NBER WORKING PAPER SERIES

# INCOME INEQUALITY AND ASSET PRICES UNDER REDISTRIBUTIVE TAXATION

Lubos Pastor
Pietro Veronesi

Working Paper 21668
http://www.nber.org/papers/w21668

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2015

This paper was written while both authors were visiting the Brevan Howard Centre for Financial Analysis at Imperial College London. Pastor is also at the National Bank of Slovakia. The authors' views are their own and there are no conflicts of interest. We are grateful for helpful comments from Doug Diamond, John Heaton, Marcin Kacperczyk, Stavros Panageas, Jacopo Ponticelli, Raman Uppal, Adrien Verdelhan, and seminar participants at Bocconi, Chicago, Imperial, MIT, and York. We are also grateful to Menaka Hampole and Oscar Eskin for excellent research assistance. This paper has been prepared for the Carnegie-Rochester-NYU Conference Series. It has been supported by the Fama-Miller Center for Research in Finance and the Center for Research in Security Prices, both located at Chicago Booth. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Income Inequality and Asset Prices under Redistributive Taxation
Lubos Pastor and Pietro Veronesi
NBER Working Paper No. 21668
October 2015
JEL No. E24,G1,H2,J24,J31,J38

**ABSTRACT**

We develop a simple general equilibrium model with heterogeneous agents, incomplete financial markets, and redistributive taxation. Agents differ in both skill and risk aversion. In equilibrium, agents become entrepreneurs if their skill is sufficiently high or risk aversion sufficiently low. Under heavier taxation, entrepreneurs are more skilled and less risk-averse, on average. Through these selection effects, the tax rate is positively related to aggregate productivity and negatively related to the expected stock market return. Both income inequality and the level of stock prices initially increase but eventually decrease with the tax rate. Investment risk, stock market participation, and skill heterogeneity all contribute to inequality. Cross-country empirical evidence largely supports the model's predictions.

Lubos Pastor
University of Chicago
Booth School of Business
5807 South Woodlawn Ave
Chicago, IL 60637
and NBER
lubos.pastor@chicagobooth.edu

Pietro Veronesi
University of Chicago
Booth School of Business
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
pietro.veronesi@chicagobooth.edu

# 1. Introduction

In recent decades, income inequality has grown in most developed countries, triggering widespread calls for more income redistribution.[1] Yet the effects of redistribution on inequality are not fully understood. We analyze these effects through the lens of a simple model with heterogeneous agents and incomplete markets. We find that redistribution affects inequality not only directly, by transferring wealth, but also indirectly through selection, by changing the composition of agents who take on investment risk. Through the same selection mechanism, redistribution affects aggregate productivity and asset prices.

Income inequality has been analyzed extensively in labor economics, with a primary focus on wage inequality.[2] While wages are clearly the main source of income for most households, substantial income also derives from business ownership and investments in financial markets, whose size has grown alongside inequality.[3] We examine the channels through which financial markets and business ownership affect inequality. To emphasize those channels, we develop a model in which agents earn no wages; instead, they earn business income, capital income, and tax-financed pensions. In our model, investment risk and differences in financial market participation are the principal drivers of income inequality.

Our model features agents heterogeneous in both skill and risk aversion. Agents optimally choose to become one of two types, "entrepreneurs" or "pensioners." Entrepreneurs are active risk takers whose income is increasing in skill and subject to taxation. Pensioners live off taxes paid by entrepreneurs. Financial markets allow entrepreneurs to sell a fraction of their own firm and use the proceeds to buy a portfolio of shares in other firms and risk-free bonds. Since entrepreneurs cannot diversify fully, markets are incomplete.

In equilibrium, agents become entrepreneurs if their skill is sufficiently high or risk aversion sufficiently low, or both. Intuitively, low-skill agents become pensioners because they would earn less as entrepreneurs, and highly risk-averse agents become pensioners because they dislike the idiosyncratic risk associated with entrepreneurship. These selection effects are amplified by higher tax rates because those make entrepreneurship less attractive. When the tax rate is high, only agents with the highest skill and/or lowest risk aversion find it optimal to become entrepreneurs. Therefore, under heavier taxation, entrepreneurs are more skilled and less risk-averse, on average, and total output is lower.

[1] For example, Alvaredo et al. (2013), Atkinson, Piketty, and Saez (2011), and many others document the growth in inequality. Piketty (2014), the Occupy Wall Street movement, and others call for redistribution.

[2] See, for example, Autor, Katz, and Kearney (2008), among many others.

[3] Non-wage income is earned by households across the whole income distribution, and it is the dominant source of income at the top. Kacperczyk, Nosal, and Stevens (2015) show that non-wage income represents 44% of total income for households that participate in financial markets in 1989 to 2013.

Inequality initially increases but eventually decreases with the tax rate. When the tax rate is zero, all agents choose to be entrepreneurs because pensioners earn no income. As the rate rises, inequality rises at first because agents who are extremely risk-averse or unskilled become pensioners. Such agents accept the low consumption of pensioners in exchange for shedding idiosyncratic risk, thereby increasing consumption inequality.[4] As the tax rate rises further, inequality declines due to the direct effect of redistribution.

There are three sources of inequality: investment risk, stock market participation, and heterogeneity in skill. Investment risk causes differences in ex-post returns on entrepreneurs' portfolios, in part due to idiosyncratic risk and in part because entrepreneurs with different risk aversions have different exposures to stocks. While entrepreneurs participate in the stock market, pensioners do not. Entrepreneurs consume more than pensioners on average, in part due to higher skill and in part as compensation for taking on more risk. Finally, not surprisingly, more heterogeneity in skill across entrepreneurs implies more inequality.

To explore the welfare implications of redistribution, we analyze inequality in expected utility, which we measure by certainty equivalent consumption. Inequality in expected utility is much smaller than consumption inequality, in part because pensioners tend to consume less than entrepreneurs but also face less risk. An increase in the tax rate reduces inequality in expected utility but it also reduces the average level of expected utility. In addition, the model yields a right-skewed distribution of consumption across agents.

The model's asset pricing implications are also interesting. First, the expected stock market return is negatively related to the tax rate. The reason is selection: a higher tax rate implies lower average risk aversion among stockholders, which in turn implies a lower equity risk premium. Second, the level of stock prices exhibits a hump-shaped relation to the tax rate. On the one hand, a higher tax rate reduces stock prices by reducing the after-tax cash flow to stockholders. On the other hand, both selection effects mentioned earlier push in the opposite direction. When the tax rate is higher, entrepreneurs are more skilled, on average, resulting in higher expected pre-tax cash flow, and they are also less risk-averse, resulting in lower discount rates. Both selection effects thus induce a positive relation between stock prices and the tax rate. The net effect is such that the stock price level initially rises but eventually falls with the tax rate. This pattern in market prices feeds back into income inequality, contributing to its hump-shaped pattern. Finally, the model implies a positive relation between the tax rate and aggregate productivity. The reason, again, is selection: a higher tax rate implies that those who create value in the economy are more skilled.

[4] In our simple model, consumption equals income, so consumption inequality equals income inequality.

While our main contribution is theoretical, we also conduct simple cross-country empirical analysis to examine the model's predictions. We use data for 34 OECD countries in 1980 through 2013. We measure the tax burden by the ratio of total taxes to GDP, inequality by the top 10% income share and the Gini coefficient, productivity by GDP per hour worked, the level of stock prices by the aggregate market-to-book ratio, and market returns by the returns on the country's leading stock market index. The evidence is broadly consistent with the model. The tax burden is strongly positively related to productivity, as predicted by the model. The relation between inequality and the tax burden is negative, consistent with the model, but without exhibiting concavity. The relation between the average stock market return and the tax burden is negative, as predicted, but its significance varies depending on which controls are included in the regression. The relation between the level of stock prices and the tax burden is concave and largely negative, as predicted, but the negativity is significant only after the inclusion of three macroeconomic controls.

This paper spans several strands of literature: income inequality, redistributive taxation, entrepreneurship, and asset pricing with heterogeneous preferences and incomplete markets. The vast literature on income inequality focuses largely on labor income, as noted earlier. A recent exception is Kacperczyk, Nosal, and Stevens (2015) who show empirically that inequality in capital income contributes significantly to total income inequality. Kacperczyk et al. also analyze inequality in a model of endogenous information acquisition. In their model, agents have the same risk aversion but different capacities to learn. In addition, assets differ in their riskiness. In our model, assets have the same risk but agents differ in their risk aversion. We also model skill differently, as the ability to deliver a high return on capital rather than the ability to learn about asset payoffs. The two models are complementary, generating different predictions for inequality through different mechanisms.[5]

In our incomplete-markets model, agents can hedge against idiosyncratic risk by trading stocks as well as by borrowing and lending. In addition, agents can escape idiosyncratic risk completely by becoming pensioners and consuming tax revenue. Redistributive taxation thus effectively represents government-organized insurance that supplements the insurance obtainable by trading in financial markets. The insurance benefits of redistribution come at the expense of growth due to a reduced incentive to invest. The tradeoff between insurance benefits and incentive costs of taxation is well known in the optimal taxation literature.[6] Unlike that literature, we do not solve for the optimal tax scheme. Instead, we simply assume proportional taxation, take the tax rate as given, and focus on its implications for

[5] Another mechanism through which capital income can affect inequality has been verbally proposed by Piketty (2014). His capital accumulation arguments extend those of Karl Marx.

[6] This tradeoff features in the models of Eaton and Rosen (1980), Varian (1980), and others.

income inequality and asset prices.

In our model, the tax rate affects the selection of agents into entrepreneurship. Selection based on skill has its roots in Lucas (1978); selection based on risk aversion goes back to Kihlstrom and Laffont (1979). In those models, the alternative to entrepreneurship is working for entrepreneurs; in our model, it is living off taxes paid by entrepreneurs. We show that heavier taxation amplifies both selection effects, with interesting implications for inequality and asset prices. Hombert, Schoar, Sraer, and Thesmar (2014) review other reasons, besides skill and risk aversion, for which agents become entrepreneurs. Hombert et al. also extend Lucas (1978) by making entrepreneurship risky and adding government insurance for failed entrepreneurs. In contrast, in our model, redistribution does not provide insurance against poor ex-post realizations. Instead, it insures agents against being born with low skill or high risk aversion. Agents endowed with such characteristics choose to live off taxes.

Given our emphasis on financial markets, our work has parallels in the asset pricing literature. Like us, Fischer and Jensen (2015) also analyze the effects of redistributive taxation on asset prices. In their model as well as ours, tax revenue is exposed to stock market risk. However, their model has only one type of agents (and thus no selection effects), one risky asset (and thus no idiosyncratic risk), and output that comes from a Lucas tree (and thus does not depend on taxation). Moreover, they focus on stock market participation rather than inequality. Studies that relate inequality to asset prices, in frameworks very different from ours, include Gollier (2001), Johnson (2012), and Favilukis (2013). More broadly, our work is related to the literatures on asset pricing with heterogeneous preferences[7] and uninsurable idiosyncratic income shocks.[8] While we do not calibrate our incomplete-market model with heterogeneous preferences to quantitatively match the data[9], we add endogeneous agent type selection and redistributive taxation. Finally, our paper is related to the literature exploring the links between asset prices and government policy.[10]

The paper is organized as follows. Section 2. develops our model and its implications. Section 3. discusses the model's predictions in more detail. Section 4. analyzes the special case of homogeneous risk aversion. Section 5. reports the empirical results. Section 6. concludes. The proofs of all theoretical results, as well as some additional empirical results, are in the Internet Appendix, which is available on the authors' websites.

[7]See, for example, Dumas (1989), Bhamra and Uppal (2014), and Garleanu and Panageas (2015).

[8]See, for example, Constantinides and Duffie (1996) and Heaton and Lucas (1996).

[9]Studies that calibrate incomplete-market models with heterogeneous preferences to match the data include Gomes and Michaelides (2008) and Gomes, Michaelides, and Polkovnichenko (2013), among others.

[10]See, for example, Croce et al. (2012), Pástor and Veronesi (2012, 2013), and Kelly et al. (2015).

## 2. Model

There is a continuum of agents with unit mass. Each agent $i$ is endowed with a skill level $\mu_i$, risk aversion $\gamma_i$, and $B_{i,0}$ units of capital at time 0. Agents are heterogeneous in both skill and risk aversion but their initial capital is the same, $B_{i,0} = B_0$.

Agents with more skill are more productive in that they earn a higher expected return on their capital if they choose to invest it. Each agent $i$ can invest $B_0$ in a constant-return-to-scale production technology that requires this agent's skill to operate. This technology produces $B_{i,T}$ units of capital at a given future time $T$:

$$B_{i,T} = B_0 \; e^{\mu_i T + \varepsilon_T + \varepsilon_{i,T}} \; , \tag{1}$$

where $\varepsilon_T$ and $\varepsilon_{i,T}$ are aggregate and idiosyncratic random shocks, respectively. These shocks are distributed so that all $\varepsilon_{i,T}$ are i.i.d. across agents and $\mathrm{E}(e^{\varepsilon_T}) = \mathrm{E}(e^{\varepsilon_{i,T}}) = 1$. Agent $i$'s skill, $\mu_i$, is therefore equal to the expected rate of return on the agent's capital:

$$\mathrm{E}\left[\frac{B_{i,T}}{B_0}\right] = e^{\mu_i T} \; . \tag{2}$$

Agent $i$ has a constant relative risk aversion utility function over consumption at time $T$:

$$U\left(C_{i,T}\right) = \frac{C_{i,T}^{1-\gamma_i}}{1-\gamma_i} \; , \tag{3}$$

where $C_{i,T}$ is the agent's consumption and $\gamma_i > 0$ is the coefficient of relative risk aversion.[11]

At time 0, each agent decides to become either an entrepreneur or a pensioner. Entrepreneurs invest in risky productive ventures and are subject to proportional taxation. If an agent becomes an entrepreneur, he starts a firm that produces a single liquidating dividend $B_{i,T}$ at time $T$. An entrepreneur can use financial markets to sell off a fraction of his firm to other entrepreneurs at time 0. The proceeds from the sale can be used to purchase stocks in the firms of other entrepreneurs and risk-free zero-coupon bonds. Each entrepreneur faces a constraint inspired by moral hazard considerations: he must retain ownership of at least a fraction $\theta$ of his own firm. Due to this friction, markets are incomplete.

The second type of agents, pensioners, do not invest; they live off taxes paid by entrepreneurs. We interpret pensioners as including not only retirees but also anyone collecting income from the government without directly contributing to total output, such as government workers (whose contribution is indirect and difficult to quantify), people on disability,

[11] The mathematical expressions presented here assume $\gamma_i \neq 1$. For $\gamma_i = 1$, the agent's utility function is $\log(C_{i,T})$ and some of our formulas require slight algebraic modifications. See the Internet Appendix.

etc. Becoming a pensioner leads to an immediate depreciation of the agent's initial capital endowment. Pensioners cannot sell claims to their pensions in financial markets.

While the agents' initial capital $B_0$ could be physical or human, the latter interpretation seems more natural. We think of $B_0$ as the capacity to put in a certain amount of labor. This interpretation helps justify two of our assumptions. First, all agents are endowed with the same amount of $B_0$, which can be thought of as eight hours per day. (The skill aspect of human capital is included in $\mu_i$.) Second, by becoming pensioners, agents lose their $B_0$. That is, entrepreneurs deploy their labor productively whereas pensioners do not.

Finally, there is a given tax rate $\tau > 0$. For simplicity, we do not model how the government chooses $\tau$. The sole purpose of taxes is redistribution. All taxes are collected from entrepreneurs at time $T$ and equally distributed among pensioners.

## 2.1. The Agents' Decision

At time 0, each agent chooses one of two options: (1) invest and become an entrepreneur, or (2) do not invest and become a pensioner. Let $\mathcal{I}$ denote the set of agents who decide to invest. The set $\mathcal{I}$ is determined in equilibrium as follows:

$$\mathcal{I} = \left\{ i : V_0^{i,yes} \geq V_0^{i,no} \right\} \ , \tag{4}$$

where $V_0^{i,yes}$ and $V_0^{i,no}$ are the expected utilities from investing and not investing, respectively:

$$V_0^{i,yes} = \mathrm{E}\left[ U\left( C_{i,T} \right) \mid \text{investment by agent } i \right] \tag{5}$$

$$V_0^{i,no} = \mathrm{E}\left[ U\left( C_{i,T} \right) \mid \text{no investment by agent } i \right] \ . \tag{6}$$

As we show below, both $V_0^{i,yes}$ and $V_0^{i,no}$ depend on $\mathcal{I}$ itself: each agent's utility depends on the actions of other agents. Solving for the equilibrium thus involves solving a fixed-point problem. Before evaluating the agents' utilities, we compute their consumption levels.

### 2.1.1. Pensioners' Consumption

Pensioners' only source of consumption at time $T$ is tax revenue, which is the product of the tax rate and the tax base. The tax base is total capital accumulated at time $T$. Since only entrepreneurs engage in production, that capital is given by

$$B_T = \int_{\mathcal{I}} B_{i,T} di \ , \tag{7}$$

so that total tax revenue is $\tau B_T$.[12] Let $m(\mathcal{I}) = \int_{\mathcal{I}} di$ denote the measure of $\mathcal{I}$, that is, the fraction of agents who become entrepreneurs. Since tax revenue is distributed equally among $1 - m(\mathcal{I})$ pensioners, the consumption of any given pensioner at time $T$ is given by

$$C_{i,T} = \frac{\tau B_T}{1 - m(\mathcal{I})} \qquad \text{for all } i \notin \mathcal{I} \ . \tag{8}$$

This consumption, and thus also $V_0^{i,no}$ in equation (6), clearly depend on $\mathcal{I}$.

**Proposition 1:** *Given $\mathcal{I}$, pensioner $i$'s consumption at time $T$ is equal to*[13]

$$C_{i,T} = \tau e^{\varepsilon_T} B_0 \, \mathrm{E}^{\mathcal{I}}\left[e^{\mu_j T} | j \in \mathcal{I}\right] \frac{m(\mathcal{I})}{1 - m(\mathcal{I})} \ . \tag{9}$$

Each pensioner's consumption is the same since $C_{i,T}$ is independent of $i$. Pensioners' consumption increases with $m(\mathcal{I})$ for two reasons: a higher $m(\mathcal{I})$ implies a higher tax revenue as well as fewer tax beneficiaries. In other words, the pie is larger and there are fewer pensioners splitting it. An increase in $\tau$ has a positive direct effect on $C_{i,T}$ by raising the tax rate but also a negative indirect effect by reducing the tax base, as we show later.

Since pensioners do not invest, they do not bear any idiosyncratic risk. Yet their consumption is not risk-free: it depends on the aggregate shock $\varepsilon_T$ because tax revenue depends on $\varepsilon_T$. This result illustrates the limits of consumption smoothing by redistribution.[14]

### 2.1.2. Entrepreneurs' Consumption

Entrepreneur $i$'s firm pays a single dividend $B_{i,T}$, given in equation (1). The fraction $\theta$ of this dividend goes to entrepreneur $i$; $1 - \theta$ goes to other entrepreneurs who buy the firm's shares at time 0. Let $M_{i,0}$ denote the equilibrium market value of firm $i$ at time 0. Entrepreneur $i$ sells $1 - \theta$ of his firm for $(1 - \theta) M_{i,0}$ and uses the proceeds to buy financial assets for diversification purposes. The entrepreneur can buy two kinds of assets: shares in other entrepreneurs' firms and risk-free zero-coupon bonds maturing at time $T$, which are in zero

[12] For notational simplicity, we denote by $\int_{\mathcal{I}} di$ the integral across agents $i$ in a given set $\mathcal{I}$ without explicitly invoking the joint distribution of $\mu_i$ and $\gamma_i$. While much of our analysis is general, we also consider specific functional forms for this distribution in some of the subsequent analysis. Also note that in $\int_{\mathcal{I}} B_{i,t} di$, each agent's capital is scaled by $di$ to take into account the agents' infinitesimal size. Given the continuum of agents, each agent's capital is given by $B_{i,t} di$, but to ease notation, we refer to it simply as $B_{i,t}$. In the same way, we simplify notation for other agent-specific variables such as consumption and firm market value.

[13] The notation $\mathrm{E}^{\mathcal{I}}(x_{i,T} | i \in \mathcal{I})$ denotes the average value of $x_{i,T}$ across all agents in set $\mathcal{I}$. The notation $\mathrm{E}(x_{i,T})$, used elsewhere, is the expected value as of time 0 of the random variable $x_{i,T}$ realized at time $T$.

[14] In our simple model, there is no intertemporal smoothing. In more complicated models, the government could in principle provide more insurance to pensioners by saving in good times and spending more in bad times, though the practical difficulties of saving in good times are well known.

net supply. Let $N_0^{ij}$ denote the fraction of firm $j$ purchased by entrepreneur $i$ at time 0 and let $N_0^{i0}$ be the entrepreneur's (long or short) position in the bond. The entrepreneur's budget constraint is

$$(1-\theta)\, M_{i,0} = \int_{\mathcal{I}\setminus i} N_0^{ij} M_{j,0}\; dj \;+\; N_0^{i0}\;, \tag{10}$$

where the price of a risk-free bond yielding one unit of consumption at time $T$ is normalized to one (i.e., the bond is our numeraire). Entrepreneur $i$'s consumption at time $T$ is therefore

$$C_{i,T} = (1-\tau)\,\theta B_{i,T} + (1-\tau)\int_{\mathcal{I}\setminus i} N_0^{ij} B_{j,T} dj + N_0^{i0} \quad \text{ for all } i \in \mathcal{I}\;. \tag{11}$$

The first term is the after-tax dividend that the entrepreneur pays himself from his own firm. The second term is the after-tax dividend from owning a portfolio of shares of other entrepreneurs' firms. The last term is the number of bonds bought or sold at time 0.

Each entrepreneur chooses a portfolio of stocks and bonds $\left\{N_0^{ij}, N_0^{i0}\right\}$ by maximizing his expected utility $V_0^{i,yes}$ from equation (5). These equilibrium portfolio allocations depend on $\mathcal{I}$, and so does the integral in equation (11); therefore, $C_{i,T}$ and $V_0^{i,yes}$ depend on $\mathcal{I}$ as well.

**Proposition 2**. *Given $\mathcal{I}$, entrepreneur $i$'s consumption at time $T$ is equal to*

$$C_{i,T} \;=\; (1-\tau)\, B_0 e^{\mu_i T}\left[\left\{\theta\left(e^{\varepsilon_T+\varepsilon_{i,T}} - Z\right) + (1-\theta)\,\alpha\left(\gamma_i\right)\left(e^{\varepsilon_T} - Z\right) + Z\right\}\right]\;, \tag{12}$$

*where $\alpha\left(\gamma_i\right)$ and $Z$ are described in Proposition 4. The entrepreneur's asset allocation is*

$$N_0^{ij} \;=\; (1-\theta)\,\alpha(\gamma_i)\,\frac{M_{i,0}}{M_0^P} \tag{13}$$

$$N_0^{i0} \;=\; (1-\theta)\,\left[1-\alpha(\gamma_i)\right]\, M_{i,0}\;, \tag{14}$$

*where $M_0^P$ is the total market value of all entrepreneurs' firms: $M_0^P = \int_{\mathcal{I}} M_{i,0} di$.*

The entrepreneur's consumption in equation (12) increases in $\mu_i$, indicating that more skilled entrepreneurs tend to consume more. We use the qualifier "tend to" because more skilled entrepreneurs can get unlucky by earning unexpectedly low returns on their investments, leading to lower consumption. To emphasize the return component of an entrepreneur's consumption, we rewrite equation (12) as follows:

$$C_{i,T} \;=\; M_{i,0}\left[\theta\left(1+R^i\right) + (1-\theta)\,\alpha\left(\gamma_i\right)\left(1+R^{Mkt}\right) + (1-\theta)\left(1-\alpha\left(\gamma_i\right)\right)\right], \tag{15}$$

where $R^i$ is the stock return of firm $i$ between times 0 and $T$ and $R^{Mkt}$ is the return on the aggregate stock market portfolio over the same period. These returns are defined as[15]

$$R^i \;=\; \frac{(1-\tau)B_{i,T}}{M_{i,0}} - 1 \tag{16}$$

$$R^{Mkt} \;=\; \frac{(1-\tau)B_T}{M_0^P} - 1\;. \tag{17}$$

[15]It can be shown that $1+R^i = \frac{e^{\varepsilon_T+\varepsilon_{i,T}}}{Z}$ and $1+R^{Mkt} = \frac{e^{\varepsilon_T}}{Z}$.

The entrepreneur's consumption in equation (15) is the product of the entrepreneur's initial wealth $M_{i,0}$ and the return on his portfolio, which includes his own firm, the aggregate stock market portfolio, and bonds. After selling $1-\theta$ of his own firm, the entrepreneur invests the fraction $1-\alpha(\gamma_i)$ of the proceeds in bonds and the fraction $\alpha(\gamma_i)$ in an equity portfolio. To see that this equity portfolio is the aggregate stock market, first note from equation (13) that agent $i$ buys the same fractional number of shares of any stock $j \neq i$. Agents whose firms are more valuable can afford to buy more shares in other firms (i.e., $N_0^{ij}$ is increasing in $M_{i,0}$), but they buy the same number of shares in each firm (i.e., $N_0^{ij}$ does not depend on $j$) because all stocks have the same exposure to risk. Yet each agent is more exposed to firms with higher $\mu_j$'s because their shares have higher market valuations. Specifically, agent $i$'s position in stock $j$ as a fraction of the agent's liquid equity portfolio is

$$w_j = \frac{N_0^{ij} M_{j,0}}{(1-\theta)\alpha(\gamma_i) M_{i,0}} = \frac{M_{j,0}}{M_0^P} \,. \tag{18}$$

Since $w_j$ are market capitalization weights, the equity part of each entrepreneur's liquid financial wealth is the aggregate cap-weighted market portfolio whose return is $R^{Mkt}$.

Finally, equation (14) shows that the bond allocation decreases in $\alpha(\gamma_i)$. Since bonds are in zero net supply, high $\alpha(\gamma_i)$'s correspond to negative bond allocations ($N_0^{i0} < 0$, that is, the agent borrows to invest more in the stock market) while low $\alpha(\gamma_i)$'s correspond to positive allocations. Since $\alpha(\gamma_i)$ is decreasing in $\gamma_i$, in equilibrium we have more risk-averse entrepreneurs lending to less risk-averse ones.[16]

### 2.1.3. Who Becomes an Entrepreneur?

Having solved for equilibrium consumption levels in Propositions 1 and 2, we immediately obtain the expected utilities $V_0^{i,yes}$ and $V_0^{i,no}$ from equations (5) and (6). We can then use equation (4) to derive the condition under which agents choose to become entrepreneurs.

**Proposition 3**: *Given $\mathcal{I}$, agent $i$ becomes an entrepreneur if and only if*

$$\begin{aligned} \mu_i \;\; > \;\; & \frac{1}{T}\left[\log\left(\frac{\tau}{1-\tau}\right) + \log\left(\frac{m(\mathcal{I})}{1-m(\mathcal{I})}\right) + \log\left(\mathrm{E}^{\mathcal{I}}\left[e^{\mu_j T} | j \in \mathcal{I}\right]\right)\right] \\ & + \frac{1}{T(1-\gamma_i)}\log\left(\frac{\mathrm{E}\left[e^{(1-\gamma_i)\varepsilon_T}\right]}{\mathrm{E}\left[\left(\theta\left(e^{\varepsilon_T+\varepsilon_{i,T}}-Z\right)+(1-\theta)\alpha(\gamma_i)\left(e^{\varepsilon_T}-Z\right)+Z\right)^{1-\gamma_i}\right]}\right). \end{aligned} \tag{19}$$

[16]We can prove $\alpha'(\gamma_i) < 0$ formally under the assumption that $\alpha(\gamma_i) > 0$ for all $i \in \mathcal{I}$, i.e., that none of the agents short the market portfolio. That assumption, which is sufficient but not necessary, holds for many probability distributions of $\gamma_i$ since the average value of $\alpha(\gamma_i)$ across all entrepreneurs must be one in equilibrium. The proof is in the Internet Appendix, along with the proofs of all other theoretical results.

Equation (19) shows that only agents who are sufficiently skilled—those with sufficiently high $\mu_i$—become entrepreneurs. This statement holds other things, especially $\gamma_i$ and $\mathcal{I}$, equal. Note that $\mu_i$ does not appear on the right-hand-side of (19), except as a negligible part of $\mathrm{E}^{\mathcal{I}}\left[e^{\mu_j T} | j \in \mathcal{I}\right]$. Entrepreneurs thus tend to be more skilled than pensioners.

This selection effect is amplified by higher tax rates. The right-hand side of (19) increases in the tax rate $\tau$, holding $\mathcal{I}$ constant. A higher $\tau$ thus discourages entrepreneurship by raising the hurdle for $\mu_i$ above which agents become entrepreneurs. Moreover, a higher $\tau$ implies a higher average value of $\mu_i$ among entrepreneurs. Intuitively, when the tax rate is high, only the most skilled agents find it worthwhile to become entrepreneurs.

While the effect of skill on the agent's decision is clear, the effect of risk aversion is not, as the right-hand-side of (19) depends on $\gamma_i$ in a non-linear fashion. For many parametric assumptions, though, the right-hand side is increasing in $\gamma_i$. One example in which we can formally prove this monotonicity is $\theta \to 1$; see Section 3. Another example is one in which all risk is idiosyncratic (i.e., $\varepsilon_T = 0$). In both examples, entrepreneurs bear much more risk than pensioners, which is plausible. In such scenarios, we thus obtain another selection effect: agents with higher $\gamma_i$ are less likely to become entrepreneurs. Intuitively, highly risk-averse agents avoid entrepreneurship because they dislike the associated idiosyncratic risk.

It is possible to construct counterexamples in which the selection effect goes the other way. The common feature of such examples is that entrepreneurs bear little risk. Consider $\theta = 0$, so that entrepreneurs bear no idiosyncratic risk. In that case, the right-hand side of (19) is initially increasing but eventually decreasing in $\gamma_i$. The reason is that when $\theta = 0$, both types of agents are exposed only to aggregate risk $\varepsilon_T$. Entrepreneurs with high $\gamma_i$'s can reduce their exposure to this risk (i.e., $\alpha(\gamma_i)$) by buying bonds whereas pensioners' exposure to market risk is fixed, as shown in equation (9).[17] Agents with sufficiently high $\gamma_i$'s become entrepreneurs because doing so allows them to choose low $\alpha(\gamma_i)$ and thus face less risk than

[17] Another way to highlight the pensioners' exposure to market risk is to rewrite equation (9) as

$$C_{i,T} = \frac{\tau M_0^P}{[1 - m(\mathcal{I})]\,[1-\tau]}\left(1 + R^{Mkt}\right) \qquad \text{for all } i \notin \mathcal{I} \; .$$

The market value of total endowment at time 0, before tax, is $M_0^P/(1-\tau)$. Any given pensioner's share of this value is $\tau/[1 - m(\mathcal{I})]$. This share earns the market rate of return between times 0 and $T$. For additional insight, note that the ratio in parentheses in the last term of equation (19) can be rewritten as

$$\text{ratio} = \frac{\mathrm{E}\left[\left(1 + R^{Mkt}\right)^{(1-\gamma_i)}\right]}{\mathrm{E}\left[\left(1 + \theta R^i + (1-\theta)\alpha(\gamma_i)\, R^{Mkt}\right)^{1-\gamma_i}\right]} \; .$$

This ratio captures the relative risk of being a pensioner (numerator) versus an entrepreneur (denominator). When $\theta = 0$ and the agent's risk aversion is average in that $\alpha(\gamma_i) = 1$, the numerator equals the denominator and it makes no difference from the risk perspective whether the agent is a pensioner or an entrepreneur.

they would as pensioners. In practice, though, entrepreneurs do bear idiosyncratic risk (i.e., $\theta > 0$) and that risk is typically large.[18] Therefore, it seems plausible to assume that $\theta$ and the volatility of $\varepsilon_{i,T}$ are large enough so that entrepreneurs bear significantly more risk than pensioners. In such realistic scenarios, we obtain the selection effect emphasized in the previous paragraph: entrepreneurs tend to be less risk-averse than pensioners.

Proposition 3 also shows that a higher mass of entrepreneurs makes it less appealing for any given agent to become an entrepreneur. Mathematically, the right-hand side of equation (19) is increasing in $m(\mathcal{I})$. Intuitively, a higher $m(\mathcal{I})$ makes it more attractive to be a pensioner because there is a larger tax revenue to be shared among fewer pensioners.

In equilibrium, $m(\mathcal{I})$ is always strictly between zero and one. If there were no entrepreneurs ($m(\mathcal{I}) = 0$), the total tax base would be zero, implying zero income for pensioners; as a result, somebody always becomes an entrepreneur. If everybody were an entrepreneur, though, ($m(\mathcal{I}) = 1$), there would be a large unallocated tax to be shared, and it would be worthwhile for some agents to quit, shed idiosyncratic risk, and enjoy positive tax-financed consumption. Mathematically, when $m(\mathcal{I}) \to 0$, the right-hand side of equation (19) goes to $-\infty$, and when $m(\mathcal{I}) \to 1$, the right-hand side goes to $+\infty$.

## 2.2. The Equilibrium

The equilibrium in our model is characterized by the consumption levels and portfolio allocations from Propositions 1 and 2, the agent selection mechanism from Proposition 3, and the conditions for market clearing and asset pricing. The latter conditions are presented in the following proposition, which highlights the equilibrium's fixed-point nature.

**Proposition 4**: *The equilibrium state price density $\pi_T$ is given by*

$$\pi_T = \int_{\mathcal{I}} \left[1 + \theta\left(\frac{e^{\varepsilon_T + \varepsilon_{i,T}}}{Z} - 1\right) + (1-\theta)\alpha(\gamma_i)\left(\frac{e^{\varepsilon_T}}{Z} - 1\right)\right]^{-\gamma_i} di \ , \tag{20}$$

*where $Z$ is the equilibrium price as of time 0 of a security that pays $e^{\varepsilon_T}$ at time $T$, given by*

$$Z = \frac{\mathrm{E}[\pi_T e^{\varepsilon_T}]}{\mathrm{E}[\pi_T]} \ , \tag{21}$$

*$\alpha(\gamma_i)$ satisfies the first-order condition*

$$0 = \mathrm{E}\left[\left\{\theta\left(e^{\varepsilon_T + \varepsilon_{i,T}} - Z\right) + \alpha(1-\theta)\left(e^{\varepsilon_T} - Z\right) + Z\right\}^{-\gamma_i}\left(e^{\varepsilon_T} - Z\right)\right] \tag{22}$$

[18] See, for example, Heaton and Lucas (2000).

*as well as the market-clearing condition*

$$\int_{\mathcal{I}} \alpha\left(\gamma_i\right) w_i \, di = 1 \ , \tag{23}$$

*$w_i$ are the market capitalization weights from equation (18), and $\mathcal{I}$ is determined by (19).*

The proposition relies on a fixed-point condition: given $Z$, we can compute $\alpha(\gamma_i)$ for every $i \in \mathcal{I}$, which then allows us to compute $\pi_T$, which then allows us to compute $Z$. An additional fixed-point relation is that the condition (19), which determines the set $\mathcal{I}$ of entrepreneurs, depends on $\mathcal{I}$ itself. In this section, we assume that distributional assumptions are such that the equilibrium conditions are well defined and the fixed-point system has a solution. We prove the existence of a solution in the special cases in Sections 3. and 4. Assuming such existence here, we characterize the equilibrium properties of asset prices below.

## 2.3. Asset Prices

The state price density from equation (20) can be rewritten in terms of asset returns:

$$\pi_T = \int_{\mathcal{I}} \left[1 + \theta R^i + (1-\theta)\alpha\left(\gamma_i\right) R^{Mkt}\right]^{-\gamma_i} di \ . \tag{24}$$

Note that $\pi_T$ depends on the full distribution of $\gamma_i$ across entrepreneurs.

Consider a security that pays $e^{\varepsilon_T}$ at time $T$. Because the price of this security at time 0 is $Z$ and $\mathrm{E}(e^{\varepsilon_T}) = 1$, the expected return of this security is given by

$$r = \frac{1}{Z} - 1 \ . \tag{25}$$

This expected return plays a key role in the following two propositions.

**Proposition 5**: *The expected return on any stock $i$ between times 0 and $T$ is*

$$\mathrm{E}\left(R^i\right) = r \ . \tag{26}$$

Since $r$ does not depend on $i$, all stocks have the same expected return. This result follows from the fact that all stocks have the same risk exposure. While the stocks of more skilled entrepreneurs have higher expected dividends, such stocks trade at higher prices so that expected returns are equalized across stocks. As a result, the expected return on the aggregate stock market portfolio is also given by equation (26).

The expected return depends on the tax rate $\tau$ through the selection effect of $\tau$ on the risk aversions of agents who become entrepreneurs. This is because the expected return is determined by $Z$, which depends on the state price density in equation (20), which in turn depends on the risk aversions of all entrepreneurs. The right-hand side of equation (19) is increasing in $\tau$, as noted earlier. If it is also increasing in $\gamma_i$, which seems realistic (see our discussion of Proposition 3), then an increase in $\tau$ leads more high-$\gamma_i$ agents to become pensioners. A higher $\tau$ thus reduces the average risk aversion of entrepreneurs. The lower average risk aversion of stockholders then depresses the equity risk premium.

**Proposition 6**: *(a) The market-to-book ratio (M/B) of entrepreneur i's firm is*

$$\frac{M_{i,0}}{B_0} = \frac{(1-\tau)\, e^{\mu_i T}}{1+r} \, . \tag{27}$$

*(b) The M/B of the aggregate stock market portfolio is*

$$\frac{M_0^P}{B_0^P} = \frac{(1-\tau)\, \mathrm{E}^{\mathcal{I}}\left[e^{\mu_j T} | j \in \mathcal{I}\right]}{1+r} \, , \tag{28}$$

*where $B_0^P = m(\mathcal{I})\, B_0$ is the total amount of capital invested at time 0.*

Equation (27) shows in elegant simplicity that stock prices are equal to expected cash flows adjusted for risk. The firm's expected after-tax dividend, $B_0\,(1-\tau)\, e^{\mu_i T}$ (see equation (2)), is discounted at the rate $r$, which performs adjustment for risk. There is no discounting beyond the risk adjustment; as noted earlier, we use the risk-free bond as numeraire, thereby effectively setting the risk-free rate to zero.

The market portfolio's M/B is very similar, except that expected dividends are averaged across entrepreneurs. The dependence of M/B on the tax rate $\tau$ is ambiguous. On the one hand, a higher $\tau$ reduces M/B by reducing the after-tax cash flow through the $(1-\tau)$ term. On the other hand, a higher $\tau$ increases M/B by increasing the average skill among entrepreneurs, and thus also $\mathrm{E}^{\mathcal{I}}\left[e^{\mu_j T} | j \in \mathcal{I}\right]$, due to the first selection effect discussed earlier. Finally, a higher $\tau$ increases M/B by reducing average risk aversion, and thus also $r$, through the second selection effect discussed above.

## 2.4. Income Inequality

Next, we analyze the model's implications for income inequality across agents. Since all income is received and consumed at time $T$, income and consumption coincide in our simple model. We therefore focus on the inequality in consumption at time $T$, which is equivalent to income inequality. We normalize each agent's consumption by its average across all agents:

$$s_{i,T} = \frac{C_{i,T}}{\overline{C}_T} \, , \tag{29}$$

where $\overline{C}_T = \int C_{i,T} di = B_T$. Our first measure of inequality, which we adopt for its analytical tractability, is the variance of $s_{i,T}$ across agents:

$$\text{Var}(s_{i,T}) = \int (s_{i,T} - 1)^2 \, di \ . \tag{30}$$

Note that the cross-sectional mean of $s_{i,T}$ is equal to one, by construction.

**Proposition 7**: *The variance of consumption across agents at time $T$ is given by*

$$\begin{aligned} \text{Var}(s_{i,T}) \quad &= \quad \frac{\tau^2}{1 - m(\mathcal{I})} + \frac{(1-\tau)^2}{m(\mathcal{I})} \frac{\text{E}^{\mathcal{I}}\left[e^{2\mu_j T} | j \in \mathcal{I}\right]}{\text{E}^{\mathcal{I}}\left[e^{\mu_j T} | j \in \mathcal{I}\right]^2} \times \\ & \quad \times \text{E}^{\mathcal{I}}\left[\left(\frac{1 + \theta R^j + (1-\theta)\,\alpha(\gamma_j)\, R^{Mkt}}{1 + R^{Mkt}}\right)^2 \ \mid j \in \mathcal{I}\right] - 1 \ . \end{aligned} \tag{31}$$

This expression highlights three sources of inequality. The first one is heterogeneity in skill across entrepreneurs: the fraction $\text{E}^{\mathcal{I}}\left[e^{2\mu_j T} | j \in \mathcal{I}\right] \ / \ \text{E}^{\mathcal{I}}\left[e^{\mu_j T} | j \in \mathcal{I}\right]^2$ is intimately related to the coefficient of variation in $e^{\mu_j T}$ across entrepreneurs. Not surprisingly, a larger dispersion in skill translates into larger consumption inequality.

The second source of inequality is differences in ex-post returns on the entrepreneurs' investment portfolios. These differences affect inequality through the term in brackets in the second line of equation (31). Different firms earn different returns $R^j$, due to idiosyncratic risk. Moreover, entrepreneurs have different exposures to the market portfolio, due to differences in $\alpha(\gamma_j)$. Even if all idiosyncratic risk could be diversified away (i.e., $\theta = 0$), cross-sectional heterogeneity in $\gamma_i$ would create ex-post inequality because agents with different risk aversions take different positions in the market portfolio.

The third source of inequality is that entrepreneurs consume more than pensioners on average, for two reasons.[19] First, entrepreneurs tend to be more skilled. Second, they tend to take more risk for which they are compensated by earning a risk premium. Since pensioners' consumption is not exposed to idiosyncratic risk, the inequality in utility between the two types of agents is smaller than the inequality in income. In other words, income inequality exaggerates the dispersion in happiness across agents.

To clarify this third source of inequality, note that $s_{i,T}$ has a mixture distribution:

$$\begin{aligned} s_{i,T} \quad &= \quad \frac{1-\tau}{m(\mathcal{I})} \times \frac{e^{\mu_i T}}{\text{E}^{\mathcal{I}}\left[e^{\mu_j T} | j \in \mathcal{I}\right]} \times \frac{1 + \theta R^i + (1-\theta)\,\alpha(\gamma_i) R^{Mkt}}{1 + R^{Mkt}} \qquad \text{for } i \in \mathcal{I} \qquad (32) \\ &= \quad \frac{\tau}{1 - m(\mathcal{I})} \qquad \text{for } i \notin \mathcal{I} \ . \qquad (33) \end{aligned}$$

[19] We can prove this inequality formally in two special cases: when $\theta \to 1$ and when all risk is idiosyncratic (i.e., $\varepsilon_T = 0$). In both cases, entrepreneurs bear significantly more risk than pensioners, which is realistic.

From equation (32), the average consumption across entrepreneurs is $(1-\tau)/m(\mathcal{I})$. If all entrepreneurs consumed at that level, the third source of inequality would be the only source, and we would have $\mathrm{Var}(s_{i,T}) = \frac{\tau^2}{1-m(\mathcal{I})} + \frac{(1-\tau)^2}{m(\mathcal{I})} - 1$, a simpler version of equation (31).

In addition to variance, we measure inequality by the percentage of income received by the top 10% of the population. Denoting the cumulative density function of $s_{i,T}$ by $F(s_{i,T})$, we compute the top 10% relative income share as

$$\mathrm{Top}_{10}(s_{i,T}) = \int_{s_{10}}^{\infty} s_{i,T}\, dF(s_{i,T}) \ , \tag{34}$$

where we choose $s_{10}$ such that $F(s_{10}) = 0.90$. Given the mixture distribution of $s_{i,T}$,

$$F(s_{i,T}) \;\; = \;\; F(s_{i,T}|i \in \mathcal{I})\, m(\mathcal{I}) + 1_{\left\{s_{i,T} > \frac{\tau}{1-m(\mathcal{I})}\right\}} (1 - m(\mathcal{I})) \ . \tag{35}$$

The first term on the right-hand side cannot be computed without more structure. After imposing such structure, we obtain $F(s_{i,T})$ in closed form in the following section.

# 3. Results under Additional Assumptions

In this section, we make additional assumptions that allow us to prove the existence of the equilibrium and characterize it analytically. The key assumption is $\theta \to 1$, so that entrepreneurs are allowed to sell only a negligible fraction of their firm in capital markets. In addition, we assume that both shocks from equation (1) are normally distributed:

$$\varepsilon_T \;\; \sim \;\; N\left(-\frac{1}{2}\sigma^2 T,\, \sigma^2 T\right) \tag{36}$$

$$\varepsilon_{i,T} \;\; \sim \;\; N\left(-\frac{1}{2}\sigma_1^2 T,\, \sigma_1^2 T\right) \ . \tag{37}$$

The non-zero means ensure that $\mathrm{E}(e^{\varepsilon_T}) = \mathrm{E}(e^{\varepsilon_{i,T}}) = 1$ and the specific structure for the variances helps when we choose parameter values later in this section.

Equation (19) then simplifies so that agents become entrepreneurs if and only if

$$\mu_i - \frac{\gamma_i}{2}\sigma_1^2 \; > \; \frac{1}{T}\left[\log\left(\frac{\tau}{1-\tau}\right) + \log\left(\frac{m(\mathcal{I})}{1-m(\mathcal{I})}\right) + \log\left(\mathrm{E}^{\mathcal{I}}\left[e^{\mu_j T}|j \in \mathcal{I}\right]\right)\right] \ . \tag{38}$$

The selection effects mentioned earlier are now particularly easy to see: agents with higher skill ($\mu_i$) and lower risk aversion ($\gamma_i$) are more likely to become entrepreneurs. To provide additional insights, we rewrite the right-hand side of equation (38) as follows:

$$\mu_i - \frac{\gamma_i}{2}\sigma_1^2 \; > \; \frac{1}{T}\left[\left\{\log\left(\frac{\tau}{1-m(\mathcal{I})}\right) - \log\left(\frac{1-\tau}{m(\mathcal{I})}\right)\right\} + \log\left(\mathrm{E}\left[\frac{B_T}{B_0^P}\right]\right)\right] \ . \tag{39}$$

The difference in the curly brackets reflects the difference between the average consumption levels of pensioners ($\frac{\tau}{1-m(\mathcal{I})}$) and entrepreneurs ($\frac{1-\tau}{m(\mathcal{I})}$), as shown in equations (32) and (33). A larger difference indicates a larger opportunity cost to being an entrepreneur. The last term on the right-hand side reflects the expected growth of total capital. A higher value implies a higher expected tax base and thus a higher expected consumption for pensioners, which again indicates a higher hurdle for being an entrepreneur. Agent $i$ becomes an entrepreneur only if his $\mu_i$ is high enough and $\gamma_i$ is low enough to overcome these aggregate effects.

To obtain closed-form solutions for the equilibrium quantities, we add the assumption that $\mu_i$ and $\gamma_i$ are independently distributed across agents as follows:

$$\mu_i \sim N\left(\overline{\mu}, \sigma_\mu^2\right) \tag{40}$$

$$\gamma_i \sim N\left(\overline{\gamma}, \sigma_\gamma^2\right) 1_{\{\gamma_i > 0\}} \ . \tag{41}$$

That is, skill $\mu_i$ is normally distributed with mean $\overline{\mu}$ and variance $\sigma_\mu^2$. Risk aversion $\gamma_i$ is truncated normal, with truncation at zero and underlying normal distribution with mean $\overline{\gamma}$ and variance $\sigma_\gamma^2$. Given these distributional assumptions, we solve for the equilibrium mass of entrepreneurs $m(\mathcal{I})$. We prove that

$$\frac{\partial m(\mathcal{I})}{\partial \tau} < 0 \ , \tag{42}$$

so that a higher tax rate shrinks the pool of entrepreneurs. This is intuitive since taxes represent a transfer from entrepreneurs to pensioners. A higher tax rate gives agents an incentive to become recipients of taxes rather than their payers. We also solve for equilibrium asset prices and both measures of inequality. All formulas are in the Internet Appendix.

Next, we illustrate the model's implications for income inequality, productivity, and asset prices. We preserve the assumption $\theta \to 1$ and choose the following parameter values for the distributions in equations (36), (37), (40), and (41): $\sigma = 10\%$ per year, $\sigma_1 = 30\%$ per year, $T = 10$ years, $\overline{\mu} = 0$, $\sigma_\mu = 5\%$ per year, $\overline{\gamma} = 3$, and $\sigma_\gamma = 0.5$. These choices are of limited importance as our conclusions are robust to a wide range of plausible parameter values.

## 3.1. Selection Effects

Figure 1 shows how agents decide to become entrepreneurs or pensioners. Each point with coordinates $(\gamma_i, \mu_i)$ represents an agent with skill $\mu_i$ and risk aversion $\gamma_i$. The circular contours outline the joint probability density of $\mu_i$ and $\gamma_i$ across agents, indicating confidence regions containing 50%, 90%, 99%, and 99.9% of the probability mass. The threshold lines correspond to the tax rates $\tau$ of 0.1%, 5%, 20%, and 70%. For a given $\tau$, all agents located above the threshold line become entrepreneurs; those below the line become pensioners. We see that agents whose skill is sufficiently high or risk aversion sufficiently low become entrepreneurs. The linear tradeoff between $\mu_i$ and $\gamma_i$ is also clear from equation (38).

The figure also shows that higher taxes discourage entrepreneurship: as $\tau$ rises, the threshold line shifts upward, shrinking the region of entrepreneurs. This effect is much more dramatic for low tax rates: raising $\tau$ from 0.1% to 5% reduces the region by more than raising it from 20% to 70%. When $\tau = 0$, nobody becomes a pensioner because there is no tax revenue for pensioners to consume. When $\tau$ rises from zero to a small value, being a pensioner becomes attractive to agents who are extremely unskilled or extremely risk-averse. Such agents choose the near-zero consumption of pensioners because the prospect of starting a firm and bearing its idiosyncratic risk is even worse. As $\tau$ rises further, the ranks of pensioners grow increasingly slowly, for two reasons. First, the rising mass of pensioners means that each pensioner's share of the tax revenue shrinks. Second, the tax revenue itself grows increasingly slowly, and it begins falling for $\tau$ high enough (the Laffer curve).

Supporting these arguments, Figure 2 shows that $m(\mathcal{I})$ declines with $\tau$ in a convex manner: it reaches the value of 0.5 quickly, at $\tau = 16\%$, but then it declines more slowly, reaching 0.1 at $\tau = 61\%$. Figure 2 also plots the average consumption levels of entrepreneurs and pensioners ($\frac{1-\tau}{m(\mathcal{I})}$ and $\frac{\tau}{1-m(\mathcal{I})}$, respectively). Pensioners consume almost nothing when $\tau$ is near zero, but their consumption grows with $\tau$. Entrepreneurs consume more than pensioners on average for any $\tau$, in part due to higher skill and in part due to compensation for risk. Interestingly, the spread between the two consumption levels widens as $\tau$ rises. The reason is that as $\tau$ grows, entrepreneurs grow increasingly more skilled and less risk-averse compared to pensioners, so their initial wealth is increasingly high and so is their amount of risk-taking. As a result, the income difference between the average entrepreneur and the average pensioner increases with $\tau$. However, when we measure inequality by the variance of consumption across individual agents (equation (31)), we see a hump-shaped pattern.

## 3.2. Sources of Inequality

To understand the hump-shaped pattern in inequality, we decompose the consumption variance from equation (31) into three components and plot them in Figure 3. The first component, plotted at the bottom, is due to the difference between the entrepreneurs' and pensioners' average consumption levels.[20] When $\tau$ is small, so is this component because there are hardly any pensioners (i.e., $m(\mathcal{I}) \approx 1$). Even though the difference between the

[20]This component is equal to $\frac{\tau^2}{1-m(\mathcal{I})} + \frac{(1-\tau)^2}{m(\mathcal{I})} - 1$, as noted earlier.

average consumption levels is large (see Figure 2), this difference does not contribute much to total variance since almost all agents are entrepreneurs. When $\tau$ rises, the component initially rises, for two reasons. First, the difference between the average consumption levels grows with $\tau$, as discussed in the previous paragraph. Second, the mass of pensioners grows as well, making this difference more important. But when $\tau$ grows so large that most agents are pensioners, this difference becomes less important again, leading to a hump-shaped pattern in the first component. In other words, as $\tau$ keeps rising, the fraction of pensioners keeps growing, and inequality declines as more and more agents become equally poor.

The second component of inequality is due to heterogeneity in skill across entrepreneurs. For most values of $\tau$, this is the smallest of the three components. The component declines when $\tau$ rises because the rising threshold for $\mu_i$ reduces the heterogeneity in $\mu_i$ among entrepreneurs. Loosely speaking, when the tax rate is high, heterogeneity in skill does not matter much because all entrepreneurs are highly skilled.

The third component, plotted at the top of Figure 3, is due to differences in returns on the entrepreneurs' investments. This investment risk component, driven by pure luck, is the largest source of inequality for any $\tau$.[21] The component initially rises because a higher $\tau$ selects entrepreneurs whose firms are more valuable. Random fluctuations in firm values are then bigger in units of consumption, pushing consumption variance up. The component eventually declines with $\tau$ because the mass of entrepreneurs shrinks. That is, when the tax rate is high, investment risk does not matter much because few agents invest.

In addition to inequality in consumption, we also compute inequality in expected utility to gain some insight into the welfare implications of redistribution. We express expected utility in consumption terms, based on certainty equivalent consumption levels. Agent $i$'s certainty equivalent consumption, $CE_{i,T}$, is the risk-free consumption that makes the agent equally happy as his equilibrium risky consumption $C_{i,T}$:

$$\frac{(CE_{i,T})^{1-\gamma_i}}{1-\gamma_i} = \mathrm{E}\left[\frac{C_{i,T}^{1-\gamma_i}}{1-\gamma_i}\right] . \tag{43}$$

The certainty equivalent consumption levels for the two types of agents are given by

$$\begin{aligned} CE_{i,T} &= B_0\,(1-\tau)\,e^{\mu_i T} e^{-\frac{1}{2}\gamma_i\left(\sigma^2+\sigma_1^2\right)T} && \text{for } i \in \mathcal{I} && (44) \\ &= B_0\,\tau\,\frac{m\left(\mathcal{I}\right)}{1-m\left(\mathcal{I}\right)}\,\mathrm{E}^{\mathcal{I}}\left[e^{\mu_j T}|j \in \mathcal{I}\right] e^{-\frac{1}{2}\gamma_i\sigma^2 T} && \text{for } i \notin \mathcal{I}\ . && (45) \end{aligned}$$

Since pensioners do not employ their skill, their $CE_{i,T}$'s do not depend on $\mu_i$, but they do depend on $\gamma_i$ because pensioners face aggregate risk. Entrepreneurs' $CE_{i,T}$'s depend on both

[21] In the same spirit, Kacperczyk, Nosal, and Stevens (2015) show empirically that inequality in income derived from financial markets contributes significantly to total income inequality.

$\mu_i$ and $\gamma_i$. We scale each agent's $CE_{i,T}$ by the average $CE_{i,T}$ across all agents, analogous to the scaling in equation (29): $s^{CE}_{i,T} = \frac{CE_{i,T}}{\int CE_{i,T} di}$. We then calculate the variance of $s^{CE}_{i,T}$ across agents, our measure of inequality in expected utility, and plot it against $\tau$.

Figure 4 shows that inequality in expected utility (solid line) is much smaller than inequality in consumption (dotted line). One reason is that realized consumption reflects realizations of random shocks whereas expected utility does not. Another, more subtle, reason is that many risk-averse agents prefer the safer consumption of a pensioner to the riskier consumption of an entrepreneur even though the latter consumption is higher on average. Such agents consume relatively little, enhancing consumption inequality, but their expected utility is relatively high due to the lower risk associated with a pensioner's income.

Figure 4 also shows that, unlike inequality in consumption, inequality in expected utility is a decreasing function of $\tau$. Heavier taxation thus implies less dispersion in ex-ante happiness. However, the average $CE_{i,T}$ across all agents (dashed line) also decreases with $\tau$ because higher $\tau$ implies less investment and thus a lower expected total output (dash-dot line). In other words, a higher $\tau$ makes agents more equal in utility terms, but it also makes the average agent worse off. As $\tau$ rises toward one, all agents become equally unhappy.

Panel A of Figure 5 plots the distribution of realized consumption across agents. Since all pensioners consume the same amount, we plot their consumption by a vertical line whose height indicates the mass of pensioners (1.5% for $\tau = 0.1\%$, 58% for $\tau = 20\%$, and 91% for $\tau = 70\%$). The entrepreneurs' consumption is plotted by a probability density. Consumption is highly right-skewed, for two reasons. First, it is right-skewed among entrepreneurs, due to its convexity in $\mu_i$ and random shocks (see equation (12)). Second, most entrepreneurs consume more than pensioners, due to higher skill and larger risk exposure. This is realistic—right skewness in consumption is well known to exist in the data.

Panel B of Figure 5 plots the distribution of certainty equivalent consumption across agents. Each of the three lines plots a mixture of two distributions, one for each type of agents. Unlike in Panel A, there are no vertical lines; even though all pensioners consume the same amount, their utilities differ due to different risk aversions (see equation (45)). The distributions in Panel B are right-skewed, due to convexity in $\mu_i$, but less so than in Panel A because of the absence of convexity in random shocks (see equation (44)).

Figure 6 plots our second measure of inequality: the income share of the top 10% of agents (equation (34)). This is the measure we use in our empirical analysis. Similar to the first measure, the top income share is a concave function of $\tau$, but its peak occurs earlier so its relation to $\tau$ is largely negative. This pattern is robust to changes in the degree of agent heterogeneity, as we see when we vary $\sigma_\mu$ (Panel A) and $\sigma_\gamma$ (Panel B) around their baseline values of $\sigma_\mu = 5\%$ and $\sigma_\gamma = 0.5$. While the effect of $\sigma_\gamma$ on inequality is small, the effect of $\sigma_\mu$ is large: as mentioned earlier, more dispersion in skill implies more inequality.

## 3.3. Productivity

Figure 7 plots expected aggregate productivity against $\tau$. We compute expected productivity as the annualized expected growth rate of total capital, or $(1/T)\mathrm{E}\left[B_T/(m(\mathcal{I})B_0) - 1\right]$. Productivity increases with $\tau$ due to the selection effect described earlier: a higher $\tau$ implies a higher average level of skill among entrepreneurs. When $\tau$ is high, only the most productive agents are willing to become entrepreneurs. The amount of invested capital is then small but this capital grows fast due to entrepreneurs' high productivity.

Productivity also depends on $\sigma_\mu$ and, to a lesser extent, $\sigma_\gamma$. An increase in $\sigma_\mu$ raises expected productivity in two ways. First, it amplifies the selection effect whereby only sufficiently skilled agents become entrepreneurs. Second, there is a convexity effect whereby more dispersion in individual growth rates increases the aggregate growth rate. For example, if half of agents have high skill and half have low skill, aggregate growth is faster than if all agents have average skill because the high-skill agents more than compensate for the low-skill agents in terms of aggregate growth.[22] In contrast, an increase in $\sigma_\gamma$ depresses productivity because it strengthens the importance of $\gamma_i$ at the expense of $\mu_i$ in the entrepreneur selection mechanism. As a result of the weaker selection on $\mu_i$, an increase in $\sigma_\gamma$ reduces the average $\mu_i$ among entrepreneurs, thereby reducing expected productivity.

We interpret the expected growth rate of capital as productivity because it captures the ratio of output ($B_T$) to input ($m(\mathcal{I})B_0$). As discussed earlier, a natural interpretation of the input is the capacity to work for a given number of hours. Under that interpretation, our productivity variable is income per hour worked, which is also the measure of productivity that we use in our empirical analysis.

## 3.4. Asset Prices

Figure 8 plots the expected return on the market portfolio, annualized, as a function of $\tau$. The expected return falls as $\tau$ rises. This result follows from selection: a higher $\tau$ implies that entrepreneurs are less risk-averse, on average. Given their lower risk aversion, agents

[22]A closely related convexity effect is emphasized by Pástor and Veronesi (2003, 2006) who argue that uncertainty about a firm's growth rate increases the firm's value.

demand a lower risk premium to hold stocks, resulting in a lower expected market return. Like all of our main results, this pattern is robust to changes in $\sigma_\mu$ and $\sigma_\gamma$.

Both $\sigma_\mu$ and $\sigma_\gamma$ affect the expected return. A higher $\sigma_\mu$ lifts the expected return because it strengthens the importance of $\mu_i$ at the expense of $\gamma_i$ in the entrepreneur selection mechanism. As a result of the weaker selection on $\gamma_i$, a higher $\sigma_\mu$ implies a higher average $\gamma_i$ among entrepreneurs, which pushes up the expected return. The effect of $\sigma_\gamma$ on the expected return is parameter-dependent because the state price density depends on the full distribution of $\gamma_i$ across entrepreneurs (see equation (24)). On the one hand, a higher $\sigma_\gamma$ implies a lower average $\gamma_i$ among entrepreneurs through the selection effect discussed above. On the other hand, a higher $\sigma_\gamma$ increases the mass of high-$\gamma_i$ entrepreneurs who have a disproportionately high effect on the state price density. While the former effect reduces the expected return, the latter effect increases it. In Panel B of Figure 8, the latter effect is stronger. But the former effect can be stronger if, for example, $\sigma_\mu$ is low enough and $\tau$ high enough.

Figure 9 plots the level of stock prices, measured by the market portfolio's M/B ratio, as a function of $\tau$. M/B exhibits a concave and mostly negative relation to $\tau$: it increases with $\tau$ until $\tau = 15\%$ but then it decreases. This nonlinear pattern results from the interaction of three effects. On the one hand, a higher $\tau$ directly reduces each firm's market value by reducing the after-tax cash flow to stockholders. On the other hand, both selection effects push the aggregate stock price level up. First, a higher $\tau$ implies that entrepreneurs are more skilled, on average, pushing up the average firm's expected cash flow. Second, a higher $\tau$ implies that entrepreneurs are less risk-averse, on average, pushing down the discount rate. The selection effects prevail initially because they are very strong for small values of $\tau$, as shown in Figure 1, but the direct effect prevails eventually.

Figure 9 also shows that stock prices are substantially affected by both types of heterogeneity across agents. An increase in $\sigma_\mu$ raises M/B by increasing expected cash flow. While a higher $\sigma_\mu$ also increases the discount rate, the former effect prevails. An increase in $\sigma_\gamma$ reduces M/B in two ways, by reducing expected cash flow and increasing the discount rate. But we focus on the dependence of M/B on $\tau$, which is robust to changes in $\sigma_\mu$ and $\sigma_\gamma$.

# 4. Special Case: Common Risk Aversion

In this section, we consider a special case of our model in which all agents have the same risk aversion: $\gamma_i = \gamma$. By removing heterogeneity in risk aversion, this case highlights the effects of this heterogeneity on asset prices and income inequality. To obtain closed-form solutions, we make the distributional assumptions (36), (37), and (40). We no longer assume $\theta \to 1$; we return to the general case in which $\theta$ can be anywhere between zero and one.

With common risk aversion, we obtain significant simplifications that provide additional insights. To begin, we have $\alpha(\gamma) = \alpha$, which implies $\alpha = 1$ (see equation (23)). As a result, the entrepreneur's bond allocation from equation (14) simplifies to $N_0^{i0} = 0$. Since all agents are equally risk-averse, there is no borrowing or lending. All entrepreneurs have the same investment portfolio: $\theta$ in their own firm and $1-\theta$ in the stock market.

## 4.1. The Agents' Decision

First, we solve for the equilibrium set $\mathcal{I}$. Agent $i$ becomes an entrepreneur if and only if

$$\mu_i > \underline{\mu} \ , \tag{46}$$

where

$$\begin{aligned}
\underline{\mu} &= K - \Theta && (47)\\
\Theta &= \frac{1}{T(1-\gamma)} \log\left( \int_{-\infty}^{\infty} \left( \theta e^{-\frac{1}{2}\sigma_1^2 T + \sigma_1 \epsilon} + (1-\theta) \right)^{1-\gamma} \phi(\epsilon; 0, T)\, d\epsilon \right) && (48)\\
K &= \overline{\mu} + \frac{1}{2}\sigma_\mu^2 T + \frac{1}{T}\left[ \log\left(\frac{\tau}{1-\tau}\right) + \log\left( \frac{1-\Phi\left(K-\Theta; \overline{\mu} + T\sigma_\mu^2, \sigma_\mu^2\right)}{\Phi\left(K-\Theta; \overline{\mu}, \sigma_\mu^2\right)} \right) \right] \ . && (49)
\end{aligned}$$

Equation (49) always has a solution for $K$, and this solution is unique. As a result, this economy always has a unique equilibrium.

The mass of agents who become entrepreneurs follows from equations (40) and (46):

$$m(\mathcal{I}) = 1 - \Phi\left(\underline{\mu}; \overline{\mu}, \sigma_\mu^2\right) \ . \tag{50}$$

This mass decreases with the tax rate. Specifically, we prove that $\partial \underline{\mu} / \partial \tau > 0$, which implies

$$\frac{\partial m(\mathcal{I})}{\partial \tau} < 0 \ . \tag{51}$$

This result is intuitive, as explained earlier. The mass of entrepreneurs also decreases with $\theta$, the share of firm $i$ that must be retained by entrepreneur $i$. We show $\partial \underline{\mu} / \partial \theta > 0$, so that

$$\frac{\partial m(\mathcal{I})}{\partial \theta} < 0 \ . \tag{52}$$

This result is also intuitive. A higher $\theta$ makes entrepreneurship less appealing because it increases each entrepreneur's exposure to idiosyncratic risk.

## 4.2. Asset Prices

The state price density $\pi_T$ from equation (20) simplifies dramatically, becoming proportional to a simple exponential function of the aggregate shock $\varepsilon_T$:

$$\pi_T \quad \propto \quad e^{-\gamma \varepsilon_T} \, . \tag{53}$$

Interestingly, $\theta$ does not affect the stochastic discount factor $(\pi_T/\pi_0)$. As noted earlier, all entrepreneurs hold $\theta$ in their own firm and $1-\theta$ in the market. Since all firms have the same risk exposure (equation (1)), everyone's position is symmetric ex ante. Therefore, the risk aversion in the economy is the common risk aversion $\gamma$ and the amount of idiosyncratic risk faced by each entrepreneur, as determined by $\theta$, does not affect equilibrium asset prices.

In contrast, $\theta$ does affect asset prices in the general case in Section 2. When risk aversions differ, agents insure each other by trading bonds: low-$\gamma_i$ agents sell bonds to high-$\gamma_i$ agents. As a result, lower-$\gamma_i$ agents acquire larger positions in the market portfolio, bringing down the value-weighted average risk aversion of those holding the market. When $\theta$ increases, all agents become more exposed to idiosyncratic risk, resulting in additional demand for bonds and thus additional changes in the agents' stock market allocations. Therefore, changes in $\theta$ shift the equilibrium risk aversion of the typical agent holding the market, thereby shifting the state price density in the general case.

We obtain explicit solutions for the M/B ratios for each firm and the market as a whole:

$$\frac{M_{i,0}}{B_0} \quad = \quad (1-\tau)\, e^{\left(\mu_i - \gamma\sigma^2\right)T} \tag{54}$$

$$\frac{M_0^P}{B_0^P} \quad = \quad (1-\tau)\, e^{\left(\overline{\mu} - \gamma\sigma^2\right)T} \left[ \left( \frac{1 - \Phi\left(\underline{\mu}; \overline{\mu} + T\sigma_\mu^2, \sigma_\mu^2\right)}{1 - \Phi\left(\underline{\mu}, \overline{\mu}, \sigma_\mu^2\right)} \right) e^{\frac{1}{2}T^2\sigma_\mu^2} \right] \, . \tag{55}$$

The firm's M/B is equal to expected after-tax cash flow adjusted for risk, as before, but now the risk adjustment is particularly simple. The market's M/B in equation (55) highlights the channels through which the level of stock prices depends on the distribution of skill. The term outside the brackets, $(1-\tau)e^{\left(\overline{\mu}-\gamma\sigma^2\right)T}$, is the expected risk-adjusted after-tax cash flow earned by the entrepreneur with average skill. The term inside the brackets is equal to one if there is no dispersion in skill $(\sigma_\mu = 0)$. If there is dispersion in skill $(\sigma_\mu > 0)$, this term is a product of two terms, both of which are greater than one. The first term, the ratio in parentheses, is greater than one due to the selection effect from equation (46). The second term, $e^{\frac{1}{2}T^2\sigma_\mu^2}$, is greater than one due to the convexity effect discussed earlier.

All three channels described in the previous paragraph operate through cash flow. There are no interesting discount rate effects due to homogeneity in risk aversion. Indeed,

$$\mathrm{E}\left(R^i\right) \;=\; e^{\gamma\sigma^2 T} - 1\ . \tag{56}$$

In contrast, a rich set of discount rate effects are present in the general case of heterogeneous risk aversion, augmenting the cash flow effects described here.

### 4.3. Income Inequality

As before in Section 2.4., we measure income inequality by the variance of scaled consumption $s_{i,T}$ across agents. This variance is now equal to

$$\mathrm{Var}(s_{i,T}) \;=\; \frac{\tau^2}{\Phi\left(\underline{\mu};\overline{\mu},\sigma_\mu^2\right)} + (1-\tau)^2\left[e^{T^2\sigma_\mu^2}\frac{1-\Phi\left(\underline{\mu};\left(\overline{\mu}+2T\sigma_\mu^2\right),\sigma_\mu^2\right)}{\left(1-\Phi\left(\underline{\mu};\left(\overline{\mu}+T\sigma_\mu^2\right),\sigma_\mu^2\right)\right)^2}\right]\left[1+\theta^2\left(e^{\sigma_1^2 T}-1\right)\right]-1. \tag{57}$$

The term in the first brackets is greater than one due to cross-sectional dispersion in skill ($\sigma_\mu^2 > 0$). This term is a product of two terms, $e^{T^2\sigma_\mu^2}$ and a ratio, both of which are greater than one. Not surprisingly, a larger $\sigma_\mu^2$ implies more income inequality. The term in the second brackets is also greater than one, due to the presence of idiosyncratic risk ($\theta > 0$). If all such risk were diversifiable ($\theta = 0$), it would generate no dispersion in income and this term would be equal to one. But when $\theta > 0$, each entrepreneur bears idiosyncratic risk whose ex-post realizations, which are commensurate to their volatility $\sigma_1$, contribute to inequality. Higher $\theta$ implies less diversification and more inequality.

## 5. Empirical Analysis

In this section, we examine the model's predictions empirically. While not all results are strong, the evidence is broadly consistent with the model.

### 5.1. Data and Variable Definitions

We collect country-level annual data for the 34 countries that are members of the Organization for Economic Co-operation and Development (OECD). The data categories include stock prices, taxes, inequality, productivity, and other macroeconomic data.

We measure the level of stock prices by the aggregate market-to-book ratio, or $M/B$. The value of $M/B$ for a given country in a given year is the ratio of $M$ to $B$, where $M$ is the total market value of equity of all public firms in the country at the beginning of the year and $B$ is the total book value of equity at the end of the previous fiscal year. If there are fewer than 10 firms over which the intra-country sums can be computed, we treat $M/B$ as missing. The data come from Datastream's Global Equity Indices databases.

Aggregate stock market index returns, $RET$, come from Global Financial Data (GFD). We download nominal returns, both in local currency terms and in U.S. dollars, directly from GFD. We convert nominal returns into real returns by using inflation data from the OECD. For each country, we use the returns on the country's leading stock market index. The stock market indices are listed in Table A1 of the Internet Appendix.

Our tax variable, $TAX$, measures total taxes to GDP. The value of $TAX$ in a given year is the ratio of total tax revenue in that year, summed across all levels of government, to GDP in the same year. These data come from the OECD Statistics database.

Our main measure of income inequality is the top 10% income share, or $TOP$, obtained from the World Top Income Database. While the database contains data on multiple percentage cutoffs, we use the top 10% share because it has the best data coverage. Our second measure of inequality is the Gini coefficient of disposable income after taxes and transfers, obtained from the OECD Income Distribution database. The data coverage for Gini is not as good as for the top 10% income share; hence we prioritize the latter measure. Both measures exhibit frequent gaps in the data. For example, for New Zealand, the Gini coefficient is 0.335 in 1995 and 0.339 in 2000, with missing data in 1996 through 1999. For Germany between 1961 and 1998, the top 10% income share is available only once every three years, ranging from 30.30% to 34.71%. Given the high persistence in these series, we use linear interpolation to fill in the missing values that are sandwiched between valid entries.

Our measure of productivity is GDP per hour worked, or $PROD$. It is measured in 2005 prices at purchasing power parity in U.S. dollars. The data come from the OECD Statistics database. The remaining macroeconomic variables also come from the OECD. Real GDP growth, or $GDPGRO$, is the growth in the expenditure-based measure of GDP. To capture the level of GDP, we use GDP per capita, or $GDPPC$, also measured in 2005 prices at purchasing power parity in dollars. Finally, $INFL$ measures consumer price inflation.

For all variables, we calculate their time-series averages at the country level. To calculate the average stock market return, we use all available data from GFD. These data begin as early as 1792 for the U.S. but as late as 1994 for Slovakia, 1995 for Poland, and 2001 for Hungary. Since stock returns are notoriously volatile and approximately independent over time, it makes sense to estimate average returns from the longest possible data series. For all other variables, which are much more persistent, we calculate their time-series averages over the period 1980 through 2013. We choose this period to make the time periods underlying the averages reasonably well aligned across variables. This is not straightforward because different datasets begin at different points in time. For example, the tax data are available for 1965 through 2013, the $M/B$ data first appear in 1981, the $PROD$ data begin in 1970, and the Gini coefficient data begin mostly in the 1980s. The data on top income shares begin at disparate times for different countries, some in the 19th century but most in the 1970s and 80s. For the time-series average to be valid, we require at least 10 annual observations. In the Internet Appendix, we show the results from cross-sectional regressions over the longer 1965–2013 period, which lead to the same conclusions as those from 1980–2013.

## 5.2. Empirical Results

Our theory makes predictions about the effects of taxes on stock prices, returns, inequality, and productivity. Since tax burdens are highly persistent over time, we examine their variation across countries. To make causal statements, we would need to assume that tax burdens are assigned to countries randomly. This assumption is clearly strong but not baseless. A country's tax burden reflects the country's preference for the degree of redistribution. Such preferences in turn reflect traditions and cultural values that are exogenous to a large extent. But even if the exogeneity assumption is violated, our empirical analysis is relevant as it examines the key econometric associations predicted by the model.

Figure 10 plots the level of stock prices, $M/B$, against the tax burden, $TAX$, across the OECD countries. For both variables, we plot their time-series averages in 1980–2013, as described earlier. Given the high degree of year-to-year persistence in both variables, it makes sense to average them over time and focus on the cross-country variation.[23] Another reason to take this approach is that a key variable examined below, the average stock market return, is a time-series average, by construction. In addition to plotting the individual country-level observations, the figure plots two lines of best fit, one from the linear cross-country regression of average $M/B$ on average $TAX$ (solid line) and the other from the quadratic regression of $M/B$ on $TAX$ and $TAX$ squared (dashed line).[24] These lines indicate a negative and concave relation between $M/B$ and $TAX$, as predicted by the model. While the concavity is statistically significant ($t = -2.84$), the negativity is not ($t = -0.81$).

Table 1 reports the results from the same regression specifications, with and without

[23]For example, the autocorrelation in $TAX$ exceeds 0.9 for 10 countries and 0.8 for 23 countries.

[24]The slope estimator from this average-on-average cross-sectional regression is sometimes referred to as the "between estimator" in panel data terminology. From now on, we suppress "average" in the description of the variables, so that $M/B$ and $TAX$ refer to a country's time-series averages of these variables.

control variables. The first column reports the above-mentioned results without controls. The remaining columns add the macroeconomic controls introduced earlier: $GDPGRO$, $INFL$, and $GDPPC$.[25] The addition of these controls strengthens the negativity of the relation between $M/B$ and $TAX$; in fact, the relation becomes statistically significant when all three controls are included ($t = -2.16$). The relation is also economically significant: a one-standard-deviation increase in $TAX$ is associated with a decrease in $M/B$ by 0.11, which is substantial relative to the standard deviation of $M/B$.[26] The concavity remains significant in all specifications. These results are consistent with the model.

Figure 11 plots the average stock market return, $RET$, against average $TAX$ across countries. Panel A plots nominal U.S. dollar returns, Panel B plots nominal local currency returns, and Panel C plots real local currency returns. In all three panels, the estimated relation between $RET$ and $TAX$ is negative, as it is in the model. The $t$-statistics range from -2.85 to -3.45 across the three panels, indicating a statistically significant relation.

The negative relation between $RET$ and $TAX$ weakens after adding the three controls, as shown in Table 2. While the relation remains mostly significant when each control is included individually, it turns statistically insignificant (or marginally significant, in Panel B) when all three controls are included at the same time. Of course, with only 33 observations, regressions with four right-hand side variables have limited power. Moreover, the relation is economically significant: a one-standard-deviation increase in $TAX$ is associated with a decrease in the average real local currency return by 1.23% per year.

Figure 12 plots our measure of income inequality, $TOP$, against $TAX$. The estimated relation is clearly negative ($t = -3.53$), as predicted by the model. In the regression of $TOP$ on $TAX$ and $TAX$ squared, though, the quadratic term does not enter significantly; in fact, its point estimate indicates convexity whereas the model predicts concavity. The addition of the three controls does not change the results compared to simple regressions. Table 3 shows a strong negative relation between $TOP$ and $TAX$ that exhibits no significant convexity or concavity. We obtain the same conclusions when we use the Gini coefficient.

Figure 13 plots our productivity measure, $PROD$, against $TAX$. The estimated relation is strongly positive ($t = 4.23$), as predicted by the model. This relation survives the inclusion of the controls, as shown in Table 4. By far the most important control is GDP per capita ($GDPPC$), which enters with a highly significant positive coefficient. This is not surprising

[25]In addition to these three controls, we have also considered the unemployment rate and the 10-year bond yield as control variables. We have dropped these controls because they are almost never statistically significant across all of our tables and because they have less data coverage than our three main controls.

[26]The cross-country standard deviations of $TAX$ and $M/B$ are 7.64% and 0.29, respectively.

since $PROD$ is GDP per hour worked. What is interesting is that even after controlling for GDP per capita, GDP per hour worked is significantly positively related to $TAX$.

We conduct various robustness tests. We consider two measures of income inequality and three measures of stock market returns, as explained earlier. We also estimate the cross-country relations between $TAX$ and the other variables in different ways, as we explain next. We summarize the results here and report the details in the Internet Appendix.

First, instead of running cross-country regressions on time-series averages, we run the cross-country regressions year by year and examine the time series of the estimated cross-sectional slope coefficients, along with 95% confidence intervals. The results are very similar to those reported here, with some additional flavor. For $M/B$, the point estimate of the slope on $TAX$ is negative in 27 of the 33 years over which we have data to run this cross-country regression (1981–2013). Moreover, this estimate is significantly negative in each of the past five years of our sample. Furthermore, five of the six years in which the point estimate is positive occur around year 2000, in which stock valuations as measured by $M/B$ were unusually high (the Internet "bubble"). If we exclude this unusual period, the relation plotted in Figure 10 is significantly negative. This evidence provides additional support for the negative relation between $M/B$ and $TAX$ predicted by the model.

When we estimate the same year-by-year cross-sectional regressions for $RET$, we find negative point estimates of the slope on $TAX$ in every single year between 1965 and 2013, and those estimates are statistically significant in almost all years. The relation between $TOP$ and $TAX$ is significantly negative in each year since 1985. The relation between $PROD$ and $TAX$ is always significantly positive. In short, the relations between $TAX$ and $RET$, $TOP$, and $PROD$ are very robust. Moreover, since our plots of the time series of cross-sectional slopes begin in 1965, they show that our choice of 1980 as the starting date for the between-estimator analysis is not crucial to our conclusions. Finally, we run panel regressions with time fixed effects, again reaching the same conclusions.

# 6. Conclusions

We analyze the effects of redistributive taxation on income inequality and asset prices. The model generates selection effects whereby entrepreneurs tend to be more skilled and less risk-averse when taxation is heavier. Through these selection effects, the model yields a rich set of predictions relating the tax burden to income inequality, aggregate productivity, and asset prices. Cross-country empirical evidence is consistent with those predictions.

Our work can be extended in many ways. Given our focus on redistribution, our simple model features only two types of agents: entrepreneurs, who pay taxes, and pensioners, who consume them. It would be natural to add a third type, "workers," who are employed by entrepreneurs, and to examine labor market implications of redistribution. It would also be interesting to endogenize the tax rate, possibly allowing for progressive taxation, and analyze the welfare implications of our results. Finally, it would be useful to extend our descriptive empirical analysis. We leave these interesting extensions for future research.

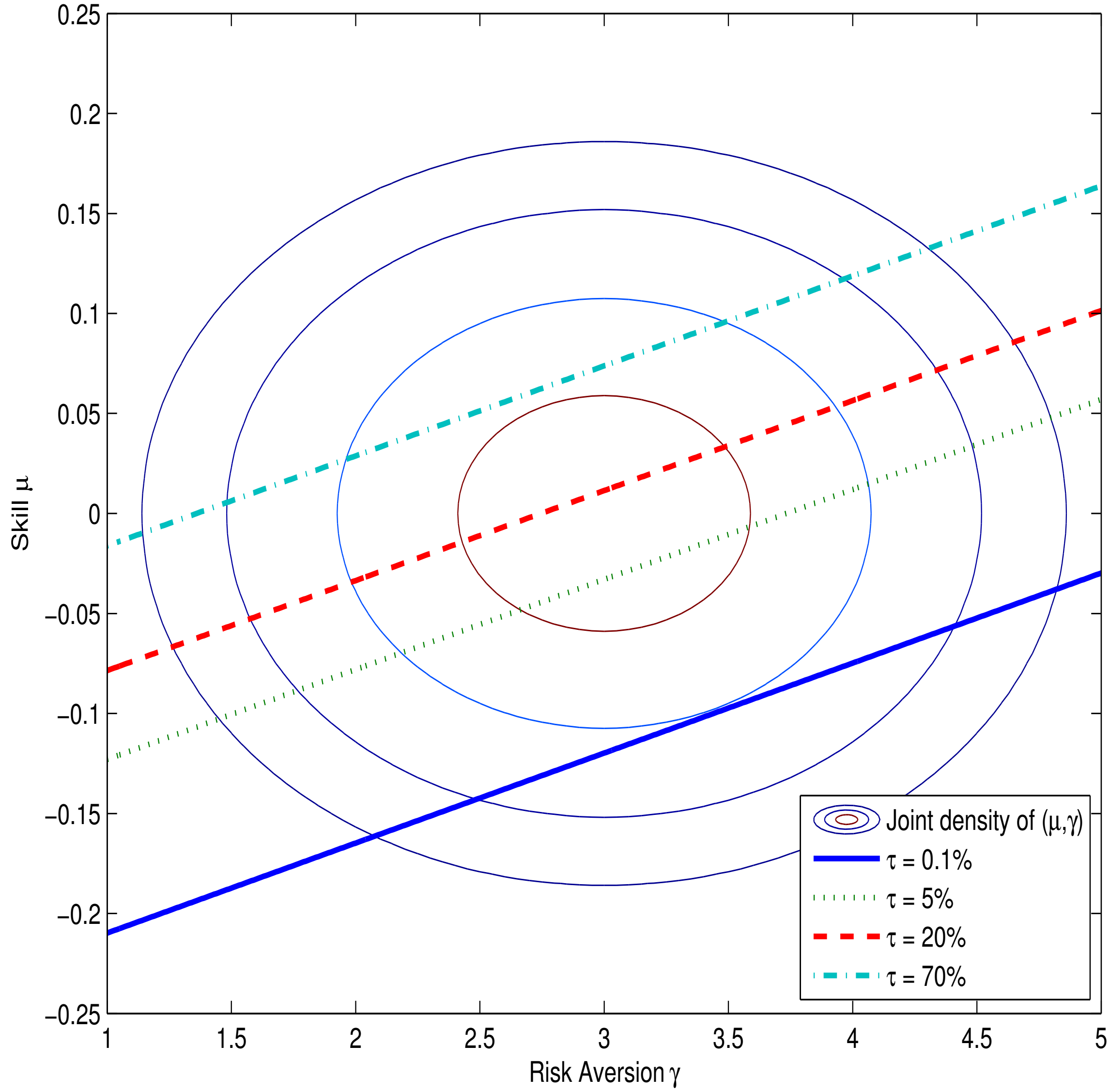



**Figure 1. The agents' decision.** Each point in this graph represents an agent with the corresponding skill $\mu_i$ and risk aversion $\gamma_i$. All agents located above the threshold line choose to become entrepreneurs; those below the line become pensioners. The four lines correspond to four different tax rates $\tau$. The circular contours outline the joint probability density of $\mu_i$ and $\gamma_i$ across agents. The four contours indicate confidence regions containing 50%, 90%, 99%, and 99.9% of the joint probability mass of $\mu_i$ and $\gamma_i$.

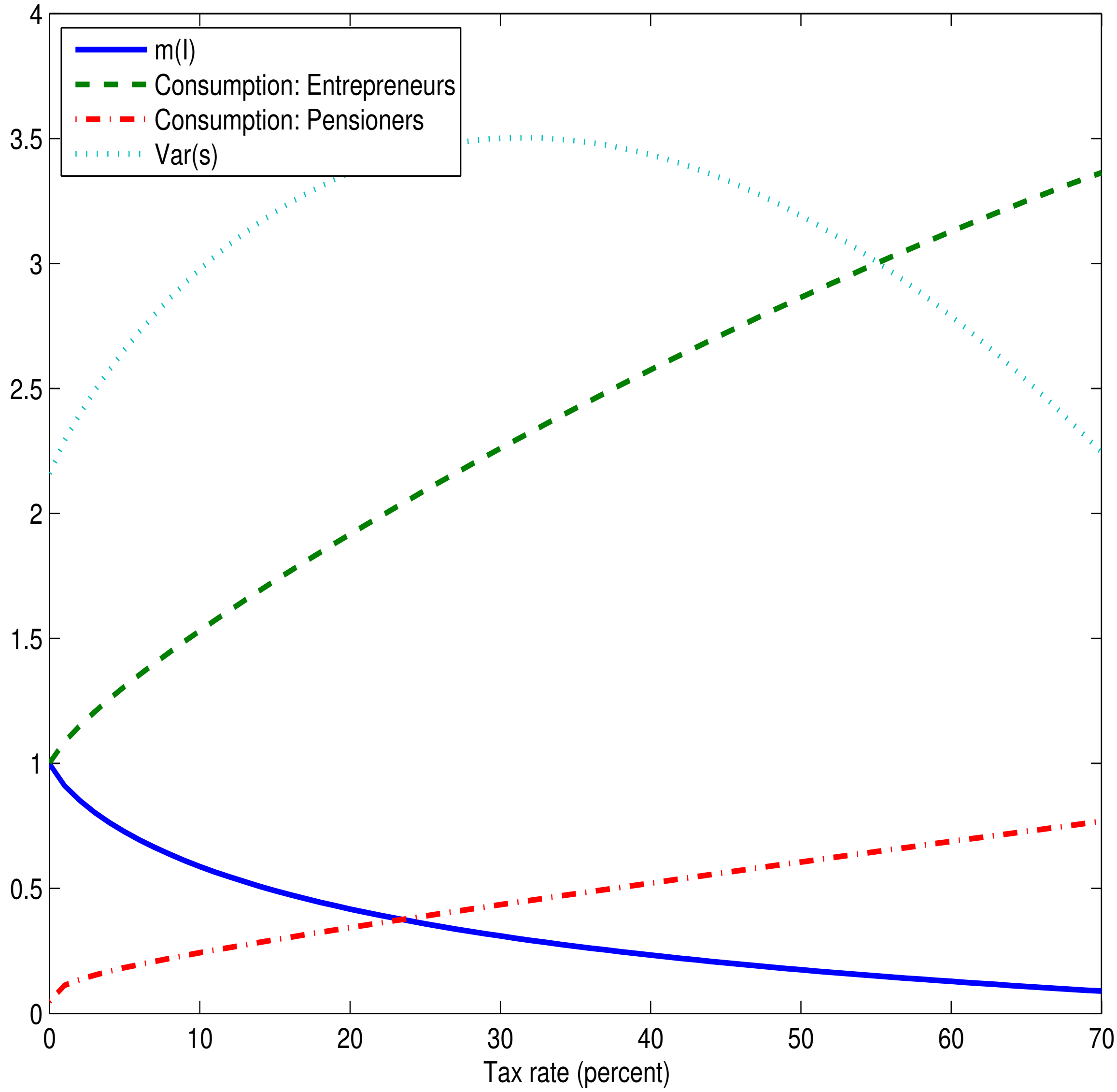



**Figure 2. The share of entrepreneurs and agents' consumption.** This figure plots four quantities as a function of the tax rate $\tau$. The solid line plots $m(\mathcal{I})$, the fraction of agents who become entrepreneurs. The dashed line plots the average consumption of entrepreneurs, which is given by $\frac{1-\tau}{m(\mathcal{I})}$. The dash-dot line plots the consumption of each pensioner, given by $\frac{\tau}{1-m(\mathcal{I})}$. The dotted line plots the variance of consumption across agents. Throughout, consumption is scaled by average consumption across all agents.

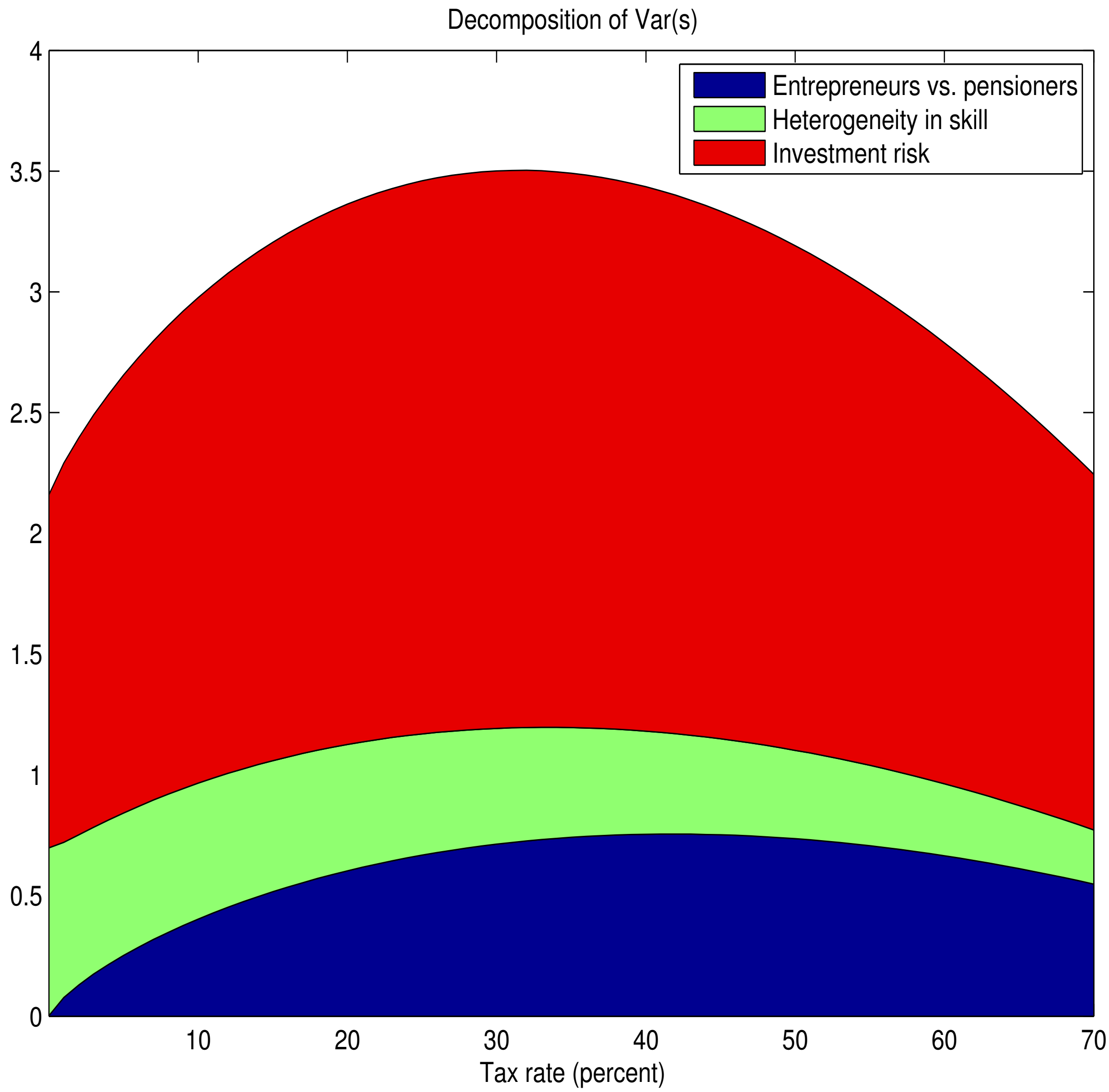



**Figure 3. Three sources of income inequality.** This figure plots three components of the variance of consumption across agents from equation (31) as a function of the tax rate $\tau$. The blue area at the bottom plots the component due to the difference between the entrepreneurs' and pensioners' average consumption levels. This component is equal to the consumption variance if all entrepreneurs consume at the same average level of $(1-\tau)/m(\mathcal{I})$. The red area at the top plots the component due to differences in ex-post returns on the entrepreneurs' investment portfolios. This component is equal to the difference between total variance and the first component under the assumption that all entrepreneurs have the same skill. Finally, the green area in the middle plots the component due to heterogeneity in $\mu_i$ across entrepreneurs. This component is obtained as the residual by subtracting the other two components from total variance.

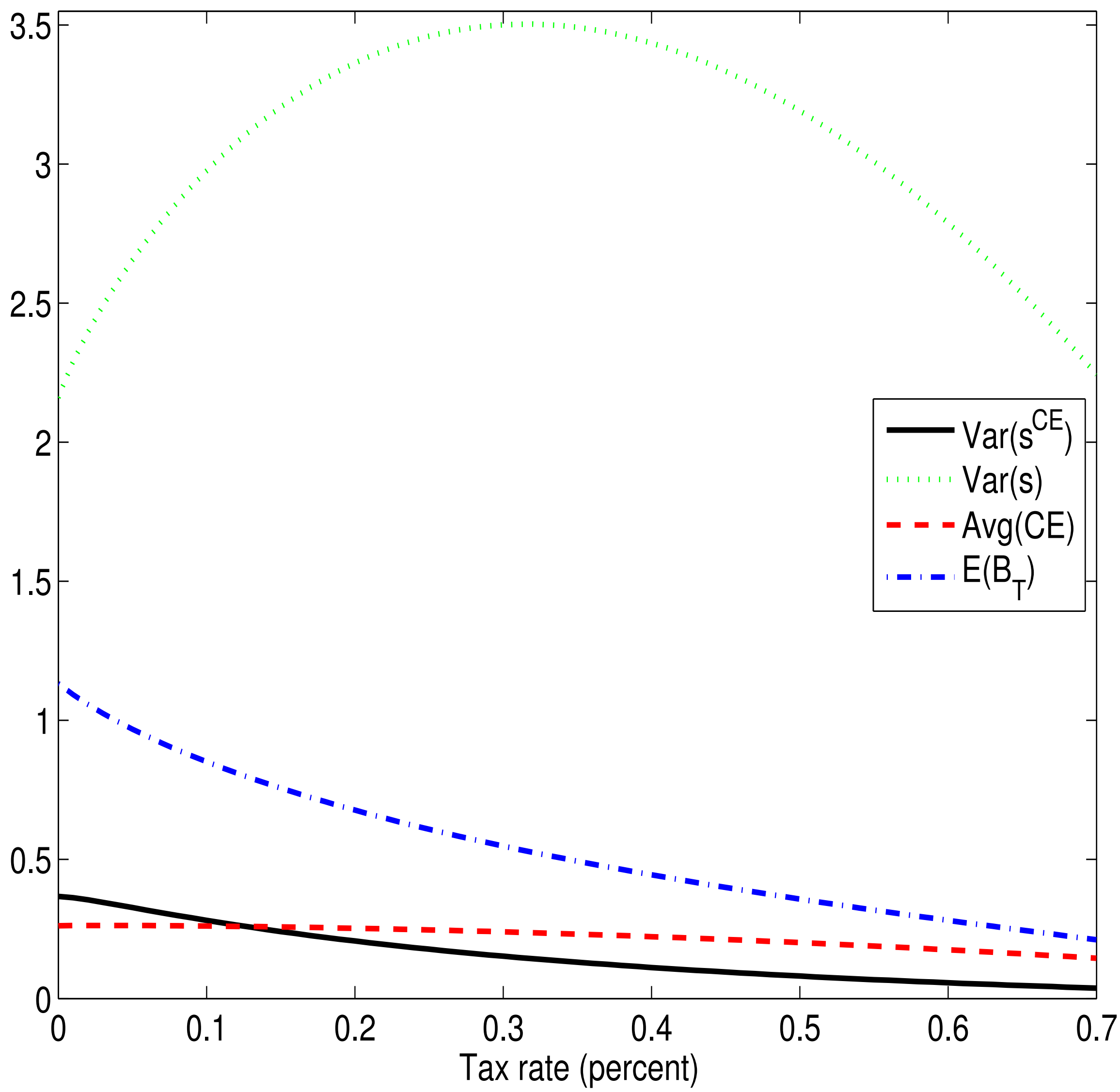



**Figure 4. Inequality in expected utility vs. inequality in consumption.** The solid line plots inequality in expected utility expressed in consumption terms, measured by the variance of certainty equivalent consumption across agents, as a function of the tax rate $\tau$. The dotted line plots inequality in consumption (or, equivalently, income), measured by the variance of consumption across agents. Both consumption and its certainty equivalent are scaled by their averages across all agents. The dashed line plots the average value of unscaled certainty equivalent consumption across all agents. The dash-dot line plots the expected value of total capital $B_T$ as of time 0. Throughout, we normalize $B_0 = 1$.

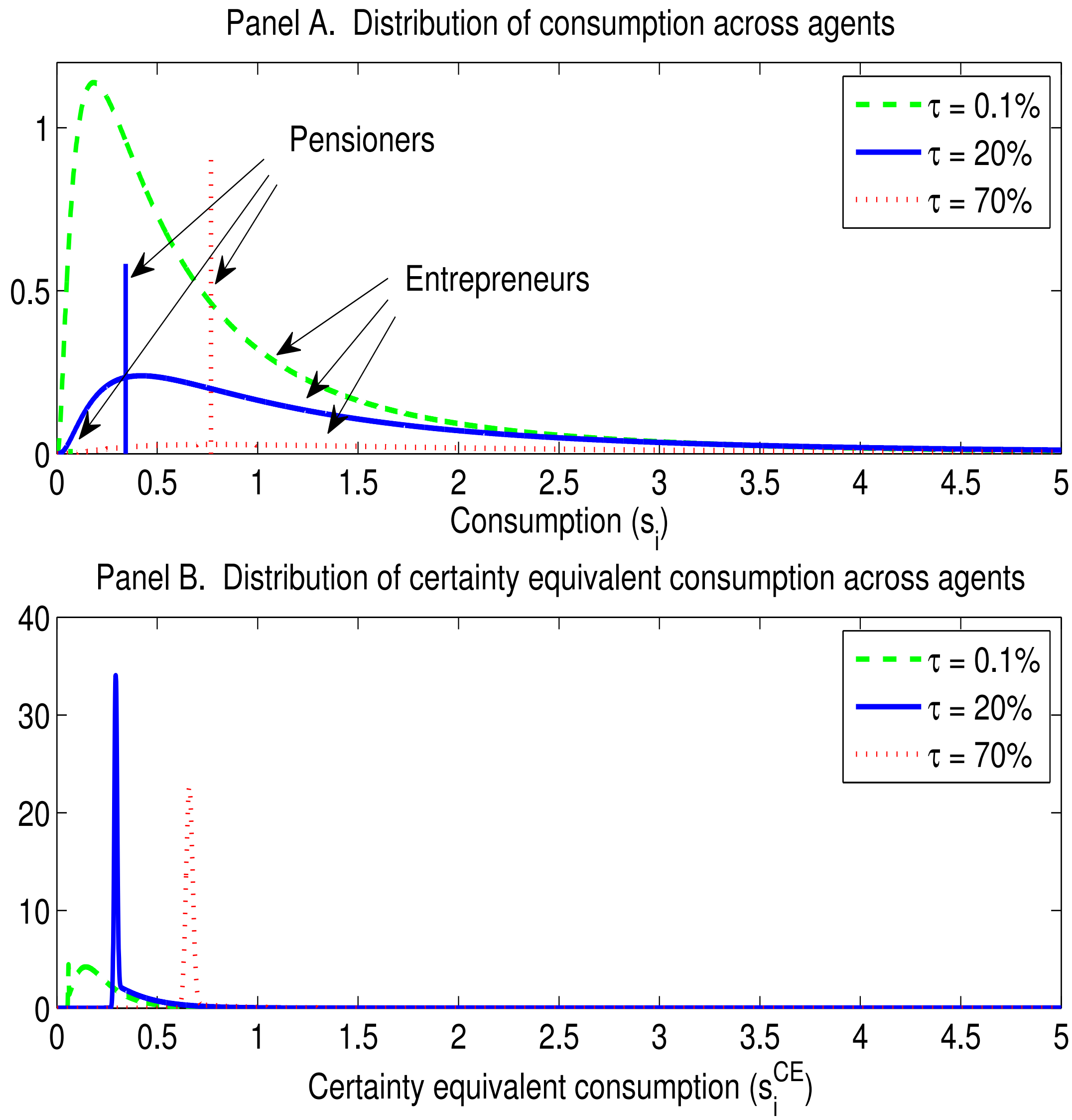



**Figure 5. Distribution of consumption across agents.** Panel A plots the distribution of consumption across agents. The consumption of pensioners is plotted by vertical lines whose height indicates the corresponding probability mass. The consumption of entrepreneurs is plotted by probability densities. Panel B plots the distribution of certainty equivalent consumption. We consider three tax rates $\tau$. Both consumption and its certainty equivalent are scaled by their averages across all agents. Throughout, we normalize $B_0 = 1$.

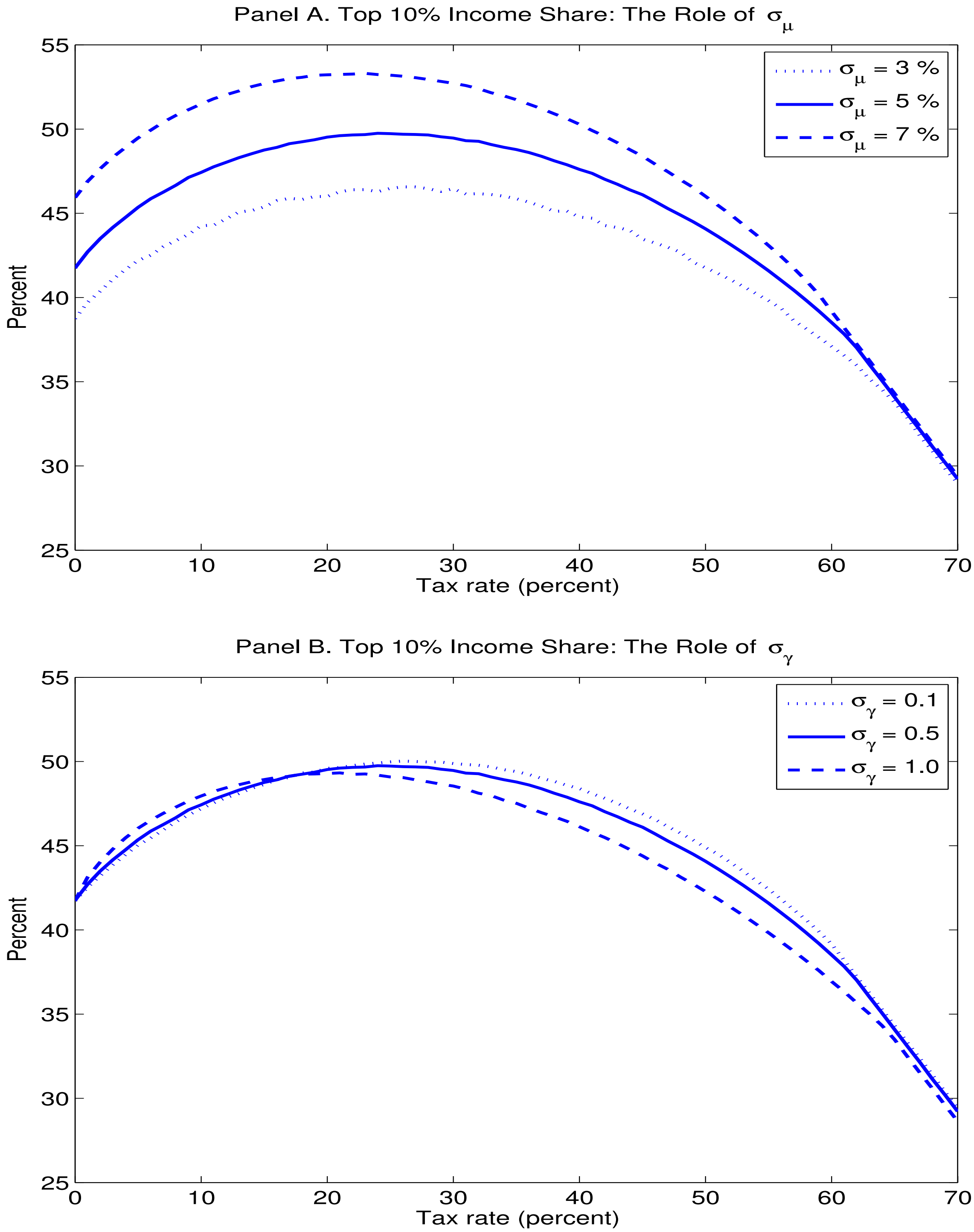
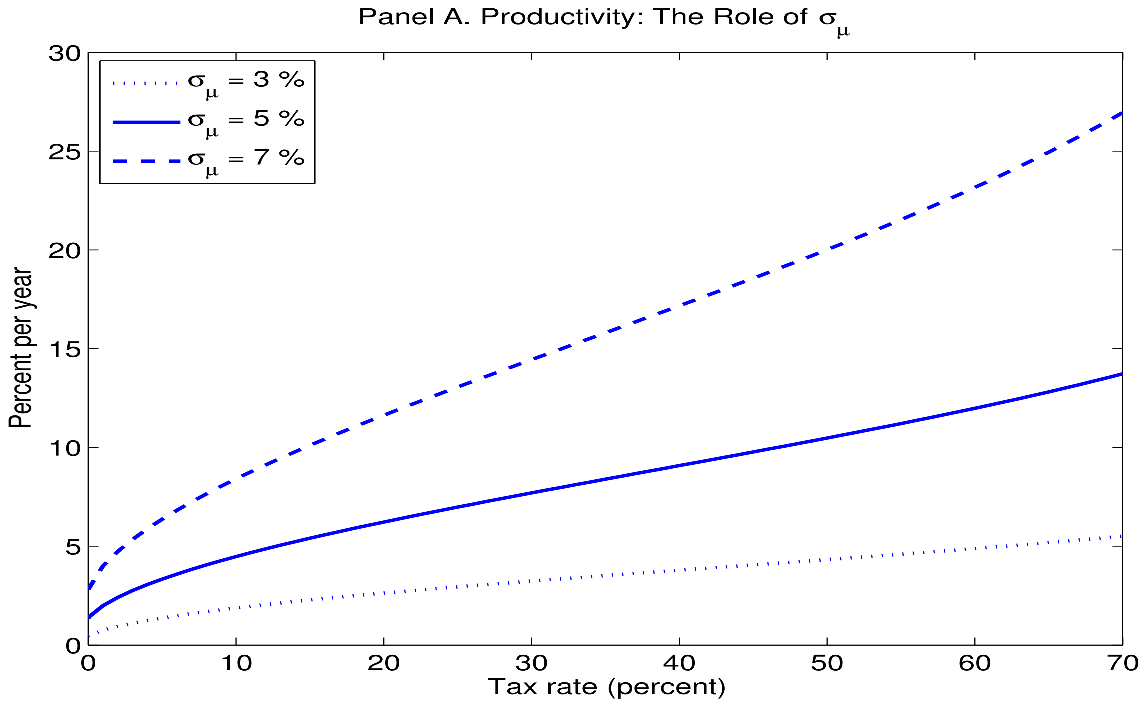



**Figure 6. Model-implied inequality.** This figure plots the income share of the top 10% of agents as a function of the tax rate $\tau$. The solid lines in both panels correspond to the baseline case in which $\sigma_\mu = 5\%$ and $\sigma_\gamma = 0.5$. Panel A varies $\sigma_\mu$ while keeping all other parameters at their baseline values. Panel B varies $\sigma_\gamma$.



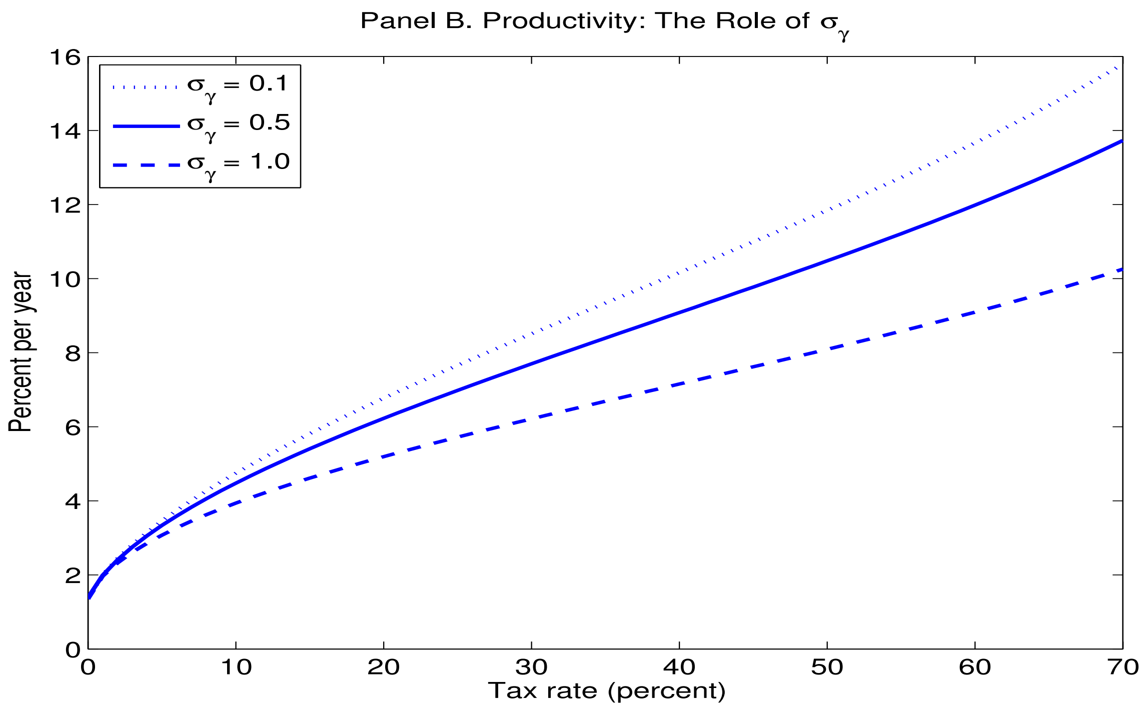




**Figure 7. Model-implied aggregate productivity.** This figure plots the model-implied expected aggregate productivity as a function of the tax rate $\tau$. Expected productivity is computed as the annualized expected growth rate of total capital, or $(1/T)\mathrm{E}\left[B_T/(m(\mathcal{I})B_0) - 1\right]$. The solid lines in both panels correspond to the baseline case in which $\sigma_\mu = 5\%$ and $\sigma_\gamma = 0.5$. Panel A varies $\sigma_\mu$ while keeping all other parameters at their baseline values. Panel B varies $\sigma_\gamma$.

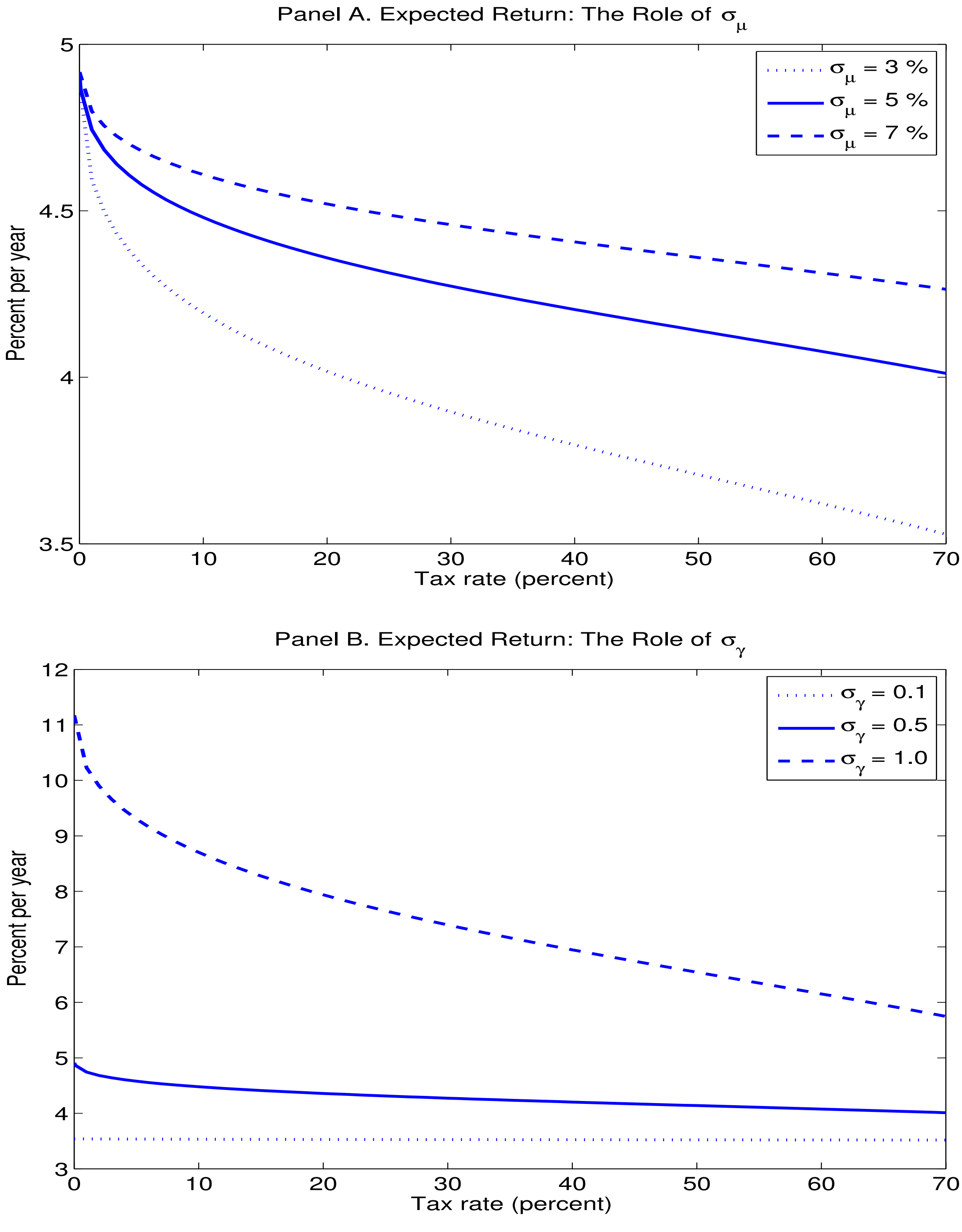



**Figure 8. Model-implied expected return.** This figure plots the model-implied expected rate of return on the aggregate market portfolio as a function of the tax rate $\tau$. The solid lines in both panels correspond to the baseline case in which $\sigma_\mu = 5\%$ and $\sigma_\gamma = 0.5$. Panel A varies $\sigma_\mu$ while keeping all other parameters at their baseline values. Panel B varies $\sigma_\gamma$.

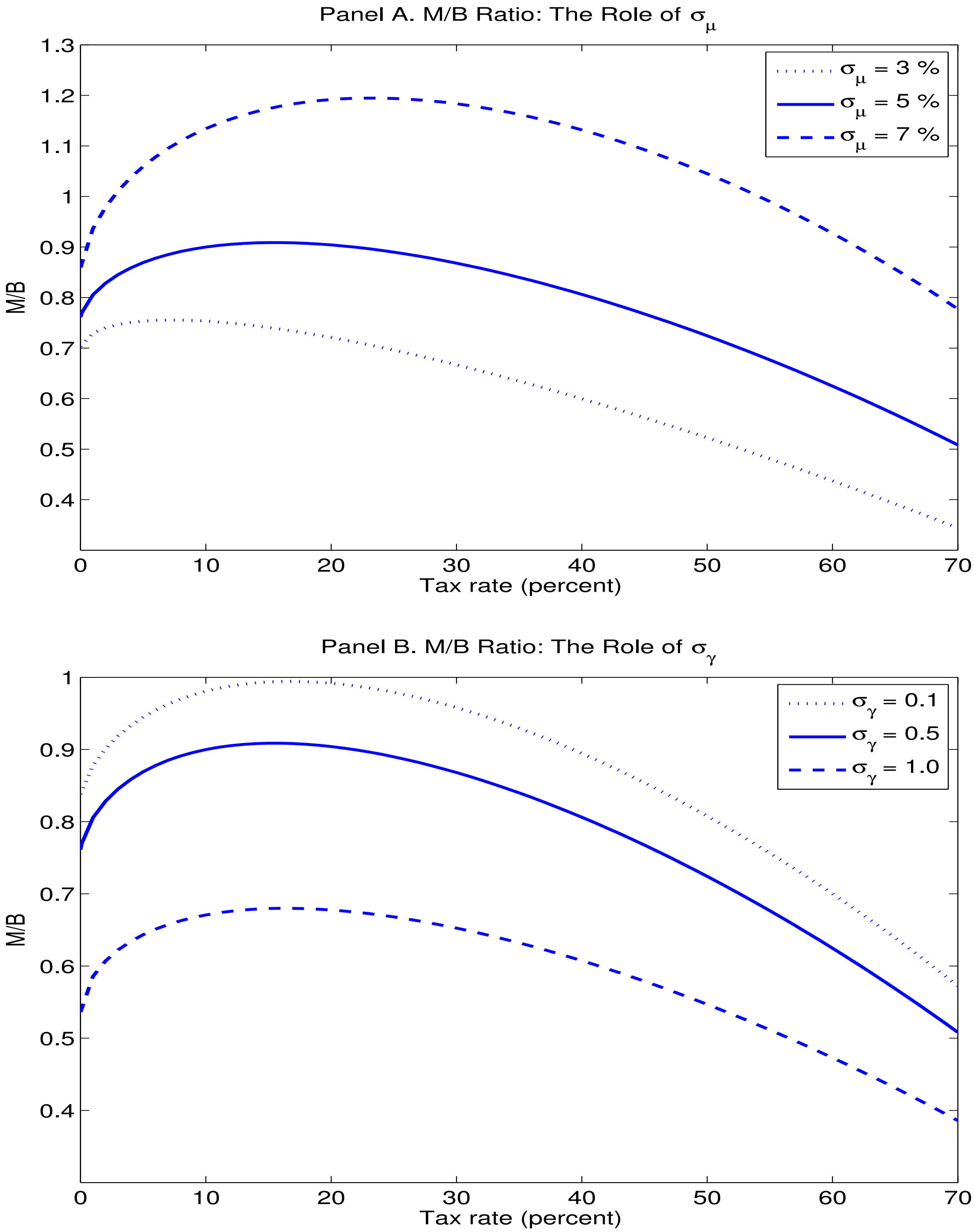



**Figure 9. Model-implied M/B.** This figure plots the model-implied M/B ratio for the aggregate market portfolio as a function of the tax rate $\tau$. The solid lines in both panels correspond to the baseline case in which $\sigma_\mu = 5\%$ and $\sigma_\gamma = 0.5$. Panel A varies $\sigma_\mu$ while keeping all other parameters at their baseline values. Panel B varies $\sigma_\gamma$.

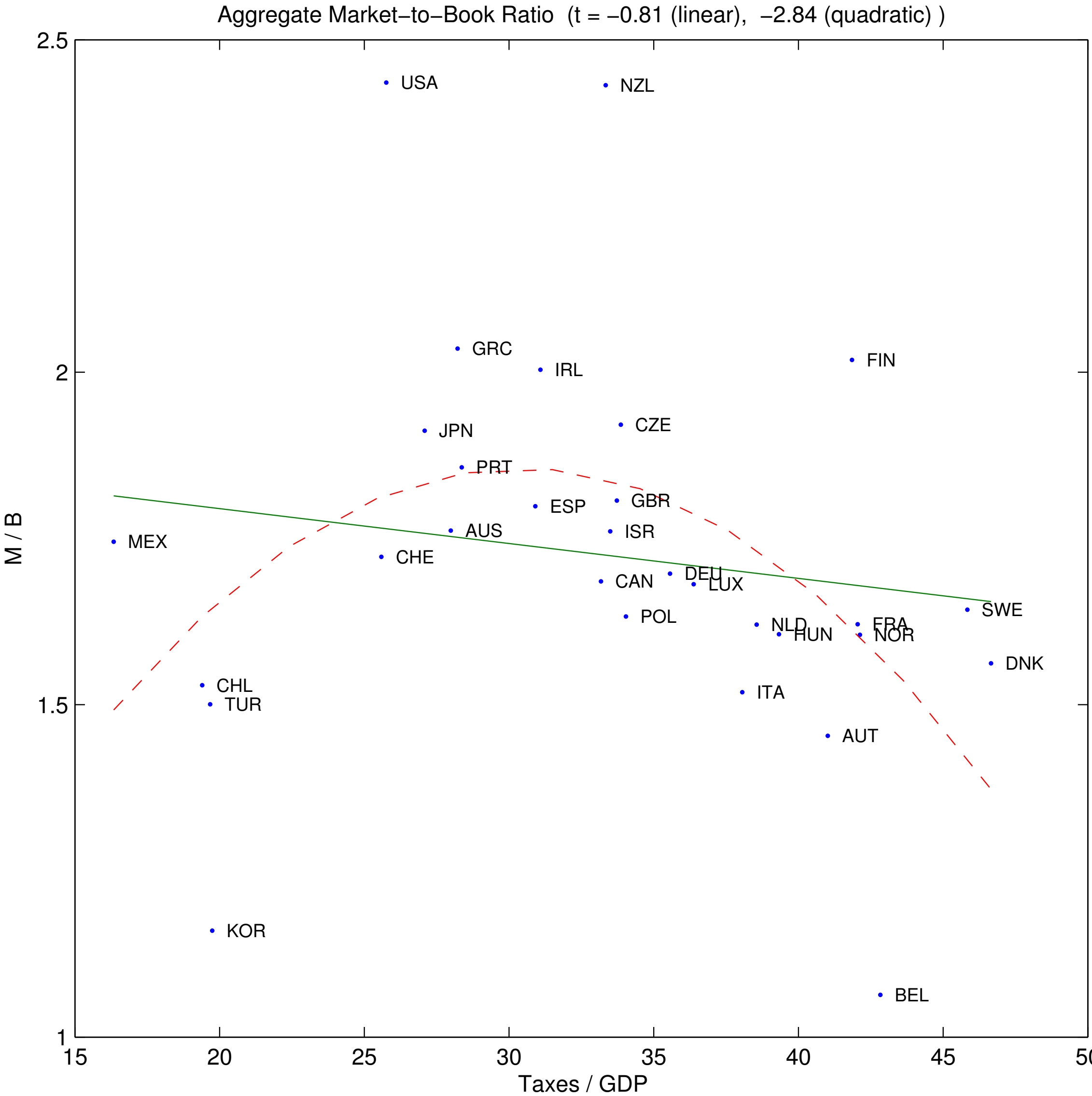



**Figure 10. Stock price level vs. taxes.** This figure plots the market-to-book ratio ($M/B$) against the tax-to-GDP ratio ($TAX$) across countries. Both variables are computed at the country level as time-series averages in 1980–2013. Each dot, labeled with the OECD country code, is a country-level observation. The solid line is the line of best fit from the linear regression of $M/B$ on $TAX$. The dashed line is the line of best fit from the quadratic regression of $M/B$ on $TAX$ and $TAX^2$. The corresponding $t$-statistics are in the panel title. The first $t$-statistic is for the slope on $TAX$ in the linear regression; the second for the slope on $TAX^2$ in the quadratic regression.

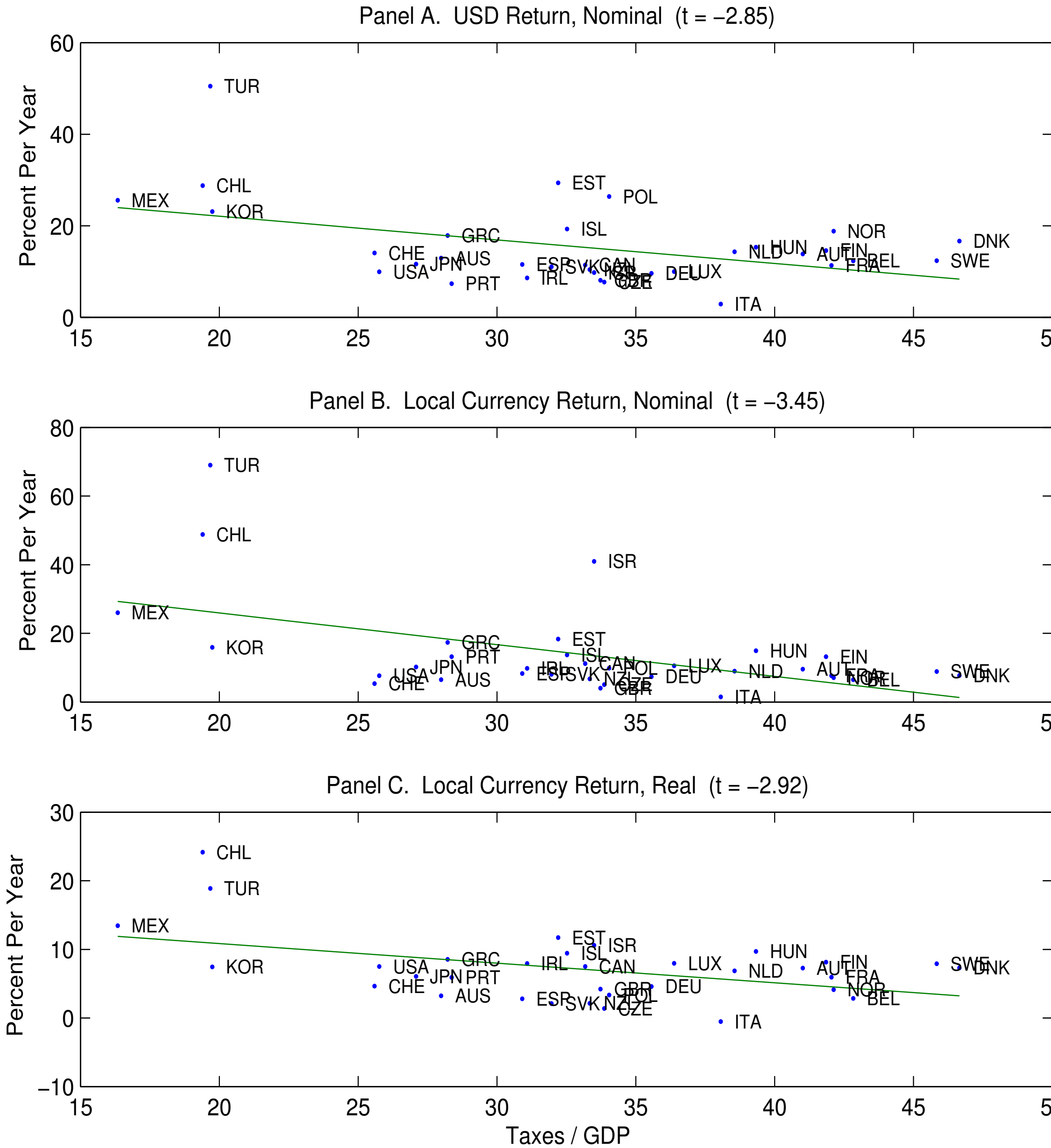



**Figure 11. Average stock market return vs. taxes.** This figure plots the average stock market index return ($RET$) against the tax-to-GDP ratio ($TAX$) across countries. Both variables are computed at the country level as time-series averages; $TAX$ over 1980–2013 and $RET$ using all available data. The three panels plot returns calculated in different ways, as explained in the panel titles. Each dot, labeled with the OECD country code, is a country-level observation. The solid line is the line of best fit from the linear regression of $RET$ on $TAX$. The corresponding $t$-statistics are in the panel titles.

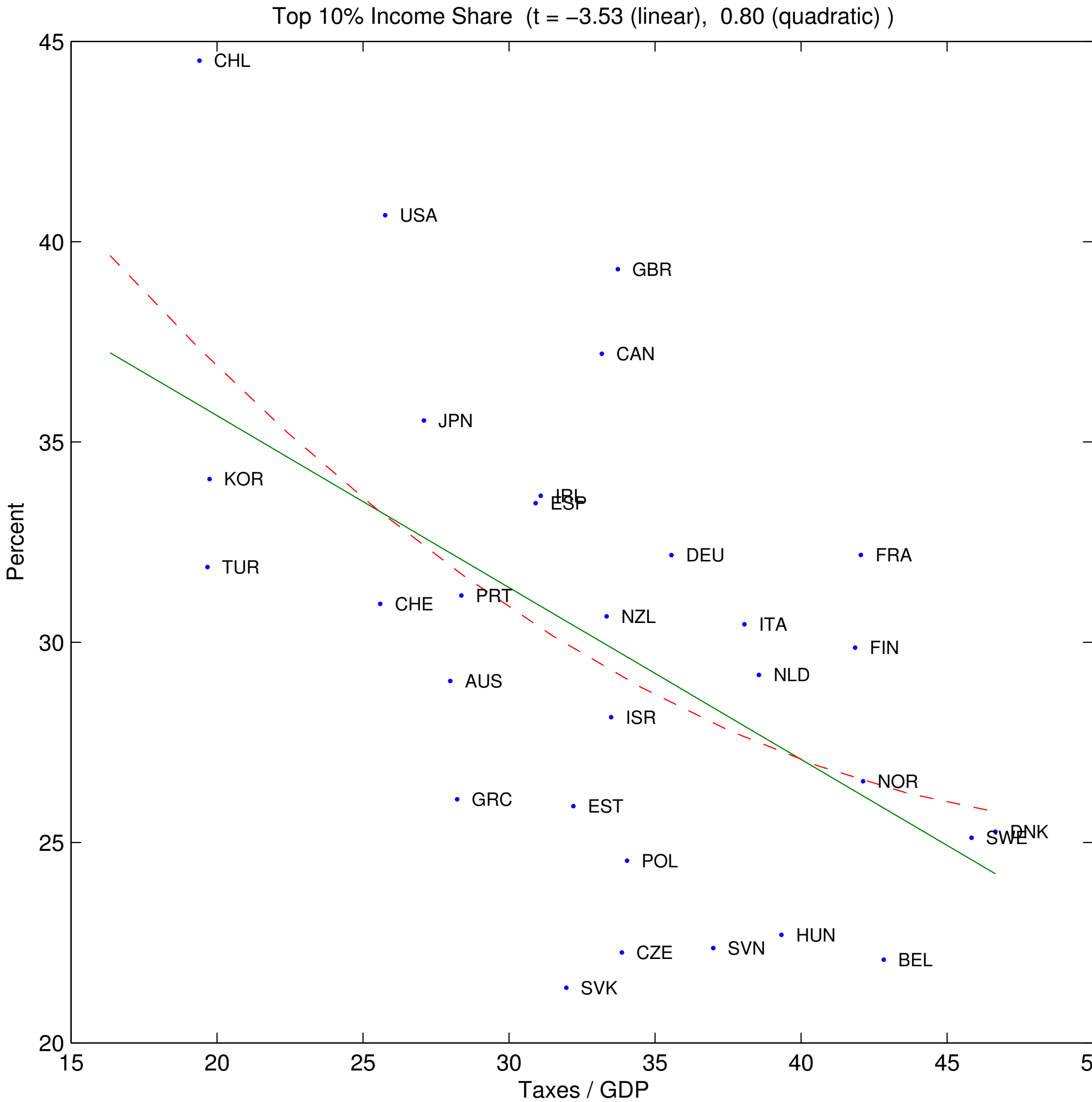



**Figure 12. Income inequality vs. taxes.** This figure plots the top 10% income share ($TOP$) against the tax-to-GDP ratio ($TAX$) across countries. Both variables are computed at the country level as time-series averages in 1980–2013. Each dot, labeled with the OECD country code, is a country-level observation. The solid line is the line of best fit from the linear regression of $TOP$ on $TAX$. The dashed line is the line of best fit from the quadratic regression of $TOP$ on $TAX$ and $TAX^2$. The corresponding $t$-statistics are in the panel title. The first $t$-statistic is for the slope on $TAX$ in the linear regression; the second for the slope on $TAX^2$ in the quadratic regression.

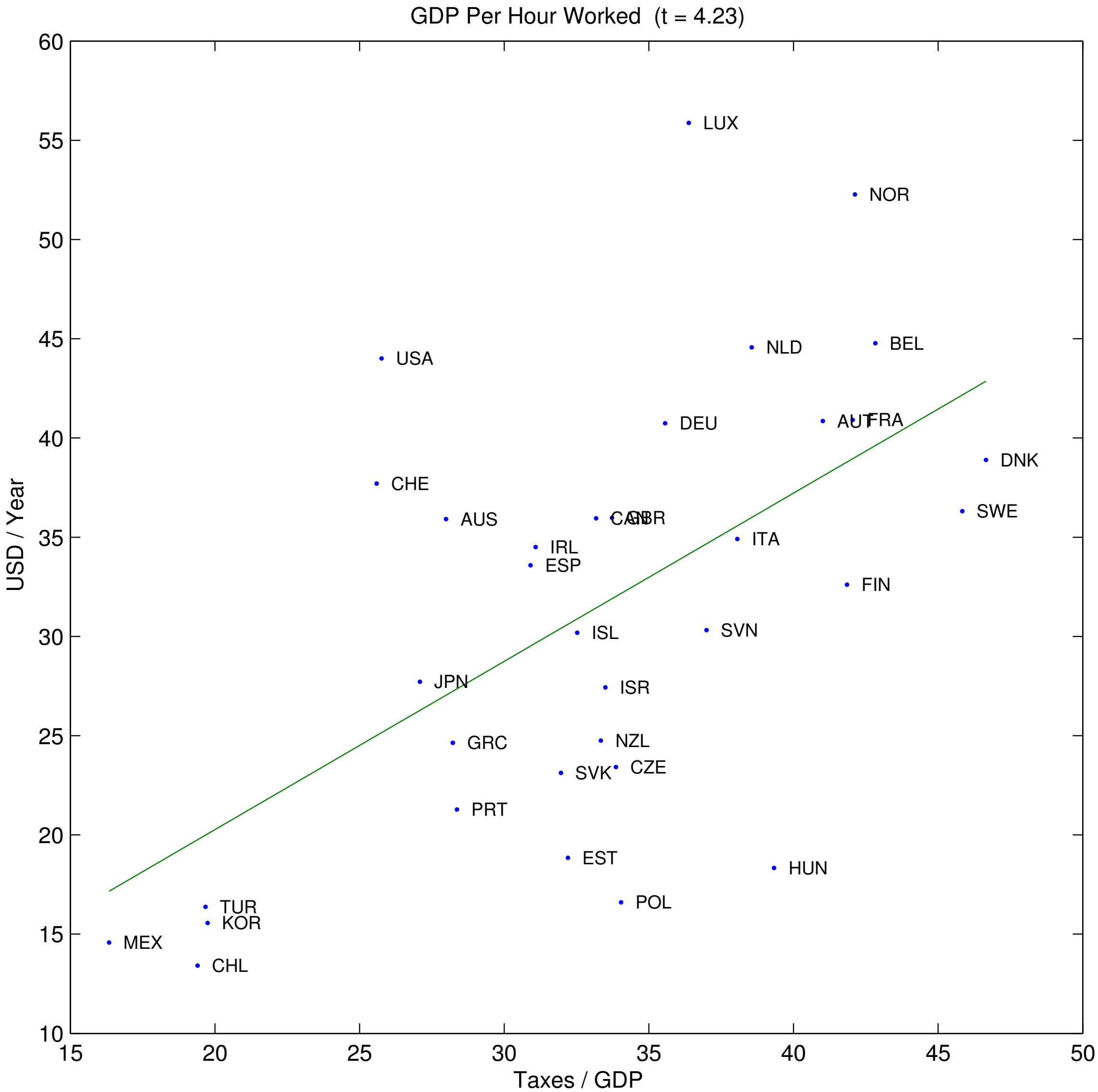



**Figure 13. Productivity vs. taxes.** This figure plots GDP per hour worked ($PROD$) against the tax-to-GDP ratio ($TAX$) across countries. Both variables are computed at the country level as time-series averages in 1980–2013. Each dot, labeled with the OECD country code, is a country-level observation. The solid line is the line of best fit from the linear regression of $PROD$ on $TAX$. The corresponding $t$-statistic is in the panel title.

**Table 1**
**Stock price level vs. taxes**

Each panel reports the results from five cross-country regressions (one per column) of the time-series average of the country's aggregate market-to-book ratio on the time-series averages of the variables given in the row labels. $TAX$ denotes the country's tax-to-GDP ratio, $GDPGRO$ is GDP growth, $GDPPC$ is GDP per capita, and $INFL$ is consumer price inflation. All time-series averages are computed from all available annual data between 1980 and 2013. $t$-statistics are in parentheses. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

| | | | | | |
|---|---|---|---|---|---|
| | | | Panel A. Linear specification. | | |
| $TAX$ | -0.0052 | $-0.0124^{*}$ | -0.0075 | -0.0088 | $-0.0149^{**}$ |
| | (-0.81) | (-1.91) | (-1.10) | (-1.24) | (-2.16) |
| $GDPGRO$ | | $-0.1106^{**}$ | | | $-0.1120^{**}$ |
| | | (-2.55) | | | (-2.48) |
| $INFL$ | | | -0.0042 | | 0.0019 |
| | | | (-0.90) | | (0.38) |
| $GDPPC$ | | | | 0.0000 | 0.0000 |
| | | | | (1.10) | (1.02) |
| Constant | $1.8996^{***}$ | $2.4106^{***}$ | $2.0120^{***}$ | $1.8410^{***}$ | $2.3064^{***}$ |
| | (8.66) | (8.54) | (8.06) | (8.32) | (7.61) |
| Observations | 30 | 30 | 30 | 30 | 30 |
| $R^2$ | 0.02 | 0.20 | 0.05 | 0.06 | 0.22 |
| | | | Panel B. Quadratic specification. | | |
| $TAX$ | $0.1116^{***}$ | $0.0851^{**}$ | $0.1067^{**}$ | $0.1046^{**}$ | $0.0785^{*}$ |
| | (2.68) | (2.08) | (2.53) | (2.39) | (1.83) |
| $TAX^2$ | $-0.0018^{***}$ | $-0.0015^{**}$ | $-0.0018^{***}$ | $-0.0017^{***}$ | $-0.0014^{**}$ |
| | (-2.84) | (-2.41) | (-2.74) | (-2.62) | (-2.21) |
| $GDPGRO$ | | $-0.0857^{**}$ | | | $-0.0883^{**}$ |
| | | (-2.09) | | | (-2.04) |
| $INFL$ | | | -0.0026 | | 0.0011 |
| | | | (-0.62) | | (0.24) |
| $GDPPC$ | | | | 0.0000 | 0.0000 |
| | | | | (0.49) | (0.51) |
| Constant | 0.1578 | 0.8662 | 0.2802 | 0.2156 | 0.9004 |
| | (0.24) | (1.25) | (0.42) | (0.33) | (1.29) |
| Observations | 30 | 30 | 30 | 30 | 30 |
| $R^2$ | 0.23 | 0.33 | 0.24 | 0.23 | 0.33 |

## Table 2
## Stock market returns vs. taxes

Each panel reports the results from five cross-country regressions (one per column) of the time-series average of the country's stock market index return on the time-series averages of the variables given in the row labels. $TAX$ denotes the country's tax-to-GDP ratio, $GDPGRO$ is GDP growth, $GDPPC$ is GDP per capita, and $INFL$ is consumer price inflation. The time-series averages of all independent variables are computed from all available annual data between 1980 and 2013. The average stock market return is computed from all available data, possibly back to 1792. $t$-statistics are in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

| | | | | | |
|---|---|---|---|---|---|
| | Panel A. USD returns, nominal | | | | |
| $TAX$ | −0.5157*** | −0.3119* | −0.3090* | −0.3065* | -0.1329 |
| | (-2.85) | (-1.74) | (-1.86) | (-1.67) | (-0.80) |
| $GDPGRO$ | | 3.0845*** | | | 2.1673** |
| | | (2.76) | | | (2.15) |
| $INFL$ | | | 0.3828*** | | 0.2564** |
| | | | (3.45) | | (2.22) |
| $GDPPC$ | | | | −0.0004** | -0.0002 |
| | | | | (-2.59) | (-1.30) |
| Observations | 33 | 33 | 33 | 33 | 33 |
| $R^2$ | 0.20 | 0.35 | 0.41 | 0.33 | 0.51 |
| | Panel B. Local currency returns, nominal | | | | |
| $TAX$ | −0.9235*** | −0.7110** | −0.5213** | −0.6285** | −0.4041* |
| | (-3.45) | (-2.54) | (-2.46) | (-2.29) | (-1.78) |
| $GDPGRO$ | | 3.2160* | | | 1.3523 |
| | | (1.84) | | | (0.98) |
| $INFL$ | | | 0.7447*** | | 0.6584*** |
| | | | (5.27) | | (4.17) |
| $GDPPC$ | | | | −0.0006** | -0.0001 |
| | | | | (-2.45) | (-0.70) |
| Observations | 33 | 33 | 33 | 33 | 33 |
| $R^2$ | 0.27 | 0.33 | 0.60 | 0.38 | 0.62 |
| | Panel C. Local currency returns, real | | | | |
| $TAX$ | −0.2856*** | −0.2143** | −0.2148** | −0.2234** | -0.1604 |
| | (-2.92) | (-2.08) | (-2.16) | (-2.10) | (-1.51) |
| $GDPGRO$ | | 1.0788* | | | 0.7794 |
| | | (1.67) | | | (1.21) |
| $INFL$ | | | 0.1310** | | 0.0946 |
| | | | (1.97) | | (1.28) |
| $GDPPC$ | | | | -0.0001 | -0.0000 |
| | | | | (-1.33) | (-0.45) |
| Observations | 33 | 33 | 33 | 33 | 33 |
| $R^2$ | 0.20 | 0.27 | 0.29 | 0.25 | 0.33 |

**Table 3**
**Income inequality vs. taxes**

Each panel reports the results from five cross-country regressions (one per column) of the time-series average of the country's share of income going to the top 10% on the time-series averages of the variables given in the row labels. $TAX$ denotes the country's tax-to-GDP ratio, $GDPGRO$ is GDP growth, $GDPPC$ is GDP per capita, and $INFL$ is consumer price inflation. All time-series averages are computed from all available annual data between 1980 and 2013. $t$-statistics are in parentheses. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

| | | | | | |
|---|---|---|---|---|---|
| | | | Panel A. Linear specification. | | |
| $TAX$ | $-0.4291^{***}$ | $-0.4943^{***}$ | $-0.4496^{***}$ | $-0.5554^{***}$ | $-0.5414^{***}$ |
| | (-3.53) | (-3.68) | (-3.96) | (-4.74) | (-4.46) |
| $GDPGRO$ | | -0.8195 | | | 0.0401 |
| | | (-1.07) | | | (0.05) |
| $INFL$ | | | $-0.0965^{**}$ | | -0.0592 |
| | | | (-2.17) | | (-1.32) |
| $GDPPC$ | | | | $0.0003^{***}$ | $0.0003^{**}$ |
| | | | | (2.80) | (2.03) |
| Constant | $44.2400^{***}$ | $48.5272^{***}$ | $46.0232^{***}$ | $40.9176^{***}$ | $42.4181^{***}$ |
| | (10.67) | (8.48) | (11.68) | (10.56) | (7.24) |
| Observations | 30 | 30 | 30 | 30 | 30 |
| $R^2$ | 0.29 | 0.32 | 0.39 | 0.44 | 0.47 |
| | | | Panel B. Quadratic specification. | | |
| $TAX$ | -1.1576 | $-1.6419^{*}$ | -1.0945 | $-1.4678^{*}$ | $-1.4594^{*}$ |
| | (-1.25) | (-1.71) | (-1.27) | (-1.79) | (-1.71) |
| $TAX^2$ | 0.0111 | 0.0171 | 0.0098 | 0.0138 | 0.0137 |
| | (0.80) | (1.21) | (0.76) | (1.12) | (1.08) |
| $GDPGRO$ | | -1.1083 | | | -0.2149 |
| | | (-1.41) | | | (-0.28) |
| $INFL$ | | | $-0.0950^{**}$ | | -0.0539 |
| | | | (-2.16) | | (-1.21) |
| $GDPPC$ | | | | $0.0003^{***}$ | $0.0003^{**}$ |
| | | | | (2.94) | (2.05) |
| Constant | $55.6051^{***}$ | $67.5821^{***}$ | $56.0600^{***}$ | $54.9936^{***}$ | $57.7052^{***}$ |
| | (3.75) | (4.04) | (4.06) | (4.20) | (3.79) |
| Observations | 30 | 30 | 30 | 30 | 30 |
| $R^2$ | 0.31 | 0.35 | 0.40 | 0.46 | 0.49 |

**Table 4**
**Productivity vs. taxes**

This table reports the results from five cross-country regressions (one per column) of the time-series average of the country's GDP per hour worked on the time-series averages of the variables given in the row labels. $TAX$ denotes the country's tax-to-GDP ratio, $GDPGRO$ is GDP growth, $GDPPC$ is GDP per capita, and $INFL$ is consumer price inflation. All time-series averages are computed from all available annual data between 1980 and 2013. $t$-statistics are in parentheses. $^{*}$ $p < 0.1$, $^{**}$ $p < 0.05$, $^{***}$ $p < 0.01$.

| | | | | | |
|---|---|---|---|---|---|
| $TAX$ | 0.8478*** | 0.7250*** | 0.7898*** | 0.3320*** | 0.3052*** |
| | (4.23) | (3.39) | (4.10) | (3.98) | (3.53) |
| $GDPGRO$ | | -1.8689 | | | -0.4969 |
| | | (-1.39) | | | (-0.95) |
| $INFL$ | | | −0.1559* | | 0.0207 |
| | | | (-1.91) | | (0.60) |
| $GDPPC$ | | | | 0.0010*** | 0.0010*** |
| | | | | (14.39) | (13.53) |
| Constant | 3.3114 | 12.2592 | 7.0245 | −4.7179* | -2.8613 |
| | (0.49) | (1.33) | (1.04) | (-1.81) | (-0.77) |
| Observations | 34 | 34 | 34 | 34 | 34 |
| $R^2$ | 0.35 | 0.38 | 0.41 | 0.91 | 0.91 |

## REFERENCES


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

Atkinson, Anthony B., Thomas Piketty, and Emmanuel Saez, 2011, Top incomes in the long run of history, *Journal of Economic Literature* 49, 3–71.

Autor, David H., Lawrence F. Katz, and Melissa S. Kearney, 2008, Trends in US wage inequality: Revising the revisionists, *Review of Economics and Statistics* 90, 300–323.

Bhamra, Harjoat S., and Raman Uppal, 2014, Asset prices with heterogeneity in preferences and beliefs, *Review of Financial Studies* 27, 519–580.

Constantinides, George M., and Darrell Duffie, 1996, Asset pricing with heterogeneous consumers, *Journal of Political Economy* 104, 219–240.

Croce, Maximiliano M., Howard Kung, Thien T. Nguyen, and Lukas Schmid, 2012, Fiscal policies and asset prices, *Review of Financial Studies* 25, 2635–2672.

Dumas, Bernard, 1989, Two-person dynamic equilibrium in the capital market, *Review of Financial Studies* 2, 157–188.

Eaton, Jonathan, and Harvey S. Rosen, 1980, Taxation, human capital, and uncertainty, *American Economic Review* 70, 705–715.

Favilukis, Jack, 2013, Inequality, stock market participation, and the equity premium, *Journal of Financial Economics* 107, 740–759.

Fischer, Marcel, and Bjarne Astrup Jensen, 2015, Taxation, transfer income and stock market participation, *Review of Finance* 19, 823–863.

Garleanu, Nicolae, and Stavros Panageas, 2015, Young, old, conservative, and bold: The implications of heterogeneity and finite lives for asset pricing, *Journal of Political Economy* 123, 670–685.

Gollier, Christian, 2001, Wealth inequality and asset pricing, *Review of Economic Studies* 68, 181–203.

Heaton, John, and Deborah Lucas, 1996, Evaluating the effects of incomplete markets on risk sharing and asset pricing, *Journal of Political Economy* 104, 443–487.

Heaton, John, and Deborah Lucas, 2000, Portfolio choice and asset prices: The importance of entrepreneurial risk, *Journal of Finance* 55, 1163–1198.

Hombert, Johan, Antoinette Schoar, David Sraer, and David Thesmar, 2014, Can unemployment insurance spur entrepreneurial activity?, NBER working paper.

Johnson, Timothy C., 2012, Inequality risk premia, *Journal of Monetary Economics* 59, 565–580.

Kacperczyk, Marcin, Jaromir B. Nosal, and Luminita Stevens, 2015, Investor sophistication and capital income inequality, Working paper, Imperial College.

Kelly, Bryan, Ľuboš Pástor, and Pietro Veronesi, 2015, The price of political uncertainty: Theory and evidence from the option market, *Journal of Finance*, forthcoming.

Kihlstrom, Richard E., and Jean-Jacques Laffont, 1979, A general equilibrium entrepreneurial theory of firm formation based on risk aversion, *Journal of Political Economy* 87, 719–748.

Lucas Jr, Robert E., 1978, On the size distribution of business firms, *Bell Journal of Economics* 9, 508–523.

Mirrlees, James A., 1971, An exploration in the theory of optimum income taxation, *Review of Economic Studies* 38, 175–208.

Pástor, Ľuboš, and Pietro Veronesi, 2003, Stock valuation and learning about profitability, *Journal of Finance* 58, 1749–1789.

Pástor, Ľuboš, and Pietro Veronesi, 2006, Was there a Nasdaq bubble in the late 1990s? *Journal of Financial Economics* 81, 61–100.

Pástor, Ľuboš, and Pietro Veronesi, 2012, Uncertainty about government policy and stock prices, *Journal of Finance* 67, 1219–1264.

Pástor, Ľuboš, and Pietro Veronesi, 2013, Political uncertainty and risk premia, *Journal of Financial Economics* 110, 520–545.

Piketty, Thomas, 2014, *Capital in the 21st Century,* Harvard University Press, Cambridge.

Varian, Hal R., 1980, Redistributive taxation as social insurance, *Journal of Public Economics* 14, 49–68.