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 presented at the November 2015 Carnegie-Rochester-NYU Conference on Public Policy. The views in this paper are the responsibility of the authors, not the institutions they are affiliated with. We are grateful to the Brevan Howard Centre for Financial Analysis at Imperial College London, which we visited while writing this paper. For helpful comments, we are grateful to our discussant Philippe Mueller, our editors Marvin Goodfriend and Burton Hollifield, Doug Diamond, John Heaton, Marcin Kacperczyk, Stavros Panageas, Jacopo Ponticelli, Raman Uppal, Adrien Verdelhan, the audience at the 2016 Jackson Hole Finance Conference, and seminar participants at Bocconi, Chicago, Imperial, LBS, LSE, MIT, Torino, and York. We are also grateful to Menaka Hampole, Oscar Eskin, and Melissa Ortiz for excellent research assistance and to the Fama-Miller Center for Research in Finance and the Center for Research in Security Prices, both at Chicago Booth, for research support. 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, Revised January 2016
JEL No. E24,G1,H2,J24,J31,J38

ABSTRACT

Our simple model features agents heterogeneous in skill and risk aversion, incomplete financial markets, and redistributive taxation. 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 equity risk premium. Both income inequality and 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 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 those 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 also affects aggregate productivity and asset prices, which, in equilibrium, feed back into inequality.

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

[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.

skilled and less risk-averse, on average, and total output is lower.

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 smaller than consumption inequality, in part because pensioners not only 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 asset price pattern feeds back into income inequality through the investment risk component, contributing to its hump-shaped pattern.

Last but not least, the model implies a positive relation between the tax rate and aggre-

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

gate productivity. The reason, again, is selection: a higher tax rate implies that those who create value in the economy—entrepreneurs—are more skilled. At the same time, heavier taxation implies fewer entrepreneurs and more pensioners, and thus lower total output. In short, a higher tax rate makes the economy smaller but more productive.

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 it does not exhibit concavity. The relation between the average stock market return and the tax burden is negative, as predicted, but not always significant. The relation between the level of stock prices and the tax burden is concave and negative, as predicted, but the negativity is significant only after the inclusion of 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

[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.

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 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. reports the empirical

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

[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. (2016).

results. Section 5. 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.

2. Model

There is a continuum of agents with unit mass. Each agent i is endowed with a skill level μ_i, risk aversion γ_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's capital endowment is technology-specific; it can be invested only in a production technology that requires this agent's skill to operate. Specifically, each agent i can invest B_0 in a constant-return-to-scale technology that produces $B_{i,T}$ units of output at a given future time T:

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

where ε_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, μ_i, is therefore equivalent 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 θ of his own firm. Due to this friction, markets are incomplete.

[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.

The second type of agents, pensioners, do not invest; instead, they live off taxes paid by entrepreneurs. We interpret pensioners as including not only retirees but also anyone collecting income from the government without making a direct contribution to total output, such as government workers (whose contribution is indirect), people on disability, etc. Because they do not invest, agents who choose to become pensioners effectively abandon their technology-specific capital endowments in exchange for future tax-financed pensions. 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. If we interpret B_0 as the capacity to put in a certain amount of labor, we can easily 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 μ_i.) Second, by becoming pensioners, agents give up 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 τ. 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. At the time they make this choice, agents know all the parameters of the model (i.e., μ_i, γ_i, τ, θ, and B_0); the only thing they do not know is the future realizations of ε_T and $\varepsilon_{i,T}$. Let $\mathcal{I}$ denote the set of agents who decide to invest and become entrepreneurs. 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 output at time T. Since only entrepreneurs engage in production, total output is given by

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

so that total tax revenue is τB_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 τ 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 ε_T because tax revenue depends on ε_T. This result illustrates the limits of consumption smoothing by redistribution.[14]

[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 μ_i and γ_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.

2.1.2. Entrepreneurs' Consumption

Entrepreneur i's firm pays a single dividend $B_{i,T}$, given in equation (1). The fraction θ 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 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 μ_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:

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

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]

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

$$R^{Mkt} = \frac{(1-\tau)B_{T}}{M_{0}^{P}} - 1\ . \qquad (17)$$

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 μ_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

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

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 γ_i, in equilibrium we have more risk-averse entrepreneurs lending to less risk-averse ones.[16]

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

[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 γ_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.

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*

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

Equation (19) shows that only agents who are sufficiently skilled—those with sufficiently high μ_i—become entrepreneurs. This statement holds other things, especially γ_i and $\mathcal{I}$, equal. Note that μ_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 τ, holding $\mathcal{I}$ constant. A higher τ thus discourages entrepreneurship by raising the hurdle for μ_i above which agents become entrepreneurs. Moreover, a higher τ implies a higher average value of μ_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 γ_i in a non-linear fashion. For many parametric assumptions, though, the right-hand side is increasing in γ_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 γ_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 γ_i. The reason is that when $\theta = 0$, both types of agents are exposed only to aggregate risk ε_T. Entrepreneurs with high γ_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 γ_i's become entrepreneurs because doing so allows them to choose low $\alpha(\gamma_i)$ and thus face less risk than 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 θ 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 ($m(\mathcal{I}) = 1$), though, 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.

[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.

[18] See, for example, Heaton and Lucas (2000) and Moskowitz and Vissing-Jorgensen (2002).

Proposition 4: *The equilibrium state price density π_T is given by*

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

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

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

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

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

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 π_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 probability distributions of ε_T, $\varepsilon_{i,T}$, μ_i, and γ_i are such that the equilibrium conditions are well defined, entrepreneurs cannot default on short positions, if any, and the fixed-point system has a solution. We prove the existence of a solution in the three special cases discussed in Section 3. 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 π_T depends on the full distribution of γ_i across entrepreneurs.

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

$$\mathrm{E}\left(R^i\right) = r \ , \tag{25}$$

where

$$r = \frac{1}{Z} - 1 \ . \tag{26}$$

Recall from Proposition 4 that Z is the equilibrium price at time 0 of a security that pays e^{ε_T} at time T. Because $\mathrm{E}(e^{\varepsilon_T}) = 1$, the expected return of this security is given by r in equation (26). Proposition 5 shows that r is also the expected return on any stock i. All stocks have the same expected return because they all 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 equal to r.

The expected return depends on the tax rate τ through the selection effect of τ 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 τ, as noted earlier. If it is also increasing in γ_i, which seems realistic (see our discussion of Proposition 3), then an increase in τ leads more high-γ_i agents to become pensioners. A higher τ 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 the risk adjustment. There is no discounting beyond this risk adjustment; as noted earlier, we use the risk-free bond as numeraire, thereby effectively setting the risk-free rate to zero. In computing M/B, we scale market value by the amount of initial capital whose natural interpretation is the book value of the firm.

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 τ is ambiguous. On the one hand, a higher τ reduces M/B by reducing the after-tax cash flow through the $(1-\tau)$ term. On the other hand, a higher τ 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 τ 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 cross-sectional variance of $s_{i,T}$, computed 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 cross-sectional variance of consumption at time T is given by*

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

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 $\alpha\left(\gamma_j\right)$. Even if all idiosyncratic risk could be diversified away (i.e., $\theta = 0$), cross-sectional heterogeneity in γ_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:

$$s_{i,T} = \frac{1-\tau}{m(\mathcal{I})} \times \frac{e^{\mu_i T}}{\mathrm{E}^{\mathcal{I}}[e^{\mu_j T}|j \in \mathcal{I}]} \times \frac{1+\theta R^i + (1-\theta)\,\alpha(\gamma_i) R^{Mkt}}{1+R^{Mkt}} \qquad \text{for } i \in \mathcal{I} \tag{32}$$

$$= \frac{\tau}{1-m(\mathcal{I})} \qquad \text{for } i \notin \mathcal{I}\ . \tag{33}$$

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 cross-sectional 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 $s_{i,T}$,

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

After imposing more 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 this limiting case, the stock market capitalization is infinitesimally small and the stock prices we calculate are shadow prices. Yet this case is particularly interesting because it

[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.

allows a clean solution while preserving all other relevant aspects of the problem, including both dimensions of agent heterogeneity and both dimensions of risk.

The special case of $\theta \to 1$ is analyzed throughout this section except for Section 3.6., in which we consider two other special cases: homogeneous risk aversion and no systematic risk. Both cases yield analytical results and admit proofs of the existence of the equilibrium even when $\theta < 1$. But until Section 3.6., we consider the general case from Section 2., with heterogeneous risk aversion and systematic risk, and assume $\theta \to 1$ instead.

We also 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

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

The selection effects mentioned earlier are now particularly easy to see: agents with higher skill (μ_i) and lower risk aversion (γ_i) are more likely to become entrepreneurs. The right-hand side of equation (38), which is independent of i, is intuitive. 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, which is unaffected by agent i's choice because any given agent is infinitesimally small. A higher value implies a higher expected tax base and thus a higher expected consumption for pensioners, which again indicates a larger opportunity cost to being an entrepreneur. Agent i becomes an entrepreneur only if his μ_i is high enough and γ_i is low enough to overcome these aggregate effects.

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

$$\mu_i \sim N\left(\overline{\mu}, \sigma_\mu^2\right) \tag{39}$$

$$\gamma_i \sim N\left(\overline{\gamma}, \sigma_\gamma^2\right) 1_{\{\gamma_i > 0\}} . \tag{40}$$

That is, skill μ_i is normally distributed with mean $\overline{\mu}$ and variance σ_μ^2. Risk aversion γ_i is truncated normal, with truncation at zero and underlying normal distribution with mean $\overline{\gamma}$

and variance σ_γ^2. Given these distributional assumptions, we solve for the equilibrium mass of entrepreneurs $m(\mathcal{I})$. We prove that

$$\frac{\partial m(\mathcal{I})}{\partial \tau} < 0 \, , \tag{41}$$

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 incentivizes agents 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), (39), and (40): $\sigma = 10\%$ 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 (γ_i, μ_i) represents an agent with skill μ_i and risk aversion γ_i. The circular contours outline the joint probability density of μ_i and γ_i across agents, indicating regions containing 50%, 90%, 99%, and 99.9% of the probability mass. The threshold lines correspond to the tax rates τ of 0.1%, 5%, 20%, and 70%. For a given τ, all agents located above the threshold line choose to 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 μ_i and γ_i is also clear from equation (38).

The figure also shows that higher taxes discourage entrepreneurship: as τ rises, the threshold line shifts upward, shrinking the region of entrepreneurs. This effect is much more dramatic for low tax rates: raising τ from 0.1% to 5% reduces the region by more than raising τ from 20% to 70%. When $\tau = 0$, nobody becomes a pensioner because there is no tax revenue for pensioners to consume. When τ 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 τ 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 τ high enough (the Laffer curve).

Supporting these arguments, Figure 2 shows that $m(\mathcal{I})$ declines with τ 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 τ is near zero, but their consumption grows with τ. Entrepreneurs consume more than pensioners on average for any τ, in part due to higher skill and in part due to compensation for risk. Interestingly, the spread between the two consumption levels widens as τ rises. The reason is that as τ 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 τ. However, when we calculate 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 τ 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 average consumption levels is large (see Figure 2), this difference does not contribute much to total variance since almost all agents are entrepreneurs. When τ rises, the component initially rises, for two reasons. First, the difference between the average consumption levels grows with τ, as discussed in the previous paragraph. Second, the mass of pensioners grows as well, making this difference more important. But when τ 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 τ 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 τ, this is the smallest of the three components. The component declines when τ rises because the rising threshold for μ_i reduces the heterogeneity in μ_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

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

the entrepreneurs' investments. This investment risk component, driven by pure luck, is the largest source of inequality for any τ.[21] The component initially rises because a higher τ 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 τ because the mass of entrepreneurs shrinks. That is, when the tax rate is high, investment risk does not matter much because there is little investment.

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 $C_{i,T}$:

$$\frac{(CE_{i,T})^{1-\gamma_i}}{1-\gamma_i} = \mathrm{E}\left[\frac{C_{i,T}^{1-\gamma_i}}{1-\gamma_i}\right] . \tag{42}$$

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

$$\begin{aligned} CE_{i,T} &= B_0\,(1-\tau)\,e^{\mu_i T} e^{-\frac{1}{2}\gamma_i\left(\sigma^2+\sigma_1^2\right)T} && \text{for } i \in \mathcal{I} && (43)\\ &= B_0\,\tau\,\frac{m\,(\mathcal{I})}{1-m\,(\mathcal{I})}\,\mathrm{E}^{\mathcal{I}}\left[e^{\mu_j T} | j \in \mathcal{I}\right] e^{-\frac{1}{2}\gamma_i \sigma^2 T} && \text{for } i \notin \mathcal{I} . && (44) \end{aligned}$$

Since pensioners do not employ their skill, their $CE_{i,T}$'s do not depend on μ_i, but they do depend on γ_i because pensioners face aggregate risk. Entrepreneurs' $CE_{i,T}$'s depend on both μ_i and γ_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): $s_{i,T}^{CE} = \frac{CE_{i,T}}{\int CE_{i,T} di}$. We then calculate the variance of $s_{i,T}^{CE}$ across agents, our measure of inequality in expected utility, and plot it against τ.

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 τ. Heavier taxation thus implies less dispersion in ex-ante happiness. However, the average $CE_{i,T}$ across all agents (dashed line) also decreases with τ because

[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.

higher τ implies less investment and thus a lower expected total output (dash-dot line). In other words, a higher τ makes agents more equal in utility terms, but it also makes the average agent worse off. As τ rises toward one, all agents become equally unhappy.

The distribution of realized consumption is right-skewed across agents, for two reasons. First, consumption is right-skewed among entrepreneurs, due to its convexity in μ_i, ε_T, and $\varepsilon_{i,T}$ (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. Certainty equivalent consumption is also right-skewed, due to convexity in μ_i, but less so than realized consumption due to the absence of convexity in random shocks (see equation (43)). We plot both distributions in the Internet Appendix.

Panel A of Figure 5 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 τ, but its peak occurs earlier so its relation to τ is largely negative.

3.3. Productivity

Panel B of Figure 5 plots expected aggregate productivity, computed as the annualized expected growth rate of total capital, or $(1/T)\mathrm{E}\left[B_T/(m(\mathcal{I})B_0) - 1\right]$, against τ. Productivity increases with τ due to the selection effect described earlier: a higher τ implies a higher average level of skill among entrepreneurs. When τ 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. In other words, heavier taxation implies lower total output but higher 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 output per hour worked, which is also the measure of productivity that we use in our empirical analysis.

3.4. Asset Prices

Panel C of Figure 5 plots the expected return on the market portfolio, annualized, as a function of τ. The expected return falls as τ rises, due to selection: a higher τ implies that entrepreneurs are less risk-averse, on average. Given their lower risk aversion, agents demand a lower risk premium to hold stocks, resulting in a lower expected market return.

Panel D of Figure 5 plots the level of stock prices, measured by the market portfolio's M/B ratio, as a function of τ. M/B exhibits a concave and mostly negative relation to τ: it increases with τ until $\tau = 15\%$ but then it decreases. This nonlinear pattern results from the interaction of three effects. On the one hand, a higher τ 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 τ implies that entrepreneurs are more skilled, on average, pushing up the average firm's expected cash flow. Second, a higher τ 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 τ, as shown in Figure 1, but the direct effect prevails eventually.

3.5. Varying Heterogeneity in Skill and Risk Aversion

The patterns in Figure 5 are robust to changes in σ_μ and σ_γ. When we vary these parameters around their baseline values, we observe some level shifts but the patterns remain very similar to those in Figure 5: hump shapes in Panels A and D, growth in Panel B, and decline in Panel C. We summarize the results here and show the plots in the Internet Appendix.

Higher values of σ_μ raise the values of all four variables from Figure 5. A higher σ_μ implies more dispersion in skill and, consequently, more inequality. A higher σ_μ also raises expected aggregate 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] A higher σ_μ lifts the expected return because it strengthens the importance of μ_i at the expense of γ_i in the entrepreneur selection mechanism. As a result of the weaker selection on γ_i, a higher σ_μ implies a higher average γ_i among entrepreneurs, which pushes up the expected return. Finally, an increase in σ_μ raises M/B by increasing expected cash flow by enough to overcome the expected return effect.

The effect of σ_γ on inequality is small and that on the expected return is parameter-dependent. But an increase in σ_γ reduces productivity because it strengthens the importance

[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.

of γ_i at the expense of μ_i in the selection of entrepreneurs. Due to the weaker selection on μ_i, an increase in σ_γ reduces the average μ_i among entrepreneurs, thereby depressing productivity. Since lower productivity means lower expected cash flow, a higher σ_γ also reduces M/B. While the effects of σ_μ and σ_γ are interesting, we focus on the effects of τ.

3.6. Other Special Cases

Sections 3.1. through 3.5. analyze the special case of our model in which θ, the share of firm i that must be retained by entrepreneur i, approaches one (i.e., $\theta \to 1$). In this section, we consider two other special cases for which we can derive analytical results and prove the existence of the equilibrium: homogeneous risk aversion (Section 3.6.1.) and no systematic risk (Section 3.6.2.). In both cases, we no longer assume $\theta \to 1$; instead, we consider any $\theta \in (0,1)$ and analyze the effect of θ on quantities of interest. While both of these special cases are instructive, they have less interesting asset pricing implications than the case of $\theta \to 1$. Discount rate effects are absent from both cases—homogeneity in risk aversion takes away selection on γ_i and the exclusion of systematic risk makes the aggregate stock market risk-free. To save space, we only summarize the results for these two special cases in this section. A formal analysis of both cases is in the Internet Appendix.

3.6.1. Special Case: Common Risk Aversion

In this section, we assume that all agents have the same risk aversion: $\gamma_i = \gamma$. To obtain closed-form solutions, we also make the distributional assumptions (36), (37), and (39).

With common risk aversion, we obtain $\alpha(\gamma_i) = \alpha$, 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: θ in their own firm and $1-\theta$ in the stock market.

In equilibrium, agent i becomes an entrepreneur if and only if he is sufficiently skilled:

$$\mu_i > \underline{\mu} \,, \tag{45}$$

where $\underline{\mu}$ is a given cutoff. The mass of entrepreneurs, $m(\mathcal{I})$, decreases with the tax rate as well as with θ. Specifically, we prove that $\partial\underline{\mu}/\partial\tau > 0$ and $\partial\underline{\mu}/\partial\theta > 0$, so that

$$\frac{\partial m(\mathcal{I})}{\partial \tau} < 0 \tag{46}$$

$$\frac{\partial m(\mathcal{I})}{\partial \theta} < 0 \,. \tag{47}$$

The first result is explained earlier in the context of equation (41). The second result is also intuitive: a higher θ makes entrepreneurship less appealing because it increases each entrepreneur's exposure to idiosyncratic risk.

The implications for inequality follow from a closed-form solution for the cross-sectional variance of consumption. This variance is increasing with the cross-sectional dispersion in skill, σ_μ. The variance is also increasing with the product $\theta^2(e^{\sigma_1^2 T} - 1)$, which captures the contribution of investment risk to inequality. A higher value of σ_1 indicates higher volatility of idiosyncratic shocks, and a higher value of θ implies a larger role for those shocks. The positive effect of θ on inequality is present as long as there is idiosyncratic risk (i.e., $\sigma_1^2 > 0$), and the positive effect of σ_1 is present unless that risk is fully diversifiable (i.e., $\theta > 0$).

Solving for asset prices is straightforward because the state price density π_T from equation (20) simplifies dramatically, becoming proportional to $e^{-\gamma \varepsilon_T}$. Interestingly, the stochastic discount factor is independent of θ, unlike in the general case in Section 2. Since all firms have the same risk exposure (equation (1)), all entrepreneurs' positions are symmetric ex ante. Therefore, the risk aversion in the economy is the common risk aversion γ and stock prices are unaffected by θ. The M/B ratio of firm i is given by

$$\frac{M_{i,0}}{B_0} = (1-\tau)\, e^{\left(\mu_i - \gamma\sigma^2\right)T}, \tag{48}$$

which is the expected after-tax cash flow adjusted for risk. The risk adjustment is simple because all expected returns are equal to $e^{\gamma\sigma^2 T} - 1$. We obtain a closed-form solution for the aggregate M/B ratio by averaging equation (48) across all entrepreneurs. This M/B ratio is increasing in σ_μ due to the selection effect from equation (45) as well as the convexity effect discussed earlier. Due to the same selection effect, the effects of taxation on stock prices and productivity are similar as before, except that there are no discount rate effects.

3.6.2. Special Case: No Systematic Risk

We now assume that there are no systematic shocks: $\varepsilon_T = 0$ in equation (1). Since the average of all idiosyncratic shocks is zero, the aggregate stock market portfolio is risk-free. All entrepreneurs invest $\alpha(\gamma_i) = 1$ in the stock market and nothing in bonds.

Assuming that μ_i and γ_i are independently distributed across agents, we prove that a unique equilibrium exists in this economy. The independence assumption is sufficient but not necessary. The screen for entrepreneurship is a simpler version of Proposition 3, implying selection effects based on both μ_i and γ_i. While the selection on μ_i has the same effects as before, the selection on γ_i has no discount rate effects because the discount rate is always zero: stock prices are given by Proposition 6 with $r = 0$. Relations (46) and (47) hold in this case as well, indicating that the mass of entrepreneurs decreases with both τ and θ. The formula for consumption variance is a simpler version of equation (31), indicating a role for θ similar to that in Section 3.6.1. Besides all these results, this case also allows us to prove that the right-hand-side of condition (19) is increasing in γ_i and that entrepreneurs consume more than pensioners on average, as noted earlier (see footnote 19).

4. Empirical Analysis

In this section, we examine the model's predictions empirically. To preview the results, the evidence is broadly consistent with the model.

4.1. Data and Variable Definitions

We collect country-level annual data for all 34 members of the Organization for Economic Co-operation and Development (OECD). Our tax burden variable is denoted by TAX. The value of TAX in a given year is the ratio of total government tax revenue in that year 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, the top 10% share 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.

Aggregate stock market index returns, RET, come from Global Financial Data (GFD). We download nominal returns from GFD and convert them 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 the Internet Appendix.

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.

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 1995 for Poland. Since stock returns are notoriously volatile and roughly 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, given that 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 in the 1970s and 80s for most countries, though for some countries they begin much earlier. 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.

4.2. Empirical Results

Our theory makes predictions about the effects of taxes on inequality, productivity, stock prices, and returns. 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. Some randomness is surely present because a country's tax burden reflects the country's traditions and cultural values, which are exogenous to a large extent. Even though the tax burdens are unlikely to be fully exogenous, our empirical analysis seems relevant as it examines the key econometric associations predicted by the model. While we present only correlations, we interpret them through the model.

Panel A of Figure 6 plots our measure of income inequality, TOP, 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 TOP on average TAX (solid line) and the other from the quadratic regression of TOP on TAX and TAX^2 (dashed line).[24] These lines indicate a negative and approximately linear relation between TOP and TAX.

Table 1 shows that the negative relation is statistically significant, with the t-statistic of -3.53. The relation becomes even stronger ($t = -4.46$) after including the macroeconomic controls introduced earlier, $GDPGRO$, $INFL$, and $GDPPC$. A negative relation between TOP and TAX is consistent with the model. Table 2 shows no significant convexity or concavity in this relation. While the model predicts concavity, its lack in the data does not necessarily violate the model because the concavity is driven by very low tax rates that are rarely observed in the data (see Panel A of Figure 5). We reach the same conclusions when we use the Gini coefficient in place of TOP.

Panel B of Figure 6 shows a strong positive relation between productivity, $PROD$, and TAX ($t = 4.23$; see Table 1), as predicted by the model. This relation survives the inclusion of the controls. The most important control is GDP per capita ($GDPPC$), which enters with a highly significant positive coefficient. This is not surprising 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 ($t = 3.53$).

Panel C of Figure 6 shows a negative relation between the average stock market return, RET, and TAX ($t = -2.92$), as predicted by the model. Stock returns are in local currency terms and adjusted for inflation.[25] The negative relation weakens after adding the three controls ($t = -1.51$), as shown in Table 1. Of course, with only 33 observations, regressions of average returns on 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.

[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 TOP and TAX refer to a country's time-series averages of these variables.

[25]The results for U.S. dollar returns and nominal returns are very similar. See the Internet Appendix.

Panel D of Figure 6 shows a negative and concave relation between the level of stock prices, M/B, and TAX, as predicted by the model. While the concavity is statistically significant ($t = -2.84$), the negativity is not ($t = -0.81$). The addition of the three controls makes the negative relation statistically significant ($t = -2.16$). 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]

We conduct various robustness tests. We consider two measures of income inequality and three measures of stock market returns, as noted 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. For TOP, the point estimate of the slope on TAX is negative in each year since 1971, and it is significantly negative ever since 1985. The relation between $PROD$ and TAX is significantly positive in each year since 1971. For RET, the point estimates of the slope on TAX are negative in every year between 1965 and 2013, and they are statistically significant ever since 1980. For M/B, the slope estimate is negative in 27 of the 33 years, and five of the remaining six years occur around year 2000, in which stock valuations as measured by M/B were unusually high (the Internet "bubble"). In short, the relations of TAX to TOP, $PROD$, RET, and M/B are 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.

5. Conclusions

Our model sheds new light on the effects of redistributive taxation. 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

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

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 who can invest their wages in financial markets. The model's implications would depend on the assumptions about the wage contract and the role of workers in the production function, but we believe that our main selection effects would remain present. If all agents derive the same disutility from working, we expect agents to self-select so that workers are more skilled than pensioners (e.g., Meltzer and Richard, 1981). And if wages are subject to enough idiosyncratic risk (e.g., through bonuses, stock options, or job risk), we expect workers to be less risk-averse than pensioners. We thus expect pensioners to be less skilled and more risk-averse than both workers and entrepreneurs. Since pensioners are the only agents excluded from both production and financial markets, we expect both selection effects to operate in ways similar to our current framework. Future work can verify these conjectures and, in addition, examine the model's labor market implications.

Another interesting extension would endogenize the tax rate, possibly allowing for progressive taxation. To do so, one would need to write down the social welfare function and find the tax rate, or schedule, that maximizes it. The optimal tax rate will in general depend on the parameters of the social welfare function, on the distribution of skill and risk aversion in the economy, and on the volatilities of the aggregate and idiosyncratic shocks.

Yet another extension could add heterogeneity in the riskiness of agents' technologies. This third dimension of heterogeneity would significantly complicate the equilibrium, but we have some conjectures about the solution. Given imperfect risk-sharing, agents with riskier technologies would be less likely to become entrepreneurs. A higher tax rate would select agents with safer technologies into entrepreneurship, resulting in less risky firms and thus a lower expected market return. This negative relation between the expected return and the tax rate would be similar to the negative relation found in our model, except that the latter relation is driven by the price of risk whereas the former would be driven by the quantity of risk. We leave a more careful examination for future research.

REFERENCES

Alvaredo, F., Atkinson, A.B., Piketty, T., Saez, E., 2013. The top 1 percent in international and historical perspective. Journal of Economic Perspectives 27, 3–20.

Atkinson, A.B., Piketty, T., Saez, E., 2011. Top incomes in the long run of history. Journal of Economic Literature 49, 3–71.

Autor, D.H., Katz L.F., Kearney M.S., 2008. Trends in US wage inequality: Revising the revisionists. Review of Economics and Statistics 90, 300–323.

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

Constantinides, G.M., Duffie, D., 1996. Asset pricing with heterogeneous consumers. Journal of Political Economy 104, 219–240.

Croce, M.M., Kung, H., Nguyen, T.T., Schmid, L., 2012. Fiscal policies and asset prices. Review of Financial Studies 25, 2635–2672.

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

Eaton, J., Rosen, H.S., 1980. Taxation, human capital, and uncertainty. American Economic Review 70, 705–715.

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

Fischer, M., Jensen, B.A., 2015. Taxation, transfer income and stock market participation. Review of Finance 19, 823–863.

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

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

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

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

Hombert, J., Schoar, A., Sraer, D., Thesmar, D., 2014. Can unemployment insurance spur entrepreneurial activity? NBER working paper.

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

Kacperczyk, M., Nosal, J.B., Stevens, L., 2015. Investor sophistication and capital income inequality. Working paper, Imperial College.

Kelly, B., Pástor, L., Veronesi, P., 2016. The price of political uncertainty: Theory and evidence from the option market. Journal of Finance, forthcoming.

Kihlstrom, R.E., Laffont, J.-J., 1979. A general equilibrium entrepreneurial theory of firm formation based on risk aversion. Journal of Political Economy 87, 719–748.

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

Meltzer, A.H., Richard, S.F., 1981. A rational theory of the size of government. Journal of Political Economy 89, 914–927.

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

Moskowitz, T.J., Vissing-Jorgensen, A., 2002. The returns to entrepreneurial investment: A private equity premium puzzle? American Economic Review 92, 745–78.

Pástor, L., Veronesi, P., 2003. Stock valuation and learning about profitability. Journal of Finance 58, 1749–1789.

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

Pástor, L., Veronesi, P., 2012. Uncertainty about government policy and stock prices. Journal of Finance 67, 1219–1264.

Pástor, L., Veronesi, P., 2013. Political uncertainty and risk premia. Journal of Financial Economics 110, 520–545.

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

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

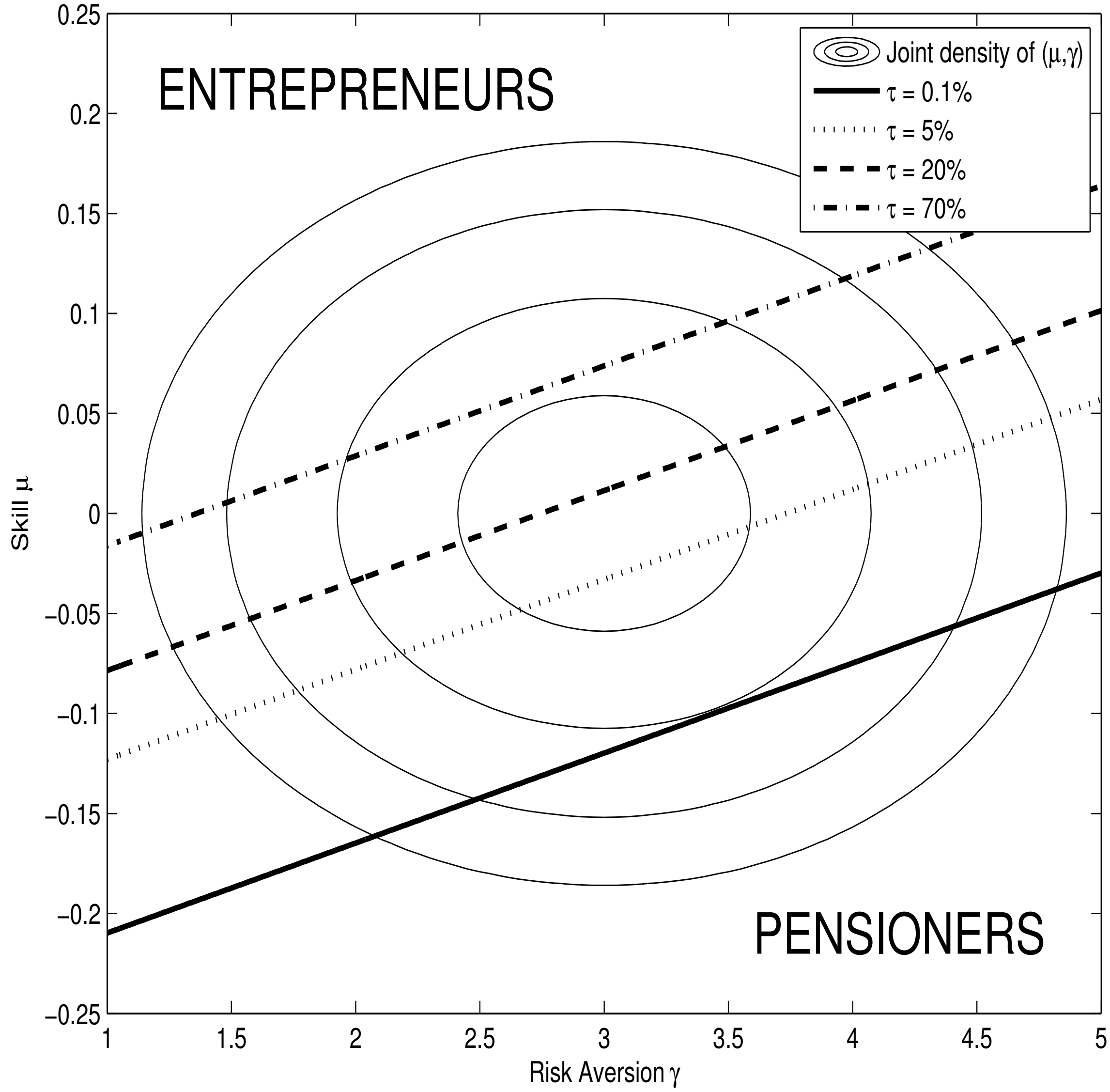

Figure 1. The agents' decision. Each point in this graph represents an agent with the corresponding skill μ_i and risk aversion γ_i. All agents located above the threshold line choose to become entrepreneurs; those below the line become pensioners. The four lines correspond to four different tax rates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

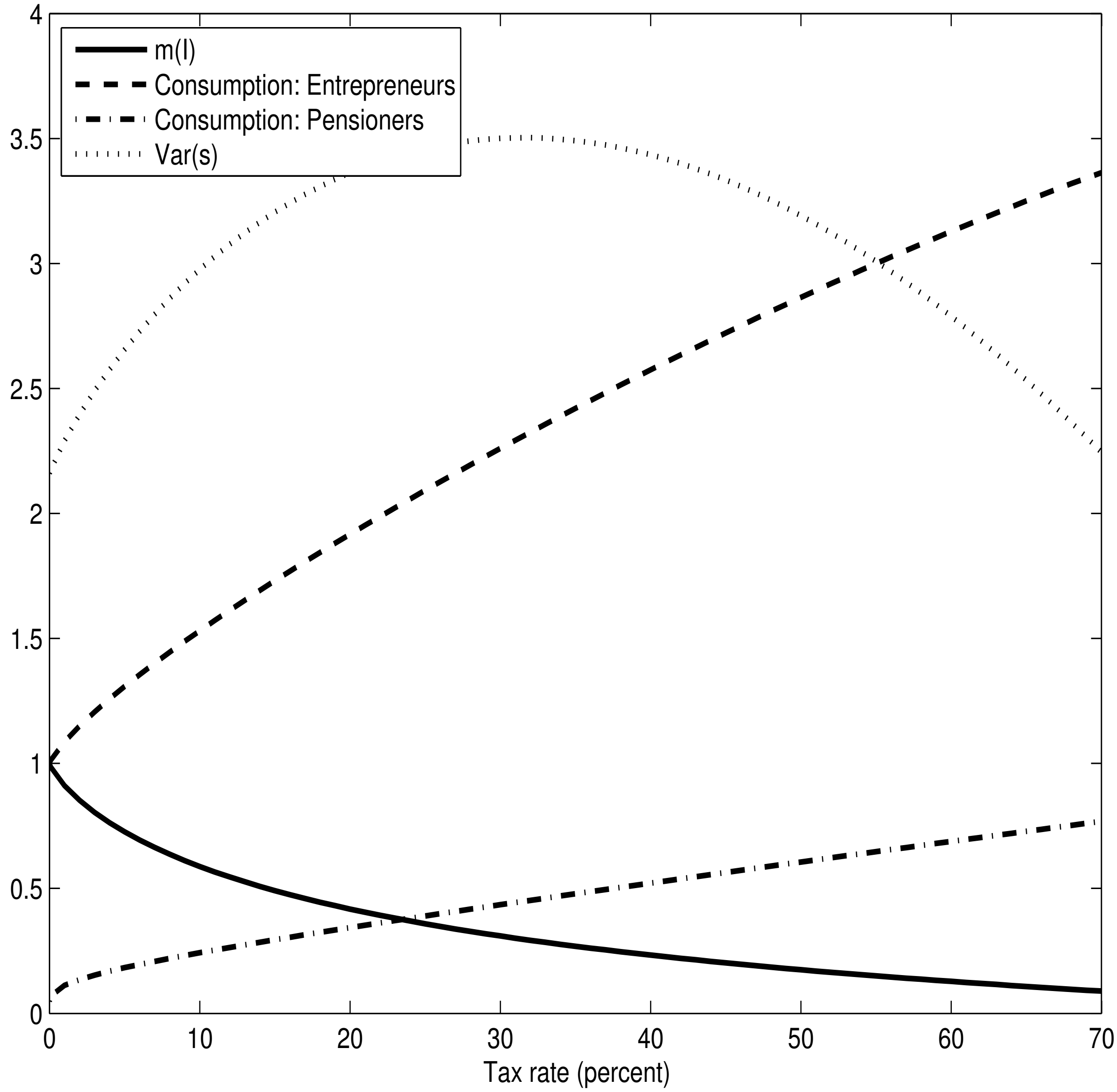

Figure 2. The share of entrepreneurs and agents' consumption. This figure plots four quantities as a function of the tax rate τ. The solid line plots $m(\mathcal{I})$, the fraction of agents who become entrepreneurs. The dashed line plots the average consumption of entrepreneurs, which is given by $\frac{1-\tau}{m(\mathcal{I})}$. 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.

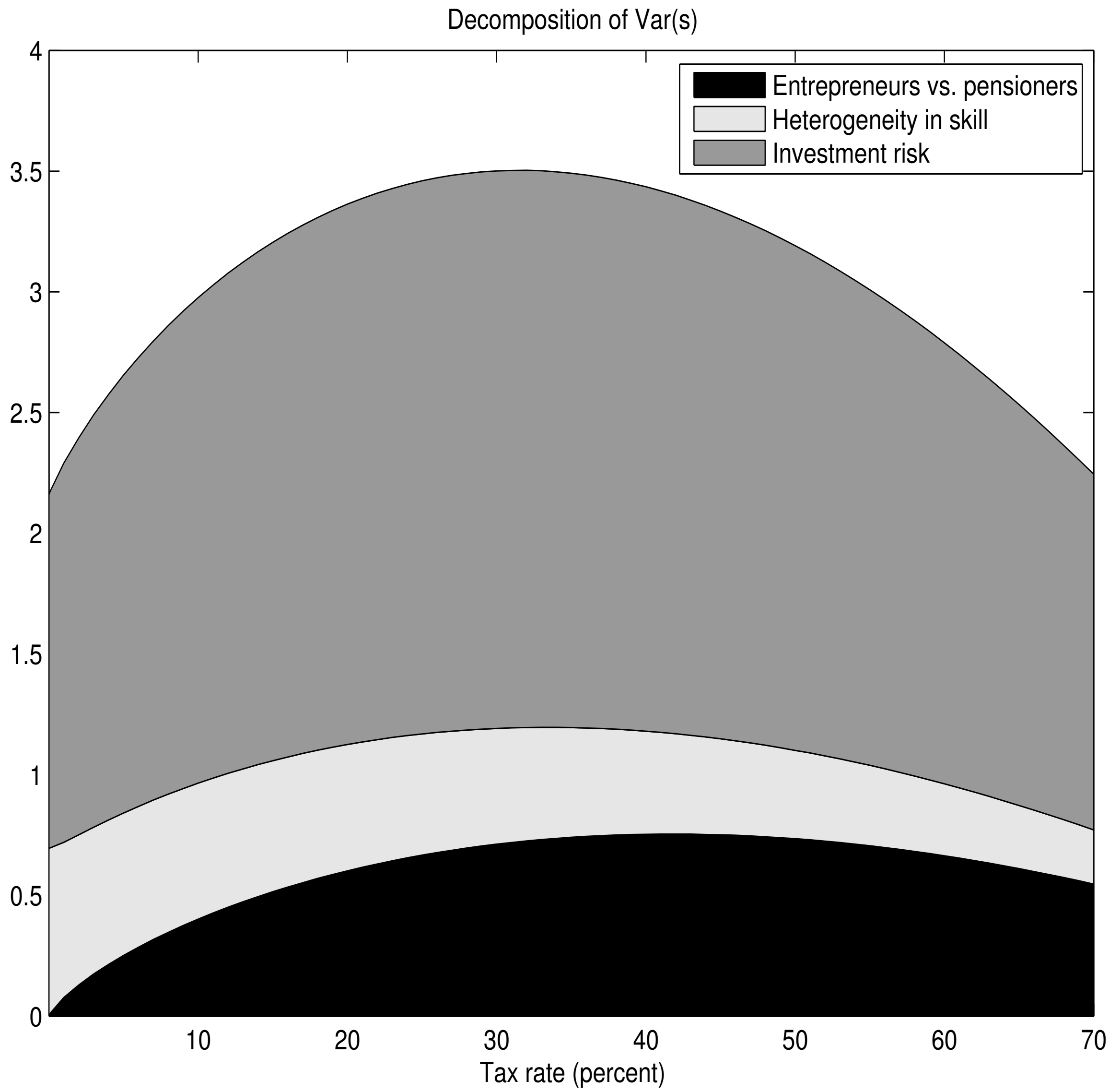

Figure 3. Three sources of income inequality. This figure plots three components of the variance of consumption across agents from equation (31) as a function of the tax rate τ. The black area at the bottom plots the component due to the difference between the entrepreneurs' and pensioners' average consumption 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 dark-grey 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 light-grey area in the middle plots the component due to heterogeneity in μ_i across entrepreneurs. This component is obtained as the residual by subtracting the other two components from total variance.

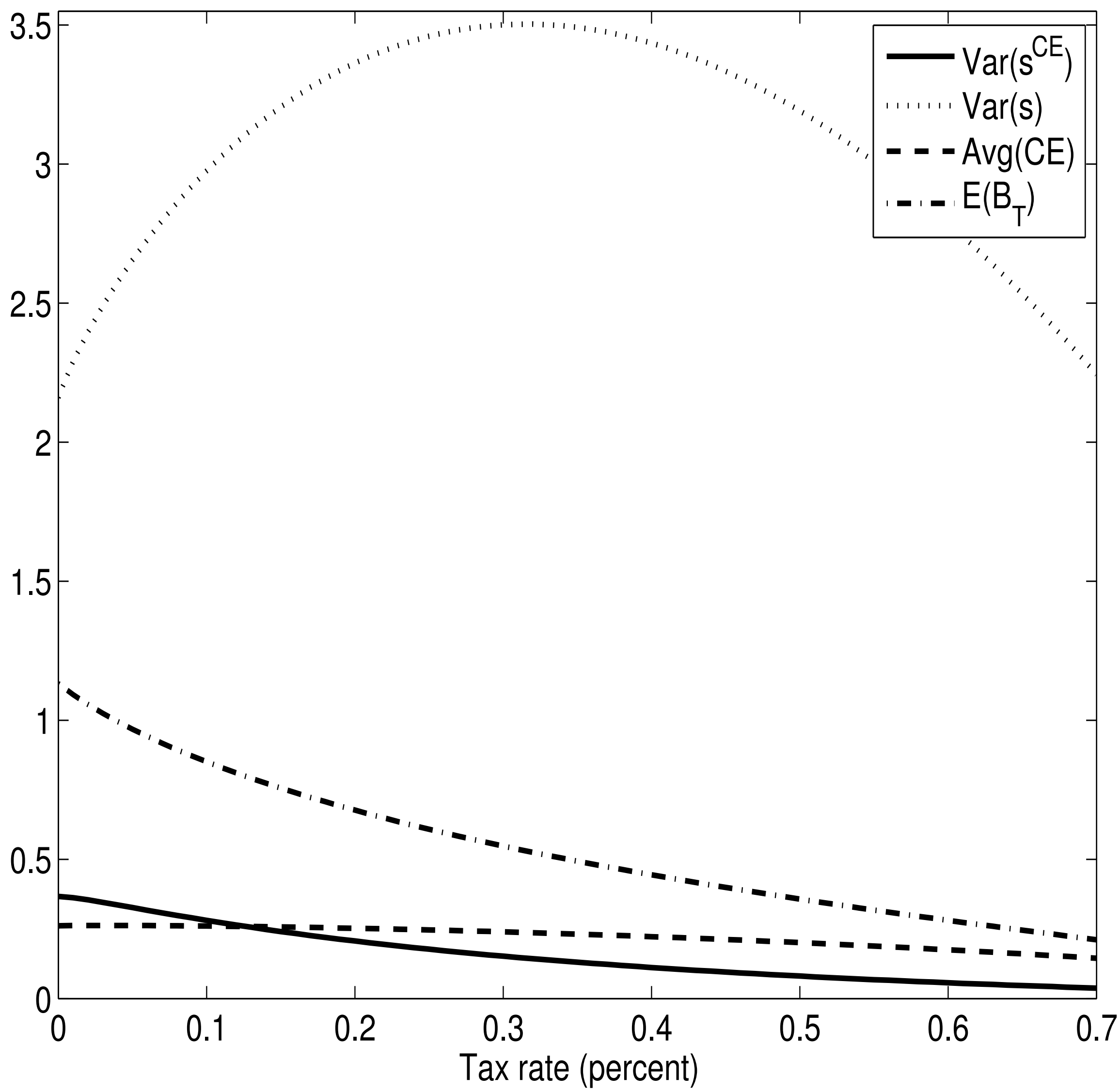

Figure 4. Inequality in expected utility vs. inequality in consumption. The solid line plots inequality in expected utility expressed in consumption terms, measured by the variance of certainty equivalent consumption across agents, as a function of the tax rate τ. The dotted line plots inequality in 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 $B_0 = 1$.

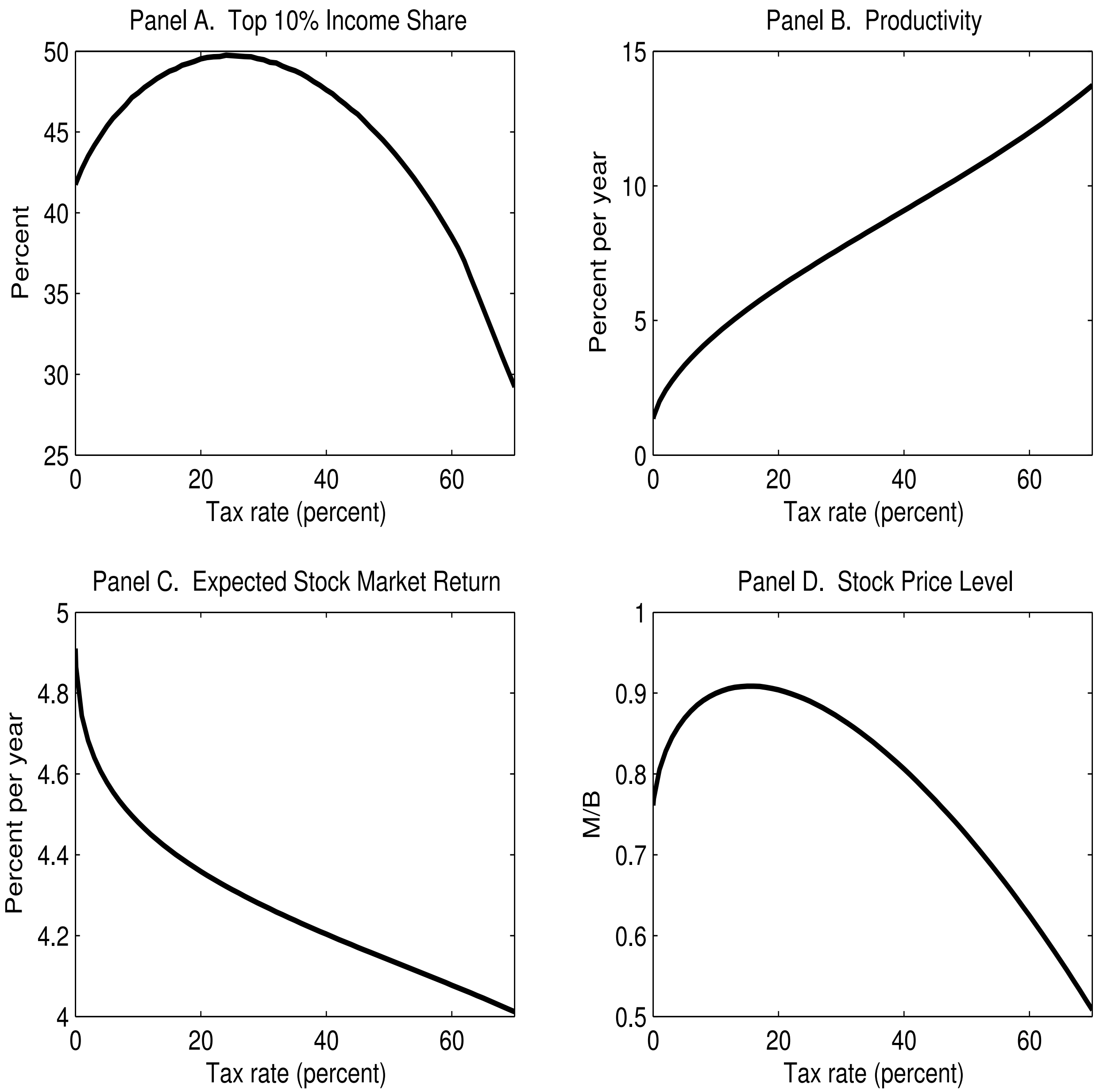

Figure 5. The effects of taxes in the model. This figure plots four variables as a function of the tax rate τ. Panel A plots income inequality, measured by the income share of the top 10% of agents. Panel B plots expected aggregate productivity, measured by the annualized expected growth rate of total capital, or $(1/T)\mathrm{E}\left[B_T/(m(\mathcal{I})B_0)-1\right]$. Panel C plots the expected rate of return on the aggregate market portfolio. Panel D plots the stock price level, measured by the market-to-book ratio of the market portfolio. All parameter values are equal to their baseline values.

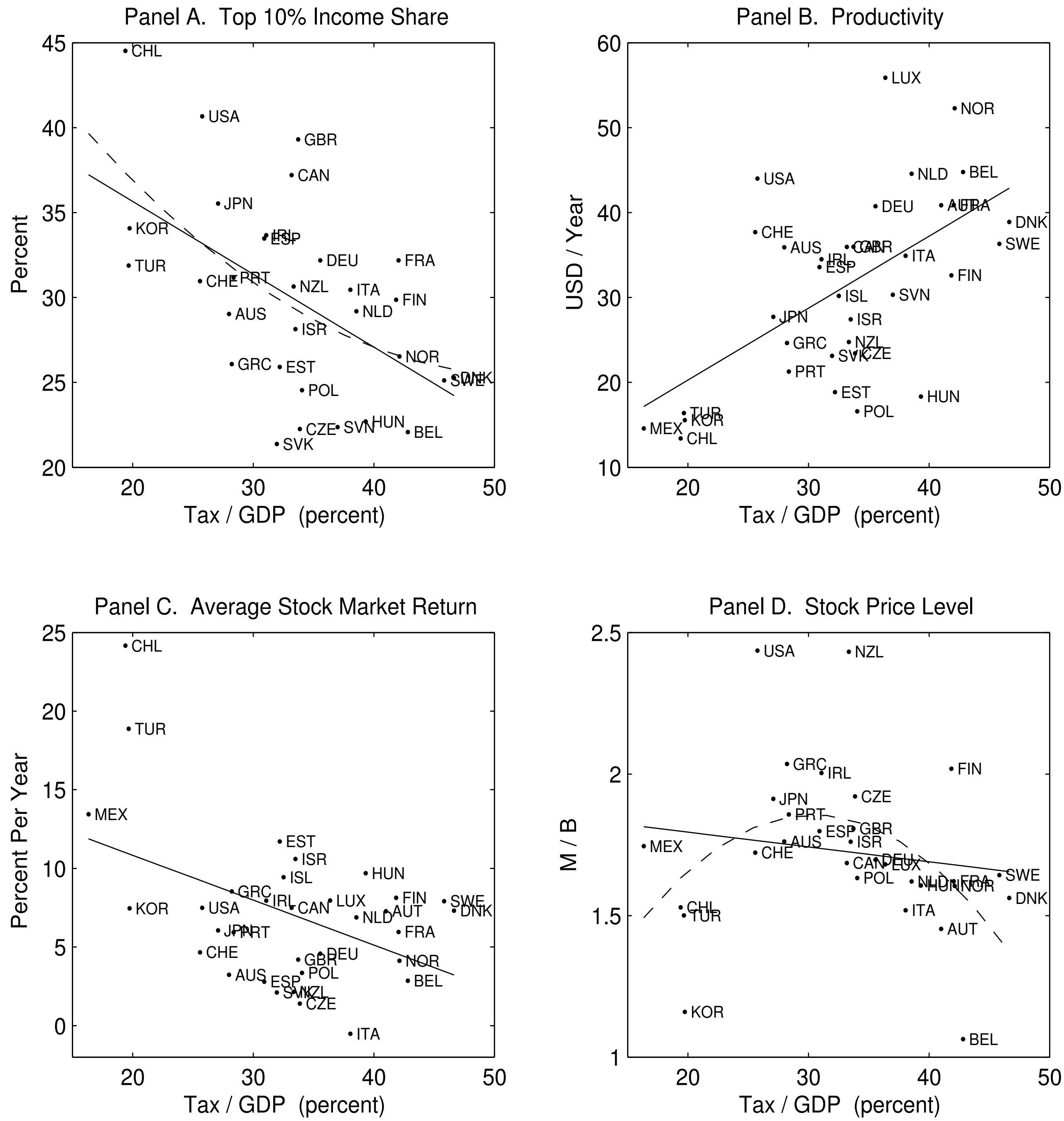

Figure 6. The effects of taxes in the data. This figure is an empirical counterpart of Figure 5. The figure plots four variables against the tax-to-GDP ratio (TAX) across countries. Panel A plots income inequality, measured by the top 10% income share (TOP). Panel B plots productivity, measured by GDP per hour worked ($PROD$). Panel C plots the average stock market index return (RET). Panel D plots the stock price level, measured by the aggregate market-to-book ratio (M/B). All variables are computed at the country level as time-series averages in 1980–2013, except for RET, which uses all available data. Each dot, labeled with the OECD country code, is a country-level observation. The solid line in each panel is the line of best fit from a linear regression. The dashed lines are the lines of best fit from quadratic regressions of TOP (Panel A) and M/B (Panel D) on TAX and TAX^2.

Table 1
The role of taxes: Linear regressions

This table reports the results from cross-country regressions of the time-series average of the dependent variable on the time-series averages of the variables given in the row labels. There are four dependent variables: (i) income inequality, measured by the country's share of income going to the top 10% (columns 1 and 2), (ii) productivity, measured by the country's GDP per hour worked (columns 3 and 4), (iii) the stock market index return, measured in local currency in real terms (columns 5 and 6), and (iv) the stock price level, measured by the country's aggregate market-to-book ratio (columns 7 and 8). The independent variables are the country's tax-to-GDP ratio (TAX), GDP growth ($GDPGRO$), GDP per capita ($GDPPC$), and consumer price inflation ($INFL$). All time-series averages are computed from all available annual data between 1980 and 2013, except for average stock returns, which are computed from all available data. t-statistics are in parentheses. The coefficient on $GDPPC$ is multiplied by 100. The intercepts are included but not reported. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	Dependent variable							
	Income inequality		Productivity		Stock return		Stock price level	
TAX	-0.43***	-0.54***	0.85***	0.31***	-0.29***	-0.16	-0.01	-0.01**
	(-3.53)	(-4.46)	(4.23)	(3.53)	(-2.92)	(-1.51)	(-0.81)	(-2.16)
$GDPGRO$		0.04		-0.50		0.78		−0.11**
		(0.05)		(-0.95)		(1.21)		(-2.48)
$INFL$		-0.06		0.02		0.09		0.00
		(-1.32)		(0.60)		(1.28)		(0.38)
$GDPPC$		0.03**		0.10***		-0.00		0.00
		(2.03)		(13.53)		(-0.45)		(1.02)
Sample size	30	30	34	34	33	33	30	30
R^2	0.29	0.47	0.35	0.91	0.20	0.33	0.02	0.22

Table 2
The role of taxes: Quadratic regressions

This table reports the results from cross-country regressions of the time-series average of the dependent variable on the time-series averages of the variables given in the row labels. There are two dependent variables: income inequality, measured by the country's share of income going to the top 10% (columns 1 and 2), and the stock price level, measured by the country's aggregate market-to-book ratio (columns 3 and 4). The independent variables are the country's tax-to-GDP ratio, both plain and squared (TAX and TAX^2), GDP growth ($GDPGRO$), GDP per capita ($GDPPC$), and consumer price inflation ($INFL$). All time-series averages are computed from all available annual data between 1980 and 2013. t-statistics are in parentheses. The coefficients on TAX^2 and $GDPPC$ are multiplied by 100. The intercepts are included but not reported. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

	Dependent variable			
	Income inequality		Stock price level	
TAX^2	1.11	1.37	-0.18***	-0.14**
	(0.80)	(1.08)	(-2.84)	(-2.21)
TAX	-1.16	-1.46*	0.11***	0.08*
	(-1.25)	(-1.71)	(2.68)	(1.83)
$GDPGRO$		-0.21		-0.09**
		(-0.28)		(-2.04)
$INFL$		-0.05		0.00
		(-1.21)		(0.24)
$GDPPC$		0.03**		0.00
		(2.05)		(0.51)
Sample size	30	30	30	30
R^2	0.31	0.49	0.23	0.33