19.4
Income Gini	0.56	0.55	0.55

Information on the share of tax payments by different income groups summarized in Table 17 provides further evidence on how lower income groups benefit more under a change in the progressivity of taxes. In this case, the share of tax payments by lower income groups is negative, indicating they are receiving transfers.[20]

[20] Total taxes paid are calculated net of transfers. Consequently, payment made by richer households may exceed 100%.

Table 17: Share of Tax Payments and Tax Rates - Welfare Maximizing

Percentiles of income	Benchmark	τ=0.15	τ_H=0.55
	Average tax rate		
Top 10%	12.3	17.2	14.1
Top 5%	15.0	24.2	15.7
Top 1%	18.6	32.0	28
	Marginal tax rate		
Top 10%	16.9	29.6	20.1
Top 5%	19.5	35.6	22.3
Top 1%	22.9	42.2	55.0
	Share of tax payments		
Income quintiles			
0-20%	1.2	-4.2	0.9
20-40%	3.4	-3.2	2.7
40-60%	6.6	0.1	5.5
60-80%	11.4	5.2	9.8
80-100%	77.5	102.2	81.0

6 Sensitivity Analysis

In this section, we examine the sensitivity of our results to certain parameters used and modeling choices that are made. In doing so, we also investigate the potential reasons for the different findings in Kindermann and Krueger (2017), Guner et al. (2016), and Badel and Huggett (2015). In general, our findings are closer to what is reported in Guner et al. (2016) and Badel and Huggett (2015). Badel and Huggett (2015) assess the consequences of increasing the marginal tax rate on top earners using a human capital model. They calculate a revenue-maximizing top tax rate of 49%. Our revenue-maximizing tax rate of 55% is comparable to their findings. Kindermann and Krueger (2017), on the other hand, find the revenue-maximizing top marginal tax rate to be 98% in the long-run. Our study distinguishes from Kindermann and Krueger (2017) in two important ways. First, our model as in Guner et al. (2016), uses Benabou's tax function to generate a realistic share of tax payments by income quintiles. In this process, we define taxable income as both labor and capital income. In Kindermann and Krueger (2017), taxes apply to labor income only. Second, in our model, the entrepreneurship sector is the main driving factor generating the right income and wealth distribution. The role of the lucky high productivity state is somewhat limited compared to that of Kindermann and Krueger (2017).

In order to investigate the sensitivity of the results to the tax base that is used, we repeat our exercise by taxing labor income only. We re-calibrate the model to the same target moments and search for the welfare-maximizing tax rate for the richest 1% of the population.

Table 18 shows the results of this experiment where we find the welfare-maximizing top marginal tax rate to be 80%.[21] This rate is much higher than the welfare-maximizing 55% marginal tax rate found in Section 5.2 where the tax base was total earnings (labor and capital income). It is also closer to the results reported in Kindermann and Krueger (2017). [22] Compared to the results in Table 14, capital stock remains fairly flat in this experiment as τ_H increases. In Table 14, where total earnings is taxed, capital stock declines by 16.5% at $\tau_H = 0.8$. In Table 18 where the tax base is composed of labor earnings only, capital stock declines by 3.5% at $\tau_H = 0.8$.

Table 18: Change in Tax at the Top 1% Earnings - Welfare Maximizing

Marginal tax for top 1%	τ_H =0	τ_H =0.1	τ_H =0.2	τ_H =0.4	τ_H =0.6	τ_H =0.8	τ_H =0.9	τ_H =1
Output	102.6	102.8	102.4	99.3	97.3	97.1	95.8	95.1
Labor supply	105.1	105.1	104.8	99.9	99.7	99.7	99.5	99.2
Capital	103.8	103.5	103.2	98.7	96.2	96.6	95.1	94.2
Interest rate	108.5	108.4	108.8	103.0	107.4	104.3	113.5	117.5
Tax revenues								
Federal income tax	79.7	90.7	99.1	110.5	115.5	123.4	122.2	120.9
State and local taxes	129.9	112.1	99.2	83.5	76.6	65.0	67.0	69.4
Corporate income tax	102.8	103.2	103.2	100.1	98.6	97.1	96.2	95.3
All taxes	100.1	100.1	100.1	99.9	99.7	99.7	99.5	99.3
Local tax rate, τ_{bal}	7.4	6.3	5.6	4.7	4.3	3.7	3.8	3.9
Average CEV								
CEV (ALL)	-2.7	-1.6	-0.8	0.2	0.5	**1.3**	0.9	0.6
CEV(Work)	-2.7	-1.6	-0.9	0.2	0.4	1.3	0.8	0.6
CEV(Entr)	-2.3	-1.2	-0.4	0.4	0.8	1.5	1.5	1.4

In addition we find that our results are not very sensitive to small changes in the values of risk aversion and Frisch elasticity parameter. Risk aversion of 1.7 and a Frisch elasticity

[21] Entrepreneurial earnings is given by $(1-\gamma)\{\theta_t\left(k_t^{\gamma}(l_t+n_t)^{1-\gamma}\right)^{\nu}-\delta k_t-r_t(k_t-a_t)-w_t n_t\}$. Entrepreneurial capital income is given by $\gamma\{\theta_t\left(k_t^{\gamma}(l_t+n_t)^{1-\gamma}\right)^{\nu}-\delta k_t-r_t(k_t-a_t)-w_t n_t\}$.

[22] The revenue-maximizing top marginal income tax rate that applies to the top 1% of earnings is also higher (85%) when the tax base is composed of labor earnings only.

of 0.71 results in the same revenue maximizing value of the progressivity parameter and the marginal tax rate that applies to top 1% of income.

Before closing the session, we would like to clarify our choice of welfare analysis. In our paper, all welfare analysis done in steady-state settings ignoring transitional dynamics. There is no doubt that the accurate welfare analysis require analyzing transitional dynamics as well. Our model had many complexities that prevented us to do this. In a version of model without labor-leisure choice in corporate and entrepreneurial sector, we were able to conduct transitional dynamics but the model failed generating important moments that are essential to analyze top taxation. Hence, we opted out the transitional dynamics.

7 Conclusion

In this paper we study optimal income taxation in a model with an entrepreneurship decision that allows the model to generate realistic wealth and income distributions. We develop a simplified life-cycle model with stochastic aging that includes explicit modeling of entrepreneurial decisions, individual heterogeneity, and endogenous labor supply. Individual heterogeneity stems from differences in entrepreneurial abilities and uninsurable productivity shocks. We use a parametric tax function proposed by Benabou (2002) to model the relationship between income and income taxes paid at the federal level. We study revenue and welfare-maximizing changes in the overall progressivity of taxes versus changing the tax rate that applies to the richest 1% of individuals only.

The revenue and welfare-maximizing tax rates we find in this environment are higher than the existing tax rates in the U.S. Furthermore, we show that focusing on the tax rates facing the richest 1% of the population results in larger government revenues but lower welfare compared to increasing the overall progressivity of taxes. When only the tax rate at the top increases, fewer entrepreneurs are impacted. As a result, the distortions created are lower and revenues are higher. However, targeting the tax rate of the richest 1% does not generate big welfare gains. An increase in the overall progressivity of taxes, on the other hand, results in a lower Wealth Gini and higher welfare.

References

Ales, Laurence and Christopher Sleet, "Taxing top CEO incomes," *American Economic Review*, November 2016, *106* (11), 3331–66.

_ , Antonio Andres Bellofatto, and Jessie Jiaxu Wang, "Taxing atlas: executive compensation, firm size and their impact on optimal top income tax rates," *Review of Economic Dynamics*, October 2017, *26*, 62–90.

Alvaredo, Facundo, Anthony B. Atkinson, Thomas Piketty, and Emmanuel Saez, "The top 1 percent in international and historical perspective," *Journal of Economic Perspectives*, September 2013, *27* (3), 3–20.

Attanasio, Orazio, James Banks, Costas Meghir, and Guglielmo Weber, "Humps and bumps in lifetime consumption," *The Journal of Business and Economic Statistics*, 1999, *17(1)*, 22–35.

Badel, Alejandro and Mark Huggett, "Taxing Top Earners: A Human Capital Perspective," *Working Paper*, 2015.

Bassetto, Marco, Marco Cagetti, and Mariacristina De Nardi, "Credit crunches and credit allocation in a model of entrepreneurship," *Review of Economic Dynamics*, January 2015, *18* (1), 53–76.

Benabou, Rolland, "How Strong Are Bequest Motives? Evidence Based on Estimates of the Demand for Life Insurance and Annuities," *Econometrica*, 2002, *70(2)*, 481–517.

Bruggemann, Bettina, "Higher Taxes at the Top: The Role of Entrepreneurs," *Working Paper*, 2017.

Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin, "Entrepreneurship and financial frictions: a macro-development perspective," *University of Notre Dame Working Paper*, 2015.

Castaneda, Ana, Javier Diaz-Gimenez, and Jose-Victor Rios-Rull, "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 2003, *111*, 818–857.

Diamond, Peter and Emmanuel Saez, “The case for a progressive tax: from basic research to policy recommendations,” *Journal of Economic Perspectives*, December 2011, *25* (4), 165–90.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri, “Heterogeneity and persistence in returns to wealth,” *NBER Working Paper Series*, 2016.

Guner, Nezih, Martin Lopez-Daneri, and Gustavo Ventura, “Heterogeneity and Government Revenues: Higher Taxes at the Top?,” *Journal of Monetary Economics*, 2016, *80* (C), 69–85.

Heer, Burkhard and Mark Trede, “Efficiency and distribution effects of a revenue-neutral income tax reform,” *Journal of Macroeconomics*, March 2003, *25* (1), 87–107.

Kaplan, Greg, “Inequality and the life cycle,” *Quantitative Economics*, 2012, *3*, 471–525.

Khun, Moritz and Jose-Victor Rios-Rull, “2013 Update on the U.S. Earnings, Income, and Wealth Distributional Facts: A View from Macroeconomics,” *Federal Reserve Bank of Minneapolis Quarterly Review*, 2016, (April), 1–75.

Kindermann, Fabian and Dirk Krueger, “High Marginal Tax Rates on the Top 1Working Paper, 2017.

Kitao, Sagiri, “Entrepreneurship, Taxation, and Capital Investment,” *Review of Econoic Dynamics*, 2008, *11*, 44–69.

Kotlikoff, Laurence J., Kent Smetters, and Jan Walliser, “Privatizing social security in the United States-comparing the options,” *Review of Econoic Dynamics*, 1999, *2(3)*, 532–574.

Lucas, Robert E., “On the size distribution of business firms,” *Bell Journal of Economics*, 1978, *9* (2), 508–523.

Malm, Arvid and Tino Sanandaji, “The role of entrepreneurship in rising wealth and income inequality,” Working Paper Series in Economics and Institutions of Innovation 398, Royal Institute of Technology, CESIS - Centre of Excellence for Science and Innovation Studies Mar 2015.

Piketty, Thomas and Emmanuel Saez, "Income Inequality in the United States, 1913-1998," *The Quarterly Journal of Economics*, 2003, *118(1)*, 1–41.

Quadrini, Vincenzo, "Entrepreneurship, Saving, and Social Mobility," *Review of Economic Dynamics*, 2000, *3*, 1–40.

Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick, "Capitalists in the twenty-first century," *Working Paper*, 2017.

Stokey, Nancy L. and Sergio Rebelo, "Growth effects of flat-rate taxes," *Journal of Political Economy*, 1995, *103(3)*, 519–550.

Survey of Consumer Finances, "The Federal Reserve Board Survey of Consumer Finances," 2010.

Appendix A: Markov Chain for Labor Productivity

We approximate the idiosyncratic labor productivity process as explained in the main text. The process is assumed to be AR(1) and a Rouwenhorst's method is used to discretize it into a five-point Markov chain. Then it is augmented for another grid point to represent exceptionally high productivity workers. The grid points y for the idiosyncratic labor productivity, which is normalized to one, are:

$$\begin{bmatrix} 0.1612 & 0.3043 & 0.5744 & 1.0840 & 2.0459 & 11.4870 \end{bmatrix}.$$

The Markov matrix for idiosyncratic labor productivity is then

$$\begin{bmatrix} 0.9168 & 0.0787 & 0.0025 & 0.0000 & 0.0000 & 0.0020 \\ 0.0197 & 0.9180 & 0.0590 & 0.0013 & 0.0000 & 0.0020 \\ 0.0004 & 0.0393 & 0.9185 & 0.0393 & 0.0004 & 0.0020 \\ 0.0000 & 0.0013 & 0.0590 & 0.9180 & 0.0197 & 0.0020 \\ 0.0000 & 0.0000 & 0.0025 & 0.0787 & 0.9168 & 0.0020 \\ 0.0000 & 0.0000 & 0.0693 & 0.0000 & 0.0000 & 0.9307 \end{bmatrix}.$$

Appendix B: Computational Algorithm

The algorithm we use to solve the benchmark of the model is as follows.

1. We construct a grid for all state and control variables. The grid points for the asset level is chosen such that the grid is more refined early in the area where the value function is steeper. The grid grows more sparsely as the value function becomes flatter. The maximum value of assets is chosen such that in the invariant distribution there is no significant probability mass falling on the highest level. This level is much higher than those levels in a standard OLG model.

2. Given the initial guesses value for interest rate r, we solve for the policy functions using value function iteration.

3. We construct the transition matrix, Γ, for all possible choices of state variables. The matrix is big and contains a lot of zeroes, so it is constructed using a sparse matrix method. Given the initial guess of the invariant distribution, denoted Φ, we iterate on $\Phi' = \Gamma \Phi$. The invariant distribution satisfies $\Phi^* = \Gamma \Phi^*$.

4. We compute the total household savings and total capital in the entrepreneurial sector. Then we use the equilibrium clearing condition in capital market (see main text) to compute total capital in the corporate sector.

5. We compute hired labor in the entrepreneurial sector and total efficient labor. Then we compute total labor in the corporate sector from the labor market clearing condition.

6. We iterate on g to satisfy the government budget constraint. We keep r and w implied by total capital and labor in the corporate sector.

7. After we satisfy the government budget constraint, we update the equilibrium interest rate using the r implied above. We iterate over the equilibrium interest rate until the interest rate from the previous iteration and the current iteration is less than the convergence criteria.

The experiments are conducted as follows.

1. For experiments regarding progressivity and revenue maximizing, we change τ, which is the progressivity parameter in Benabou's tax function. Then we compute the federal income tax and total tax revenue. All other parameters remain fixed. We keep varying τ until we find the maximized level of tax revenues.

2. For experiments regarding tax at the top 1% and revenue maximizing, we first compute the income threshold from the benchmark. Then we compute the model by imposing the top tax rate, τ_H, only if individuals have income higher than the threshold level. All other parameters also remain fixed. We keep varying τ_H until we find the maximized level of tax revenues.

3. For experiments regarding progressivity and welfare maximizing, we implement similar process as above except that we now iterate on τ_{bal}, which is a state and local tax rate, to satisfy the government budget constraint. We then compute CEV until we find the maximizing progressivity.

4. For experiments regarding tax at the top 1% and welfare maximizing, we implement similar process as above and we iterate on τ_{bal}. We compute CEV and keep searching for the τ_H that maximizes CEV.