2022

2022-14 Use It or Lose It: Efficiency and Redistributional Effects of Wealth Taxation

Fatih Guvenen

Gueorgui Kambourov

Burhan Kuruscu

Sergio Ocampo

Daphne Chen

Use It or Lose It: Efficiency and Redistributional Effects of Wealth Taxation

by

Fatih Guvenen, Gueorgui Kambourov, Burhan Kuruscu, Sergio Ocampo, and Daphne Chen

Research Report # 2022-14 December 2022

Department of Economics

Research Report Series

Department of Economics
Social Science Centre
Western University
London, Ontario, N6A 5C2
Canada

Use It or Lose It: Efficiency and Redistributional Effects of Wealth Taxation*

Fatih Guvenen Gueorgui Kambourov Burhan Kuruscu
Sergio Ocampo Daphne Chen

Abstract

How does wealth taxation differ from capital income taxation? When the return on investment is equal across individuals, a well-known result is that the two tax systems are equivalent. Motivated by recent empirical evidence documenting persistent return heterogeneity, we revisit this question. With heterogeneity, the two tax systems typically have opposite implications for both efficiency and inequality. Under capital income taxation, entrepreneurs who are more productive and therefore generate more income pay higher taxes. Under wealth taxation, entrepreneurs who have similar wealth levels pay similar taxes regardless of their productivity, which expands the tax base, shifts the tax burden toward unproductive entrepreneurs, and raises the savings rate of productive ones. This reallocation increases aggregate productivity and output. In the simulated model parameterized to match the US data, replacing the capital income tax with a wealth tax in a revenue-neutral fashion delivers a significantly higher average welfare. Turning to optimal taxation, the optimal wealth tax (OWT) is positive and yields large welfare gains by raising efficiency and lowering inequality. In contrast, the optimal capital income tax (OKIT) is negative—a subsidy—and delivers lower welfare gains than OWT, owing to the welfare losses from higher inequality. Furthermore, when the transition path is considered, the gains from OKIT turn into significant welfare losses for existing cohorts, whereas OWT continues to deliver robust welfare gains. These results suggest that moderate wealth taxation may be a more appealing alternative than capital income taxation, which can be significantly more distorting under return heterogeneity than under the equal-returns assumption. JEL Codes: E62, H21, H24.

Keywords: Wealth tax, Capital income tax, Optimal taxation, Rate of return heterogeneity, Power law models, Pareto tail, Wealth inequality.

*This draft: December, 2022. First draft: April 2014. For helpful comments, we thank Mark Aguiar, Adrien Auclert, Jess Benhabib, Alberto Bisin, Richard Blundell, V. V. Chari, Mariacristina De Nardi, Andres Erosa, Priscilla Fialho, Xavier Gabaix, Mike Golosov, Joao Gomes, Pierre-Olivier Gourinchas, Luigi Guiso, Bob Hall, Hugo Hopenhayn, Nobu Kiyotaki, Henrik Kleven, Dirk Krueger, Per Krusell, Guido Menzio, Claudio Michelacci, Ben Moll, Juan Pablo Nicolini, Thomas Piketty, Victor Ríos-Rull, Martin Schneider, Rob Shimer, Stefanie Stantcheva, Aleh Tsyvinski, Maxim Troshkin, Gabriel Zucman, and seminar and conference participants at various institutions. We thank the editor and five anonymous referees for comments that substantially improved the paper. For funding, Guvenen thanks the National Science Foundation; Kuruscu thanks the Marie Sklodowska-Curie programme under Horizon 2020; Kambourov and Kuruscu thank the Social Sciences and Humanities Research Council; and Ocampo thanks the ERC starting grant ValuRed (n. 804104). Daphne Chen made valuable contributions in the early stages of the project, before leaving for a career change. Guvenen: U of Minnesota, FRB of Minneapolis, and NBER; guvenen@umn.edu; Kambourov: U of Toronto; g.kambourov@utoronto.ca; Kuruscu: U of Toronto; burhan.kuruscu@utoronto.ca; Ocampo: Western University; socampod@uwo.ca; Chen: Econ One.

1 Introduction

We start this paper with a simple question: How does wealth taxation differ from capital income taxation? To fix ideas, let a denote wealth, r denote the rate of return on wealth, and τ_k and τ_a denote the tax rates on capital income and wealth, respectively. Under a capital income tax, the after-tax wealth of individual i is given by

$$a_i^{\text{after-tax}} = a_i + (1 - \tau_k) \times r a_i,$$

whereas under the wealth tax, it is

$$a_i^{\text{after-tax}} = (1 - \tau_a) \times a_i + r a_i.$$

In a variety of benchmark economic models, the answer to the question above is not very interesting: the two tax systems are equivalent, with $\tau_a = r\tau_k$. Partly because of this equivalence, the academic literature on capital taxes has traditionally focused on the capital income tax, with the understanding that it can be reinterpreted as a wealth tax. However, the equivalence result relies on the assumption that all individuals face the same rate of return, which we also made implicitly above by not indexing r with a subscript i. What happens if instead rates of return vary across individuals—as the empirical evidence we review below indicates?

To see some of the implications for capital taxation, consider two entrepreneurs who start out with the same wealth level—say, \$1,000 each—but earn different returns—say, $r_1 = 0\%$ and $r_2 = 20\%$. Under capital income taxation, the unproductive (first) entrepreneur will escape taxation because he generates no income, and the tax burden will fall entirely on the more productive (second) entrepreneur. Under wealth taxation, on the other hand, both entrepreneurs will pay the same amount of tax on wealth regardless of their productivity; as a result wealth taxation expands the tax base, shifts the tax burden toward the unproductive entrepreneur, and reduces (potential) tax distortions on the productive entrepreneur. To the extent that these differences in productivity are persistent, a wealth tax will gradually prune the wealth of idle entrepreneurs and boost that of successful ones, leading to a more efficient allocation of aggregate capital,[1] in turn raising productivity and output. In this sense, wealth taxation has a "use-it-or-lose-it" effect absent from capital income taxation. We expand on this example in Section 2.

[1]Throughout the paper, we use the terms *capital* and *wealth* interchangeably.

While this is a clearly stylized example, it illustrates how (rate of) return heterogeneity can drive a wedge between the implications of the two ways of taxing capital. In this paper, we study these implications in a full-fledged overlapping-generations model with rich heterogeneity and intergenerational links. The main contribution of this paper is to flesh out new economic mechanisms from wealth and capital income taxation that become operational when returns are heterogenous. As we elaborate in a moment, we find that the two taxes have very different—and sometimes opposite—implications.

There are three more considerations that motivate us to study capital taxation under return heterogeneity. First, a growing number of empirical studies cast strong doubt on the assumption of homogeneous returns. Using administrative panel datasets that track millions of individuals over time, these studies document large and persistent differences in individual returns, even after adjusting for risk and other factors (e.g., Fagereng, Guiso, Malacrino, and Pistaferri, 2020, Bach, Calvet, and Sodini, 2020, and Smith, Zidar, and Zwick, forthcoming, among others).[2] These new pieces of evidence make studying the tax implications of return heterogeneity more than a theoretical curiosity.

Second, the literature on power law models shows that return heterogeneity is a powerful modeling tool that can generate key features of wealth inequality that proved challenging to explain through other mechanisms.[3] This is an important consideration for studying capital taxation: because of the extreme concentration of wealth in the United States and in many other countries (Vermeulen, 2018), the bulk of the capital tax burden falls on a small fraction of wealthy households. In addition, a large fraction of these very wealthy households are self-made rather than inheritors, which is an important determinant of the trade-offs they face and how they respond to capital taxation. Thus, for a sound quantitative analysis of capital taxation, we believe that it is important for a model to reproduce key features of the top-end wealth distribution, such as the thick Pareto tail and the very rapid wealth accumulation of the super wealthy.[4]

Third, studying wealth taxation also has a practical motivation: it is a policy tool that has long been used by governments around the world.[5] While its popularity was on decline until recently, the last few years have seen a revival of interest in many countries. In light

[2]This result holds for both public and private equity investments, although it is larger in the latter case. We discuss this evidence in more detail in Section 4.

[3]See Gabaix (2009) and Benhabib and Bisin (2018) for reviews of this literature.

[4]For example, 54% of the 2017 US *Forbes* 400 billionaires (with a minimum wealth of $2 billion) were self-made, which implies a conservative lower bound of a 10,000-fold increase in their wealth over the life cycle. As we show in this paper, a calibrated model with return heterogeneity can generate this pattern.

[5]Until the last decade or so, France, Germany, Spain, Italy, the Netherlands, and the Nordic countries, among others, had wealth taxation. See OECD (2018) for a recent review of the use of wealth taxes.

of this reality, studying the implications of wealth taxation and how it differs from capital income taxation is an important step toward providing better guidance to policy makers.

We study an overlapping-generations model in which individuals derive utility from consumption and leisure, and have a warm-glow bequest motive. The key ingredient of the model is persistent heterogeneity in entrepreneurial productivity, which, together with incomplete financial markets that prevent the free flow of funds across individuals, allows some individuals to earn persistently higher returns on wealth than others. The model features a bond market where individuals can borrow, subject to a collateral constraint, to rent capital to use in their firm. The same bond market can also be used as a savings device, which will be optimal for individuals whose entrepreneurial productivity (hence, their return) is very low.

Each individual/entrepreneur produces a differentiated intermediate good using a proprietary technology with individual-specific productivity. These intermediates are combined in a Dixit-Stiglitz aggregator by a final goods producing firm, which pins down (together with the collateral constraint) each entrepreneur's production scale and profits. In this setup, every entrepreneur earns a monopoly profit. Individuals face idiosyncratic labor income risk, mortality risk, and various intergenerational links, although plausible variations in these details do not change the substantive conclusions. The calibrated model is consistent with key features of the US data, including top-end wealth inequality (including the Pareto tail), the amount of return heterogeneity, statistics on entrepreneurs, and the magnitude of borrowing by US businesses, among other features.

Our analysis produces four sets of results. First, in Section 5, we study a revenue-neutral tax reform that replaces the existing US tax system of capital income taxation with a flat-rate wealth tax. Comparing across stationary equilibria (we consider the transition path later), we find that this reform raises average welfare significantly—by about 7% of consumption-equivalent per year for newborn individuals in our baseline calibration. The gains come from a combination of a higher capital level and a more efficient allocation of capital generated by the use-it-or-lose-it mechanism. In the tax reform, switching to a wealth tax *lowers* the tax burden on capital (to 4% of GDP, from 6% in benchmark) because we keep the revenue and labor tax rate fixed and the economy is larger with a wealth tax. As a result, with a lower tax burden on the wealthy, their incentives for evasion may not be higher under a wealth tax.

Second, in Section 6, we conduct an optimal tax analysis in which a utilitarian government chooses flat-rate taxes on labor income and wealth to maximize the ex ante expected lifetime utility of a newborn. The optimal wealth tax (hereafter, OWT) rate is positive and relatively high, at about 3%. The high revenues from wealth taxes allow the government to reduce the tax on labor income, which is more distorting than the wealth tax in this environment. Overall, output and consumption are significantly higher in the OWT economy. The bulk of the gains come from an improved allocation of capital (as in the tax reform) and almost none from a change in the capital stock—which remains almost unchanged—in the new stationary equilibrium.

The analogous optimal capital income taxation (hereafter, OKIT) exercise delivers an optimal (linear) tax rate of –13.6%, implying a nontrivial *subsidy* to capital income. This finding may seem surprising in light of previous results in the literature that found a high positive tax rate (of about 35%), using Aiyagari-style models that share many similarities with ours (e.g., Conesa, Kitao, and Krueger, 2009). The main difference is return heterogeneity: shutting down return heterogeneity restores the high positive tax rate found in previous work. To understand why this happens, note that in Aiyagari-style models, the wealthy are *workers* who have been productive in the past, but they are not any better at investing this wealth than others, so the efficiency losses from capital income taxation are not especially large. In contrast, with return heterogeneity, entrepreneurs who earn high capital income (per unit of wealth) are precisely the productive ones today, which makes taxing capital income much more distorting and the efficiency losses especially large, making a subsidy optimal. A key takeaway from these results is that, beyond its implications for wealth taxation, accounting for return heterogeneity can matter greatly for studying capital income taxation.

Third, OWT delivers higher average welfare than OKIT. About 2/3 of the welfare gains of OWT come from the rise in the *level* of consumption and 1/3 from the decline in the *inequality* of (the marginal utility of) consumption. We also consider a progressive wealth tax, enabled by an optimally chosen tax exemption level. While it delivers only marginally higher average welfare gains, a larger fraction of welfare now comes from distributional gains. Thus, wealth taxes can yield both first- and second-order gains. This is not the case with OKIT: the large subsidies, coupled with the high labor income taxes that the policy requires, increase inequality, resulting in distributional losses.

Fourth, we extend the optimal tax analysis to incorporate transition to understand how the individuals who are alive at the time of the policy switch fare from the reform. The OKIT policy leads to widespread welfare *losses* for individuals who are alive at the time of the policy change, whereas the OWT policy continues to deliver significant welfare gains for both the newborns and the overall population. The main reason for this contrast is that OKIT works primarily by inducing higher savings during the transition, leading to a higher capital stock, which is costly. By contrast, OWT works through reallocation without (much) change in the capital stock. We discuss these results in Section 6.4.

Two elements that are common to all the experiments we study in this paper are that (i) they *replace* a capital income tax with a wealth tax—rather than *adding* a wealth tax on top of existing capital income taxes—and (ii) they are *broad-based* rather than being levied on multimillionaires or billionaires only. In this sense, they are very different from the proposals circulating in public debates today, which advocate a wealth tax as an additional tax targeted at the top. Thus, while the mechanisms we study in this paper certainly inform this broader wealth tax debate, we caution the reader about drawing direct conclusions regarding those tax proposals.

Another important difference of the wealth tax we propose is that it is levied on the *book value* of assets—not on their market value—which maximizes the efficiency gains from the use-it-or-lose-it mechanism. This is because the market value incorporates the future profit stream of the firm, which in turn depends on the productivity of the entrepreneur. Therefore, a tax on the market value falls disproportionately on more productive entrepreneurs, partially undoing the positive reallocation created by the wealth tax. Incidentally, an important practical challenge with implementing the wealth tax has been to assess market values, a problem that would be alleviated by taxing the book value.

The issue of book versus market values is one example of practical considerations that arise when implementing a wealth tax (like any other tax). A short list of these includes the possibility of exacerbating capital flight, which is already happening under capital income taxation owing to the ease of shifting intangible capital across borders (Guvenen, Mataloni, Rassier, and Ruhl, 2022b; Tørsløv, Wier, and Zucman, 2022); how to tax unrealized capital gains, which are increasingly being used as a tax shelter by some wealthy households; and whether the wealth tax should be levied on the firm side (similar to a corporate income tax) or on the household side (similar to a dividend or capital gains tax), among others. In this paper, we do not tackle these important issues, which are left for future work. We share some thoughts on how they can potentially be addressed in the concluding section.

Finally, in Section 7, we conduct various sensitivity checks and extensions, including adding a corporate sector, considering various formulations of financial constraints, studying a version with pure monopolistic rents, and allowing for nonlinear capital income taxation, among others. These changes affect the various magnitudes of welfare gains, but they do not overturn the main substantive conclusions of our analysis.

Related Literature

Although the use-it-or-lose-it feature of wealth taxes has been noted by a few authors, we are not aware of prior academic analyses of its effects.[6] Maurice Allais was among the best-known proponents of wealth taxes, and he discussed the use-it-or-lose-it rationale in Allais (1977). More recently, Piketty (2014) revived the debate on wealth taxation and proposed using a combination of capital income and wealth taxes to balance these efficiency and inequality trade-offs. Piketty focused mostly on equity considerations but also described the use-it-or-lose-it mechanism without providing a formal analysis.

The broader literature on capital taxation is vast, so we will not attempt to review it here; see Chari and Kehoe (1999), Golosov, Tsyvinski, and Werning (2006), Stantcheva (2020), and Scheuer and Slemrod (2021) for excellent surveys. Our paper is more closely related to the quantitative public finance literature that allows for incomplete markets, plausibly restricted tax instruments, and finitely lived individuals (Hubbard, Judd, Hall, and Summers, 1986; Aiyagari, 1995; Imrohoroglu, 1998; Erosa and Gervais, 2002; Garriga, 2003; Conesa, Kitao, and Krueger, 2009; Kitao, 2010). Some of these studies found that the optimal capital tax rate is positive and large. The two main differences between our analysis and these studies are the presence of heterogeneous returns and the consideration of wealth taxation. On capital income taxation, our contribution is to show that if heterogeneity is sufficiently large, it alters some key conclusions and turns the optimal policy from a tax to a subsidy. On wealth taxation, we show that its effects can be qualitatively very different from taxing capital income and yield larger and more broad-based welfare gains.

Our paper has useful points of contact with different literatures that feature (entrepreneurial) firms with heterogeneous productivity facing financial frictions. Examples include Restuccia and Rogerson (2008) and Hsieh and Klenow (2009) in the context of aggregate TFP; Buera, Kaboski, and Shin (2011), Midrigan and Xu (2014), Moll (2014), and Itskhoki and Moll (2019) in the context of economic development; and Quadrini (2000) and Cagetti and De Nardi (2006) in the context of entrepreneurship, among others. These papers do not study tax policies in general, with the exception of Itskhoki and Moll (2019), whose conclusions share some interesting similarities with ours. These authors find that along the development path, the optimal policy starts by suppressing wages to boost entrepreneurial profits and wealth accumulation, which relaxes borrowing constraints over time, yielding higher productivity and wages. In the long run, optimal policy reverses and

[6] This paper was first presented at the 2014 NBER Summer Institute and has been widely presented at seminars and conferences since then. It therefore predates the recent public debate on wealth taxation that rose to prominence during the 2020 presidential election campaign.

becomes pro-worker. In our framework, wealth taxation plays a similar role.

Finally, as noted above, this paper is also related to the growing literature on power law models of inequality. This literature shows that the thick Pareto tail of the wealth distribution, which is challenging to generate (even for some models of inequality that match the share of wealth held by the top 1%), emerges naturally in models with return heterogeneity (Benhabib, Bisin, and Zhu, 2011; Benhabib, Bisin, and Luo, 2017). Moreover, if return heterogeneity is persistent, these models also generate behavior that is consistent with the *dynamics* of wealth inequality over time (Gabaix, Lasry, Lions, and Moll, 2016; Jones and Kim, 2018). Despite the rapid growth of this literature, the implications of capital taxation in these models have not been explored, and our paper fills this gap.

2 An Illustrative Example

It is useful to elaborate on the simple one-period example described in the Introduction. Consider two brothers, Fredo and Michael, who each have $100 million of wealth. Fredo has low entrepreneurial skills, so he earns a return of $r_F = 0\%$ on his investments, whereas Michael is a highly skilled businessman and earns a return of $r_M = 20\%$. Both brothers invest their wealth in their business and make no other decisions. There is also a government that needs to finance an expenditure of $G = \$5$ million through tax revenues collected at the end of the period. The example is summarized in Table I.

First, consider capital income taxation. The required tax rate to raise $5 million is $\tau_k = 25\%$, which will be paid entirely by Michael, since he is the only one generating capital income. Michael's after-tax rate of return is 15% (down from 20%), while Fredo's return is unaffected (still 0%). Michael ends the period with $115M of wealth, up from $100M, while Fredo's wealth remains unchanged.

Now consider a wealth tax imposed on beginning-of-period wealth (right panel). First, this doubles the tax base, which now includes Fredo's wealth. The tax rate on wealth is $\tau_a = 5/200 = 2.5\%$. More importantly, now half of the tax bill is paid by Fredo, while Michael's tax bill is cut by half. As a result, Fredo's after-tax return is now –2.5%, while Michael's return is 17.5%, so wealth taxation does not compress the distribution of returns or wealth across investors as much as capital income taxation does (and does not compress at all in this specific example). Notice that wealth dispersion in this example is between wealthy investors and not across the broader population. As we will see in the quantitative analysis below, wealth taxation can deliver (large) distributional welfare gains (by raising wages) in addition to its efficiency benefits.

Table I – Summary of the Illustrative Example

	Capital Income Tax		Wealth Tax	
	Fredo	Michael	Fredo	Michael
Rates of return	$r_F = 0\%$	$r_M = 20\%$	$r_F = 0\%$	$r_M = 20\%$
Wealth	100M	100M	100M	100M
Pre-tax income	0	20M	0	20M
Tax liability	0	$20\tau_k$ =5M	$100\tau_a = 2.5$M	$100\tau_a = 2.5$M
After-tax rate of return	0%	15%	−2.5%	17.5%
After-tax wealth ratio	$\frac{W_M}{W_F} = \frac{115}{100} = 1.15$		$\frac{W_M}{W_F} = \frac{117.5}{97.5} \approx 1.20$	

Notes: $\tau_k = 25\%$; $\tau_a = 2.5\%$ is imposed on the beginning-of-period wealth. See the text for further details of this example.

To sum up, wealth taxation has two main effects that are the opposite of those created by capital income taxes. *First*, by shifting the tax burden toward the less productive entrepreneur, it allows the more productive one to keep more of his wealth, thereby reallocating the capital stock toward more productive uses. Going forward, we will refer to this first, direct effect as the "use-it-or-lose-it channel." *Second*, wealth taxes do not compress the after-tax return distribution nearly as much as capital income taxes, which effectively punish the successful entrepreneur and reward the inefficient one. In the dynamic setting we study next, this feature will deliver an endogenous (behavioral) savings response, further reallocating capital toward the more productive entrepreneur.[7]

Two final remarks are in order. First, if this one-period example were repeated for many periods, Michael would eventually hold all the wealth, so misallocation would vanish in the long run. But this prediction would be true only if return differences were completely permanent, which is not the case in real life: the fortunes of entrepreneurs vary both over time and from one generation to the next, so capital misallocation will persist even in the long run. These features will be incorporated into the dynamic model we describe next. Second, we assumed that the wealth tax was imposed on the book value of assets, not their market value; the distinction between the two is critical. This is because the latter incorporates the productivity of the entrepreneur through its effect on future earnings, so taxing the market value weakens the "use-it-or-lose-it" mechanism by raising the tax burden of more productive entrepreneurs.

[7] In addition, when the dynamic model is embedded in general equilibrium, the equilibrium response of prices (wages and interest rates) to tax policies will constitute another important effect.

3 Full OLG Model

We study an overlapping-generations model with two sectors (producing intermediate goods and a final good, respectively) and a government that raises revenues through taxes. We now describe each of these components.

3.1 Individuals

Individuals face mortality risk and can live up to a maximum of H years. When an individual dies, she is replaced by an offspring, who inherits her wealth. Individuals derive utility $u(c, 1-\ell)$ from consumption, c, and leisure, $1-\ell$ (ℓ denoting market hours), during their lifetime, as well as warm-glow utility $v(b)$ from the bequest, b, they leave upon death. They maximize expected lifetime utility:

$$\mathbb{E}_0 \left(\sum_{h=1}^{H} \beta^{h-1} \left[\phi_h u(c_h, 1-\ell_h) + (1-\phi_h) v(b) \right] \right), \tag{1}$$

where ϕ_h is the unconditional probability of survival to age h.

Individuals make three decisions every period: (i) leisure time versus labor supply to the market (until retirement age, $R < H$), (ii) consumption versus savings, and (iii) how much to produce of an intermediate good as an entrepreneur. Each individual is endowed with two types of skill: one that determines her productivity in entrepreneurial activities and another that determines her productivity as a worker. We now describe these skills, the production technologies, and the market arrangements, and then spell out each decision problem in more detail.

Entrepreneurial Ability and Productivity. The entrepreneurial *productivity* of individual i at age h, denoted z_{ih}, has two components: her entrepreneurial *ability*, $\bar{z}_i$, which is a fixed characteristic of the individual, and a second component—to be described in a moment—that captures the stochastic variation in productivity over the life cycle for a given ability level.[8] The ability component is transmitted imperfectly from a parent to her child:

$$\log\left(\bar{z}_i^{\text{child}}\right) = \rho_z \log\left(\bar{z}_i^{\text{parent}}\right) + \varepsilon_{\bar{z}_i}, \tag{2}$$

[8]We abstract from endogenous productivity arising from entrepreneurial effort. If higher returns were the result of innovation effort of entrepreneurs as in Jones and Kim (2018) and Jones (2022), switching to a wealth tax increases the entrepreneurial effort of higher ability entrepreneurs, which amplifies efficiency gains from wealth taxation, as shown in Guvenen, Kambourov, Kuruscu, and Ocampo (2022a).

where $\varepsilon_{\bar{z}_i} \sim \mathcal{N}\left(0, \sigma^2_{\bar{z}_i}\right)$. Because of this imperfect transmission, some low-ability children will inherit large fortunes from their high-ability parents, whereas some high-ability children will inherit little wealth from their low-ability parents, providing one source of capital misallocation in the model.

An entrepreneur faces many external factors that can amplify her ability (e.g., a lucky head start on a novel idea, good health and high drive) or hamper it (e.g., competitors catching up, negative health shocks, rising opportunity cost of time driven by family factors). While these shocks can conceivably happen at any age, positive factors are arguably more common at younger ages and negative ones later in life.

With this in mind, in the baseline model, we assume that high-ability entrepreneurs (specifically, those with $\bar{z}_i > \bar{z}_{\text{median}} = 1$) start life in the *fast lane*, with positive factors amplifying their productivity above their base level, $z_{ih} = \bar{z}_i^{\lambda}$, with $\lambda > 1$. (We consider alternative specifications for initial conditions later.) In every subsequent year, they face the risk of losing their place in the fast lane—for example, because of creative destruction by other entrepreneurs, as in Jones and Kim (2018)—and dropping to their base level, $z_{ih} = \bar{z}_i$, with annual probability p_1. With another probability p_2, all entrepreneurs (regardless of $\bar{z}_i$) face the risk of losing their productivity completely, $z_{ih} = 0$, and "retiring" from entrepreneurship. The evolution of z_{ih} can be summarized by the following three-state Markov chain, where $\mathbb{I}_{ih} \in \{\mathcal{H}, \mathcal{L}, 0\}$ is an indicator function:

$$z_{ih} = \begin{cases} \bar{z}_i^{\lambda} & \text{if } \mathbb{I}_{ih} = \mathcal{H} \\ \bar{z}_i & \text{if } \mathbb{I}_{ih} = \mathcal{L} \\ 0 & \text{if } \mathbb{I}_{ih} = 0 \end{cases} \qquad \text{and} \qquad \Pi_{\mathbb{I}} = \begin{bmatrix} 1 - p_1 - p_2 & p_1 & p_2 \\ 0 & 1 - p_2 & p_2 \\ 0 & 0 & 1 \end{bmatrix} \tag{3}$$

denoting the transition matrix, with $\mathbb{I}_{i0} = \mathcal{H}$ if $\bar{z}_i > \bar{z}_{\text{median}}$ and $\mathbb{I}_{i0} = \mathcal{L}$ otherwise.[9]

Modeling this stochastic variation in productivity serves three purposes. First, and most importantly, it allows for a more realistic calibration of the model to the wealth dynamics of the very wealthy, as we discuss in Section 4. Second, it introduces a second plausible source of capital misallocation (in addition to the intergenerational channel in eq. 2) that we believe is empirically relevant. Third, it provides an additional precautionary savings motive for individuals. That said, as we show in Section 7, our main conclusions are robust to shutting down this stochastic variation ($z_{ih} = \bar{z}_i$ for all h), although the

[9] We prefer this parsimonious specification with only two parameters to more general transition matrices that one could write, especially given the difficulty of pinning down these parameters from available micro data. In Section 7, we present results from alternative specifications.

model misses some features of wealth accumulation that we find important to match.

Labor Market Productivity. The specification of labor market productivity, w_{ih}, is fairly standard:

$$\log w_{ih} = \underbrace{\kappa_i}_{\text{permanent}} + \underbrace{g(h)}_{\text{life cycle}} + \underbrace{e_{ih}}_{\text{AR(1)}}, \tag{4}$$

for $h < R$, where κ_i is an individual fixed effect, $g(h)$ is a polynomial that captures the lifecycle component that is common to all individuals, and e_{ih} follows an AR(1) process:

$$e_{ih} = \rho_e e_{i,h-1} + \nu_{ih}, \tag{5}$$

with $\nu_{ih} \sim \mathcal{N}(0, \sigma_\nu^2)$. The permanent component, κ_i, is imperfectly inherited from parents:

$$\kappa_i^{\text{child}} = \rho_\kappa \kappa_i^{\text{parent}} + \varepsilon_{\kappa_i}, \tag{6}$$

where $\varepsilon_\kappa \sim \mathcal{N}\left(0, \sigma_{\varepsilon_\kappa}^2\right)$.

Individuals supply their labor services directly to the final good producer. The aggregate effective labor supply is

$$L = \int \left(w_{i,h(i)} \ell_{i,h(i)}\right) di, \tag{7}$$

where we expanded the subscript to clarify that $h(i)$ refers to the age of individual i in the current year, and $w_{i,h(i)}\ell_{i,h(i)}$ is a worker's efficiency-adjusted labor hours. We will suppress the dependence on i when it does not cause confusion. Therefore, for a given market wage rate per efficiency units of labor, $\overline{w}$, an individual's labor income is $y_{ih} = \overline{w} w_{ih} \ell_{ih}$.

3.2 Production Technology

Each active entrepreneur (i.e., those for whom $z_{ih} > 0$) produces a differentiated good according to a linear technology:

$$x_{ih} = z_{ih} k_{ih}, \tag{8}$$

where k_{ih} is the final good used in production by entrepreneur i at age h. The final good, Y, is produced according to a Cobb-Douglas technology:[10]

$$Y = Q^{\alpha} L^{1-\alpha}, \tag{9}$$

where Q is the CES composite of intermediate inputs:

$$Q = \left(\int x_{i,h(i)}^{\mu} di \right)^{1/\mu}. \tag{10}$$

To distinguish Q from the unadjusted capital stock, $K = \int k_{i,h(i)} di$, we refer to the former as the "quality-adjusted capital stock," since its level depends on the allocation of capital across entrepreneurs. Total Factor Productivity in the intermediate goods sector can be written as

$$\text{TFP}_Q = \frac{Q}{K}, \tag{11}$$

where TFP_Q captures the extent of misallocation of capital (see Appendix D for derivation).[11] In turn, equation (11) allows us to write total output as a Cobb-Douglas function of K and L: $Y = \text{TFP} \cdot K^{\alpha} L^{1-\alpha}$, where $\text{TFP} \equiv \text{TFP}_Q^{\alpha}$ is the aggregate TFP of the economy.

The final good producing sector is competitive, so the profit maximization problem is

$$\max_{\{x_{ih}\},L} \left(\int x_{i,h(i)}^{\mu} di \right)^{\alpha/\mu} L^{1-\alpha} - \int \left(p(x_{i,h(i)}) \times x_{i,h(i)} \right) di - \overline{w} L. \tag{12}$$

The first-order optimality conditions yield the inverse demand (price) function for each intermediate input and for the market wage:

$$p(x) = \alpha Q^{\alpha-\mu} L^{1-\alpha} x^{\mu-1} \qquad \overline{w} = (1-\alpha) Q^{\alpha} L^{-\alpha}. \tag{13}$$

[10] We introduce labor supply into (9) instead of (8). We made this choice to focus on entrepreneurs' saving and production decisions and abstract from other endogenous choices that introduce new channels and confound the analysis of the mechanisms we wish to better understand. This approach is in keeping with the bulk of the extant capital taxation literature that uses an entrepreneur- or capitalist-worker framework (e.g., Judd, 1985 and Straub and Werning, 2020). In Appendix F, we introduce labor supply through (8) and discuss the conditions under which this extension amplifies or dampens the effects of taxes relative to the baseline model.

[11] Specifically, equation (D.3) expresses TFP_Q in terms of producer-specific wedges that reflect the distortions in the allocation of capital.

3.3 Markets and the Government

Financial Markets. There is a bond market where intra-period borrowing and lending take place at interest rate r. This market has three important features. First, borrowed funds can only be used as capital in production; they cannot be used to finance consumption. Second, borrowing and lending take place before production but after z_{ih} is observed, so there is no uncertainty regarding an entrepreneur's ability to repay at the end of the period. Individuals with high entrepreneurial productivity relative to their private assets will choose to borrow to finance their business, whereas those with low productivity relative to their assets will find it optimal to lend for a risk-free return.

Third, borrowing is collateralized and is subject to a limit indexed to individuals' assets:

$$k_{ih} \leqslant \vartheta\left(\bar{z}_i\right) \times a_{ih}, \tag{14}$$

where $\vartheta\left(\bar{z}_i\right) \geqslant 1$ and $\vartheta'\left(\bar{z}_i\right) > 0$. When $\vartheta = 1$, individuals can use only their own assets in production; when $\vartheta = \infty$, they can borrow without a limit.[12]

We model the collateral constraints to be less stringent for higher ability entrepreneurs both because it seems to be a realistic feature and also because it is a more conservative assumption than a flat constraint $\left(\vartheta\left(\bar{z}_i\right) = \bar{\vartheta}\right)$, which implies larger welfare gains from wealth taxes (as we show later). In Section 7, we consider several alternative forms of borrowing constraints, including a version with unlimited borrowing subject to a credit spread, and a version with an aggregate corporate sector that is not subject to a borrowing limit.

Government and the Tax/Transfer Systems. In the benchmark "US economy," the government imposes flat taxes at rates τ_k on capital income, τ_ℓ on labor income, τ_c on consumption, and τ_b on bequests. We alternatively refer to this case as the "capital income tax" economy. In the alternative "wealth tax" economy, the government does not tax capital income (set $\tau_k = 0$) but imposes a flat-rate tax, τ_a, on beginning-of-period

[12] Imperfect enforceability of financial contracts can also generate a borrowing limit that is increasing in z. For example, Buera, Kaboski, and Shin (2011) assume that entrepreneurs can steal a fraction $1-\phi$ of the undepreciated part of capital plus revenues from production, $(1-\phi)\left((1-\delta)k + \mathscr{R}(zk)^{\mu}\right)$, and the only punishment is garnishment of wealth $(1+r)a$. Given that, and without taxation, they show that the borrowing limit simplifies to $\left((1+r) - \phi(1-\delta)\right)k \leqslant (1+r)a + \phi\mathscr{R}(zk)^{\mu}$. For a given a, this constraint allows higher-z entrepreneurs to borrow more. For our purposes, one drawback of this specification is that it involves prices, r and Q (and would also involve tax terms once introduced), thereby introducing new general equilibrium channels through which a tax policy works—by tightening or loosening the collateral constraints—that we wish to abstract from in this paper. We implement a simpler version of this constraint without prices in Section 7.

wealth, a_{ih}. We consider formulations with nonlinear taxes later. We denote a tax system as $\mathcal{T} \equiv (\tau_{\text{cap}}, \tau_\ell)$, where cap $\in (k, a)$.[13] The government runs a balanced budget and uses the tax revenues to fund social security pension payments to retirees and an exogenous, fixed level of government spending, G, which does not enter individuals' utility function (see Appendix A.2 for the equations).

The social security pension of a retiree ($t \geqslant R$) is determined according to the formula

$$y^R(\kappa, e_{R-1}) = \Phi(\kappa, e_{R-1})\,\overline{y}, \tag{15}$$

where $\overline{y}$ is the average labor income and $\Phi(\kappa, e)$ is the pension replacement rate function, which depends on a worker's permanent type κ and the persistent component of labor productivity at age $R-1$. The functional form of Φ is taken from the SSA's OASDI system—see Appendix A.1.

3.4 Individuals' Decision Problem

Every period, individuals make two sets of decisions: (i) the scale (k_{ih}) at which they operate their entrepreneurial business, which also determines how much they borrow or lend in the bond market, and (ii) the labor-leisure and consumption-savings decisions. The first problem does not interact with the second within a period, so it can be solved separately. The only dependence is through the appearance of a—which is predetermined—in the entrepreneur's borrowing constraint.[14]

I. Individual/Entrepreneur's Problem

For clarity, we suppress the subscripts i and h when possible. Every period, the individual/entrepreneur chooses the capital level to maximize profit:

$$\pi(a, z) = \max_{k \leqslant \vartheta(z) a} \{p(zk) \times zk - (r+\delta) k\}, \tag{16}$$

[13] Consumption and bequest taxes are not included in $\mathcal{T}$ because they will remain fixed throughout the analysis.

[14] Notice that we do not impose a fixed cost for being an entrepreneur, which is a common modeling strategy. This is for two reasons. The first is technical convenience and feasibility: the parameterized model delivers a thick Pareto tail all the way up to wealth levels about one million times higher than the average income in the economy ($100 billion), which makes it challenging to solve accurately. Eliminating a fixed cost here saves us computational complexity, which in turn makes it feasible to investigate richer tax policies that feature kinks or jumps. Second, as we will see, the process for z_{ih} we choose ensures that many individuals do not engage in entrepreneurial production, and among those who do, many earn a very small income from their businesses. As a result, the model will be able to reproduce important features of entrepreneurship in the US data. We believe these are acceptable trade-offs for our purposes, and we view our modeling strategy as a convenient alternative for modeling entrepreneurship.

where δ is the depreciation rate. The price of the differentiated good in (13) can be written as $p(zk) = \mathcal{R} \times (zk)^{\mu-1}$, where $\mathcal{R} \equiv \alpha Q^{\alpha-\mu} L^{1-\alpha}$, yielding the solution

$$k(a,z) = \min\left[\left(\frac{\mu \mathcal{R} z^{\mu}}{r+\delta}\right)^{\frac{1}{1-\mu}}, \vartheta(z)\,a\right], \tag{17}$$

with the associated maximized profit function

$$\pi(a,z) \equiv \begin{cases} \mathcal{R}\left(z\vartheta(z)\,a\right)^{\mu} - (r+\delta)\,\vartheta(z)\,a & \text{if } k(a,z) = \vartheta((z)\,a \\ (1-\mu)\,\mathcal{R} z^{\mu}\left(\frac{\mu\mathcal{R}z^{\mu}}{r+\delta}\right)^{\frac{\mu}{1-\mu}} & \text{if } k(a,z) < \vartheta(z)\,a \end{cases}. \tag{18}$$

II. Individual's Dynamic Programming Problem

Let $\mathcal{W}$ denote the post-production, after-tax wealth of the individual, which includes current-period profits plus interest income from the bond market:

$$\mathcal{W}(a,z;\tau_{cap}) \equiv \begin{cases} a + (\pi(a,z) + ra)(1-\tau_k) & \text{if } \tau_{cap} = \tau_k \\ a(1-\tau_a) + (\pi(a,z) + ra), & \text{if } \tau_{cap} = \tau_a \end{cases}. \tag{19}$$

It will also be convenient to define *total disposable resources* (after production and taxation):

$$\mathcal{Y}(a,\ell;z,e,\kappa;\mathcal{T}) \equiv \mathcal{W}(a,z,\tau_{cap}) + (1-\tau_{\ell})\,\overline{w}w(\kappa,e)\,\ell, \tag{20}$$

which the individual optimally splits between c and a'. Let $\mathbf{S} \equiv (\bar{z}, \mathbb{I}, e, \kappa)$ denote the vector of exogenous individual states and $s_{h+1} \equiv \phi_{h+1}/\phi_h$ be the conditional survival probability. The individual's dynamic problem is given by

$$\begin{aligned} V_h(a;\mathbf{S}) = &\max_{c,\ell,a'} u(c,1-\ell) + \beta\left[s_{h+1}E\left(V_{h+1}\left(a',\mathbf{S}'\right) \mid \mathbf{S}\right) + (1-s_{h+1})\,v(b)\right] \\ &\text{s.t. } (1+\tau_c)\,c + a' = \mathcal{Y}(a,\ell;z,e,\kappa;\mathcal{T}) \\ &\quad b = (1-\tau_b)\,a' \qquad \text{and } a' \geqslant 0. \end{aligned} \tag{21}$$

Retirees solve the same problem with labor income in $\mathcal{Y}$ replaced with pension income, $y^R(\kappa,e)$, with terminal condition $V_{H+1} \equiv 0$. The definition of a recursive competitive equilibrium is standard and is relegated to Appendix A.2.

4 Quantitative Analysis

In this section, we discuss the parameterization of the baseline model. There is a set of standard parameters whose calibration is straightforward. However, there are two features of the model that are crucial for the mechanisms we analyze later. These are the specifications of the entrepreneurial productivity process and financial constraints, so we discuss them in greater detail.

We also consider a second parameterization, which targets a lower level of wealth inequality—a top 1% share of 20%, versus 36% in the data and in the baseline—to gauge the sensitivity of our main conclusions along this dimension. In Section 7, we discuss a large number of additional robustness checks and extensions.

4.1 Model Parameterization

The model is calibrated to the US data. We first set the values of 9 parameters based on outside empirical evidence (reported in the top panel of Table II) and choose the remaining 11 parameters by targeting 11 data moments for the simulated model to match (bottom panel). As seen in Table III, the simulated model does a good job of matching the 11 targeted moments. We now discuss the targets and parameter choices in more detail.

Demographics. The model period is one year. Individuals enter the economy at age 20, retire at age 64 (model age $R = 45$), and quit entrepreneurial production when $\mathbb{I}_{ih} = 0$ is realized, which can happen at any age. The conditional mortality probabilities are taken from Bell and Miller (2002), and individuals die by age 100 (81 periods) with certainty.

Preferences and Technology. The utility function takes the Cobb-Douglas form:

$$u(c, \ell) = \frac{\left(c^{\gamma}(1-\ell)^{1-\gamma}\right)^{1-\sigma}}{1-\sigma}.$$

We set $\sigma = 4$, following Conesa, Kitao, and Krueger (2009), and target a K/Y ratio of 3 and 40 work hours per week (i.e., $\ell = 0.4$, assuming 100 weekly discretionary hours), which requires $\gamma = 0.445$ and $\beta = 0.9593$, given the other parameter values. The warm-glow bequest utility function takes the power form with the same curvature parameter as consumption:

$$v(b) = \chi \frac{(b + b_0)^{\gamma(1-\sigma)}}{1-\sigma}, \tag{22}$$

Table II – Parameters for the Benchmark Model

Parameter		Value
Parameters Calibrated outside of the Model		
Annual persistence for indiv. labor efficiency	ρ_e	0.9
Std. of innovations to indiv. labor efficiency	σ_ν	0.2
Interg. correlation of labor fixed effect	ρ_κ	0.5
Capital's share of output	α	0.4
Curvature of CES production function	μ	0.9
Depreciation rate	δ	0.05
Curvature of utility function	σ	4.0
Maximum (model) age	H	81
Retirement (model) age	R	45
Parameters Calibrated (Jointly) inside the Model		
Discount factor	β	0.9593
Consumption share in utility	γ	0.445
Strength of bequest motive	χ	0.20
Utility shifter for bequest ($)	b_0	26,800
(Controls) dispersion of labor fixed effect	$\sigma_{\varepsilon_\kappa}$	0.309
(Controls) dispersion of entrepr. ability	$\sigma_{\varepsilon_{\bar{z}}}$	0.277
Interg. correlation of entrepreneurial ability	$\rho_{\bar{z}}$	0.1
Productivity boost while in fast lane	λ	1.5
Annual transition rate of z_{ih}: $\mathcal{H}$ to $\mathcal{L}$	p_1	0.05
Annual transition rate of z_{ih}: $\mathcal{H}$ or $\mathcal{L}$ to 0	p_2	0.03
Slope of borrowing constraint schedule	φ	0.225

Notes: In addition to these parameters, survival probabilities, ϕ_h, are taken from Bell and Miller (2002) (omitted from the table). To keep the computational burden of the moment matching exercise feasible, the optimization was performed over a (fine) discrete grid for 4 parameters ($\rho_{\bar{z}}, \lambda, p_1, p_2$) and over a continuous space for the other 7 parameters.

as in De Nardi and Yang (2016). In this formulation, χ measures the importance of the bequest motive, and b_0 measures the degree to which bequests are luxury goods. When $b_0 > 0$ the marginal utility of bequests is bounded at $b = 0$ allowing some individuals to leave no bequests, consistent with the data. Following De Nardi and Yang (2016), we choose χ and b_0 to match a bequest-to-wealth ratio of 1.18 and the 90th percentile of the bequest distribution scaled by income, which is 4.31.[15]

We set $\alpha = 0.4$, which implies a labor share of 0.60, and set the depreciation rate to 5%. The curvature parameter of the CES aggregator, μ, is set to 0.9, which corresponds to a 10% markup over marginal cost. A higher μ implies less diminishing marginal returns to capital, which makes it easier to generate high inequality. The reverse happens as μ goes down, although the effects are mild down to a value of $\mu = 0.75$ or so. Beyond that point, the diminishing returns in entrepreneurial production become so strong that matching the right tail of the wealth distribution becomes impossible (and is no longer Pareto). We discuss the results for $\mu = 0.8$ in Section 7.

Tax System. The current US tax system is modeled as a quadruplet of (flat) tax rates: $(\tau_k, \tau_\ell, \tau_c, \tau_b)$. We set $\tau_k = 25\%$, $\tau_\ell = 22.4\%$, and $\tau_c = 7.5\%$, based on McDaniel (2007)'s calculations for the US economy between 2000 and 2003.[16] In the baseline analysis, we consider flat-rate taxes; later in the paper, we analyze nonlinear capital income taxes. In an earlier version of this paper, we allowed for progressive labor income taxes, but they did not change any of our substantive conclusions (Guvenen et al., 2019). Finally, we capture the US estate tax system with a 40% flat-tax rate on bequests.[17]

Labor Market Productivity. The deterministic lifecycle profile, $g(h)$, is a quadratic polynomial that generates a 50% rise in average labor income from age 21 to its peak at age 51. The AR(1) process has a persistence of $\rho_e = 0.9$ and a standard deviation of $\sigma_e = 0.2$, consistent with the estimates in the literature when a separate transitory shock is not modeled. The intergenerational correlation of the fixed effect is set to $\rho_\kappa = 0.5$ (Solon, 1999). With these parameters fixed, we set $\sigma_{\varepsilon_\kappa} = 0.309$ to match a cross-sectional standard deviation of log labor earnings of 0.80 (Guvenen, Karahan, Ozkan, and Song, 2021). Finally, $\overline{y}$ is set to \$90,000 to map model wealth levels into data counterparts.

[15] De Nardi and Yang (2016) report a range from 0.88% to 1.18% for the bequest/wealth ratio in the United States. Hendricks (2001) reports a similar figure, 1%, using Survey of Consumer Finances (SCF) data. An earlier estimate by Auerbach et al. (1995) for the bequest/GDP ratio is 3.6%, which translates to a 1.2% bequest/wealth ratio in our model (given $K/Y = 3$).

[16] McDaniel (2007)'s definition of capital income tax is based on capital income revenue, which is the sum of (i) taxes levied on corporate income, (ii) taxes paid by households on dividend income, capital gains, and the capital share of private business income, and (iii) property taxes paid by entities other than households. Property taxes paid by households are not included in the capital income tax but rather in the consumption tax. The labor income tax is based on tax revenues from taxes imposed on the compensation of employees (wages and salaries) and the labor share of income earned by the self-employed.

[17] Although we do not target tax revenue statistics directly, the model comes fairly close. For example, the share of aggregate tax revenues in US GDP was 25% in 2019, and the share of capital tax revenues in total tax revenues is 28% (OECD, 2021). The model counterparts are 26% and 24%, respectively.

TABLE III – Targeted and Untargeted Moments: Model versus Data

	Data	Benchmark	Low-Inequality Calibration (L-INEQ)
TARGETED MOMENTS			
Standard deviation of log earnings	0.80	0.80	0.80
Capital-to-output ratio	3.00	3.00	3.00
Average labor hours	0.40	0.40	0.40
Bequest/wealth	0.012	0.012	0.012
90th percentile of bequest distribution	4.31	4.10	6.60
Intergenerational corr. of return fixed effect	0.10	0.10	0.10
Top 1% wealth share	0.36	0.36	0.20†
Self-made billionaires (fraction)	0.54	0.56	0.26
Population share of entrepreneurs in top 1%	0.65	0.68	0.68
Wealth share of entrepreneurs	0.42	0.39	0.34
Business debt plus external funds/GDP	1.52	1.50	1.50
UNTARGETED MOMENTS			
Wealth Gini	0.82	0.78	0.66
Top 0.1% wealth share	0.15	0.23	0.10
Pareto tail index	1.52	1.51	1.81
Intergenerational correlation of wealth	0.16	0.16	0.21

Notes: †The low-inequality calibration (L-INEQ) targets a top 1% share of 20% as an alternative benchmark with lower wealth concentration.

Entrepreneurial Productivity. The stochastic process for z_{ih} is governed by five parameters —ρ_z, $\sigma_{\varepsilon_{\overline{z}}}$, λ, p_1, and p_2—that jointly determine the features of the steady state wealth distribution and the lifecycle dynamics of wealth accumulation and entrepreneurship (given the other parameters of the model). We choose these five parameters to match five empirical targets: (i) the intergenerational correlation of individual fixed effects in rates of return (0.1), (ii) the wealth share of the top 1% (36%), (iii) the share of *Forbes* 400 billionaires who are self-made (54%), (iv) the population share of entrepreneurs in the top 1% (65%), and (v) the wealth share of entrepreneurs (41.6%).[18]

The model counterpart of a self-made billionaire is an individual who is among the top 400 wealthiest in our simulated US economy (whose minimum wealth is $1.9 billion, compared with $2 billion in the 2017 *Forbes* 400 list) and who started life with less than

[18]The empirical target for (i) is from Fagereng et al. (2020), (ii) is from Vermeulen (2018), (iii) is from *Forbes*, and (iv) and (v) are from Cagetti and De Nardi (2006).

$250,000 of wealth.[19] The latter criterion is likely to be a generous upper bound for someone coming from an upper-middle-class family. If we use a $100,000 cutoff instead, the fraction of self-made in the model is 50%, versus 56% under the original definition. Notice that these cutoffs imply a lower bound of 8,000-fold to 20,000-fold wealth growth over the life cycle. It is not possible to generate this incredible speed of wealth accumulation in models of wealth inequality that rely on idiosyncratic earnings shocks or heterogeneity in patience.[20]

The model matches the fraction of self-made without overstating the top 1% share or overall wealth inequality (Gini is 0.78 versus 0.82 in the data; Wolff, 2006), thanks to the stochastic variation in z_{ih}, as specified in equation (3). In particular, the "fast lane" state early in life generates extremely fast wealth growth at the top, while the transitions to the $\mathcal{L}$ and 0 states prevent this fast growth from being too persistent and therefore overstating overall inequality.

The last two moments we target pertain to entrepreneurship. The definition of an entrepreneur in the data is not as clear-cut as that of a worker, because an individual may be self-employed but not own a business (or use any capital), or may own a private business but not manage it, or manage the business but not own it, and so on. Using data from the US Survey of Consumer Finances, Cagetti and De Nardi (2006) report statistics on both entrepreneurship and wealth inequality based on combinations of different criteria (self-employed, owns business, manages business, etc.). In our model, entrepreneurs own and manage their business, which is most similar to their definition of "active business owners." As for the model counterpart, first note that with the parameter choices for p_1 and p_2, 53.5% of individuals are in the $z_{ih} = 0$ state and therefore have no entrepreneurial production or income. Among the remaining 46.5%, the vast majority have low enough z_{ih} that their business income is fairly marginal. To get a definition of "entrepreneur" that is more comparable to an active business owner, we take individuals who earn the majority of their income from their business, as opposed to from wages and interest income.

With these definitions, the model jointly matches the wealth share of entrepreneurs (38.6% vs 41.6% in the data) and the population share of entrepreneurs within the top 1% wealthy (68.1% vs. 65% in the data) as well as their share within the top 10%,

[19]We follow *Forbes*'s definition of self-made as someone on their list who comes from, at most, an upper-middle-class family. Details of the *Forbes* classification are in Appendix B, Table B.1.

[20]Even with calibrations that match the top 1% share in steady state, it takes dynasties hundreds of years to get there starting from the median initial wealth level. This is the case, for example, in stochastic-beta models à la Krusell and Smith (1998) as well as in models with "awesome-state" idiosyncratic earnings shocks à la Castañeda, Díaz-Giménez, and Ríos-Rull (2003).

which was not targeted (40% vs. 42% in the data).[21] Although the model understates the population share of entrepreneurs we targeted (7.1% vs 11.5% in the data), other datasets and slightly different definitions yield lower empirical estimates.[22] Finally, the model generates a lifecycle pattern of entrepreneurship that is hump-shaped (with the fraction of entrepreneurs rising until late 30s and then tapering and declining in the 50s) as has been documented in the data (see Figure C.2 in Appendix C).[23]

Financial Constraints. Because of the importance of financial frictions for our analysis, we will consider several different specifications, including alternative functional forms for the collateral constraint and allowing unlimited borrowing for a subset of firms or for all firms subject to a credit spread, among others. We begin by describing the baseline specification, and discuss other formulations later.

We assume that the lowest-ability group, $\bar{z}_0$, cannot borrow at all, and the borrowing limit increases linearly with ability from there on: $\vartheta(\bar{z}) = 1 + \varphi(\bar{z} - \bar{z}_0)$.[24] We choose the slope parameter, $\varphi = 0.225$, so as to match the ratio of aggregate business borrowing to GDP, which is 1.52 in the US data. With this choice, an entrepreneur at the 90th percentile of $\bar{z}$ distribution can borrow up to 92% of her wealth.[25]

The business debt to GDP ratio can also be estimated from firm-level data. Using this approach, Asker, Farre-Mensa, and Ljungqvist (2011) report an average debt-to-asset ratio of 0.20 for publicly listed firms and a ratio of 0.31 for private firms in the United States. With a capital/output ratio of 3, and assuming as an upper bound that all capital stock

[21]Taking a higher business income cutoff of 90% (as opposed to 50%) reduces the wealth share of entrepreneurs from 38.6% to 30.5% but barely changes their population share within top 1% (67.3% vs. 68.1%), reflecting the high concentration of entrepreneurs at the top.

[22]For example, Cagetti and De Nardi (2006) consider a definition that adds the condition that an active business owner also reports being self-employed (and use it as their benchmark). This group's population share is 7.6% and wealth share is 33%, not too far from our benchmark. Using this latter definition and data from the Panel Study of Income Dynamics, Salgado (2020) reports a population share slightly above 7% in 2014. Gentry and Hubbard (2000) define an entrepreneur as an active business owner with at least $5,000 in assets and report a population share of 8.7% and a wealth share of 39% for entrepreneurs, figures similar to what our model generates. So, overall, our calibrated model is well within the range of figures reported in the literature.

[23]See, e.g., Kelley, Singer, and Herrington (2011) and Liang, Wang, and Lazear (2018). Although z_{ih} is higher at young ages, business income is typically not, because many individuals start life with low wealth, which in turn implies tight collateral constraints, limiting their productive capacity despite a high z_{ih}.

[24]In practice, $\bar{z}_0$ corresponds to the lowest point in the z grid and contains 0.6% of the population.

[25]We calculate the aggregate business borrowing as the sum of nonfinancial business liability ($22.79 trillion) and the capitalized value of external funds raised through IPOs and equity issues by US nonfinancial businesses ($4.14 trillion) for 2015. The first figure is taken from the flow of funds accounts (Federal Reserve Statistical Release, 2015Q3, Table L.102), and the second is the capitalized value at 5% of the $197.5 billion annual flow reported for 2015 by the US Department of the Treasury (2017). Dividing this sum by the US GDP of $17.65 trillion for the same year yields a ratio of 1.52.

is owned by firms, these leverage figures correspond to a debt-to-GDP ratio between 0.60 and 0.93, implying significantly tighter financial constraints than those in our benchmark. The bottom line is that our baseline model allows significant amounts of borrowing by US firms, which are likely to be on the upper end of the empirical estimates.

Alternative Parameterization: Targeting Lower Inequality

We consider an alternative calibration (hereafter, L-INEQ) that targets a much lower level of wealth inequality—specifically, a top 1% share of 20%, versus 36% above. One reason we do so is to address a possible concern that the model could be attributing part of top inequality stemming from other mechanisms to the features present in the model (specifically, persistent return heterogeneity, birth/death process, and bequests), which may in turn skew the costs and benefits of certain tax systems.[26] While the existing evidence indicates that return heterogeneity is responsible for the bulk of top-end inequality,[27] we find it useful to analyze this alternative calibration. As Table III shows, the model matches moments unrelated to wealth inequality well, and it understates all measures of wealth concentration by design. In the rest of the paper, we report the results from the L-INEQ calibration when they are relevant and describe the remaining results in Appendix E.1.

4.2 Performance of the Benchmark Model

In this section, we further assess the performance of the benchmark model for some important untargeted data moments and features, including the right tail of the wealth distribution, the distribution of individual rates of return, and the extent of capital misallocation, among others.

Wealth Inequality. A thick Pareto tail essentially means that the model can generate extremely wealthy individuals (e.g., holding \$100M or more), who are a key source of capital tax revenue as well as being the focal point of some recent wealth tax proposals. While the calibration targeted the top 1% share, this does not guarantee a Pareto tail or its thickness.

Figure 1a plots the log counter-CDF of wealth against log wealth, which should be a straight line with a slope of $-\alpha$ if wealth has a Pareto distribution: $\mathsf{P}(\omega > x) \sim x^{-\alpha}$. The blue circles are the US data (from Vermeulen, 2018), which essentially form a straight

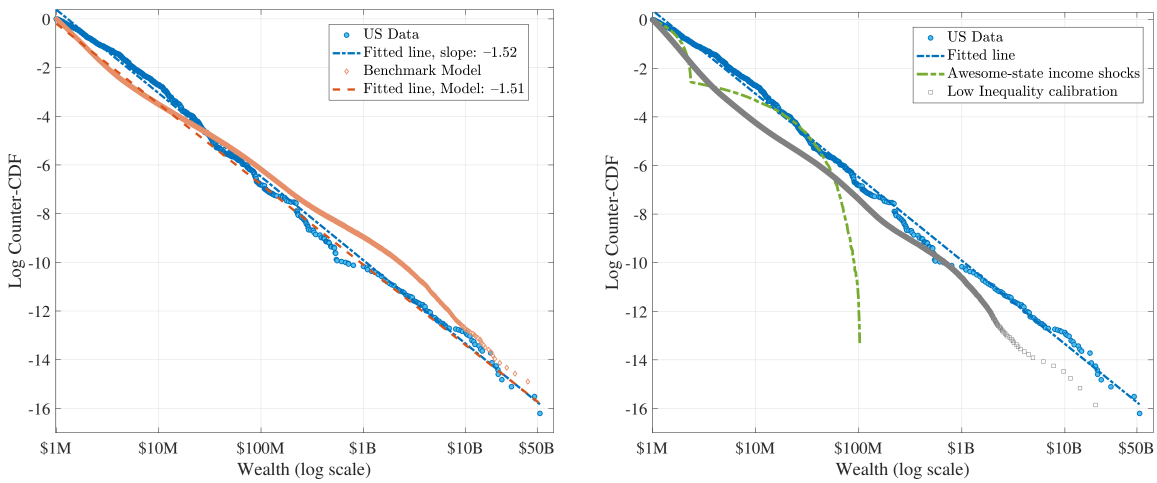
[26] Of course, this is true for every quantitative analysis, since all models are misspecified and omit other key mechanisms.

[27] For example, Bach, Calvet, and Sodini (2020) attribute 80%-90% of changes in top-end wealth inequality in Sweden to return heterogeneity.

Figure 1 – Pareto Tail of Wealth Distribution: US Data vs. Models

(a) Benchmark Model

(b) Alternative Models



Note: Both axes are in natural logs. The x-axis ticks are placed at powers of 10 for readability.

line all the way up to $50 billion, with a slope of −1.52, confirming a thick Pareto right tail. The model counterpart (orange diamonds) aligns well with the data, especially below $500 million and above $10 billion, with a slight overestimation between these points. The fitted line has a slope of −1.51, even though this moment was not targeted. That said, the top 0.1% share is somewhat higher in the model than in the data (23% versus 16%), although the robustness checks in Section 7 show that this does not affect our substantive conclusions.[28]

To further this point, the gray square markers in Figure 1b show the counterpart for the (low-inequality) L-INEQ model, which lies below the data everywhere, but still displays a nearly perfect Pareto tail, albeit with a thinner tail of (slope of −1.81 versus −1.52), and a top 0.1% share of 10% versus 16% in the data (as intended). We will analyze its implications for taxation below. Finally, for comparison, we solved and calibrated a Castañeda, Díaz-Giménez, and Ríos-Rull (2003)-style model with "awesome-state" labor income shocks to match the top 1% share (as well as some other details—see Appendix E.2). As seen in the figure, the tail is not Pareto, and the richest person in the simulated economy has about $100M in wealth—or 500 times less than in the data.

The baseline model is consistent with some other key facts about wealth and capital income, which we briefly discuss here and present in detail in Appendix B.1. For example, capital income is more concentrated than wealth in the model and the US data, and the

[28]The 16% figure is from Smith, Zidar, and Zwick (forthcoming, Table 1).

magnitudes broadly align as well.[29] The model also generates a steeper age profile for capital income than in labor income—a 2.9-fold raise between ages 25 and 50 versus a 50% rise for labor income consistent with the US data (Piketty, Saez, and Zucman, 2018). Another important statistic is the intergenerational correlation of wealth, which matters for the effects of taxation and depends on many factors (parent-child correlation in abilities, bequests, estate taxes, retirement, other motives for savings, etc.). Figure C.3 in Appendix C shows the rank-rank plot of parent-child wealth in the model (left panel), whose shape and magnitudes match up well with the data (right) from Fagereng et al. (2020, Figure 11). Further, the 90th-10th percentile gap in wealth of parents translates into a 16 percentile gap in wealth among their children in both the data and the model.

Rate of Return Heterogeneity. Rate of return heterogeneity is the key ingredient behind the results in this paper, so we examine whether our calibration implies a plausible distribution for returns. As noted earlier, the availability of individual return data is very recent, and the most detailed analysis comes from Norway (Fagereng et al., 2020) with some less detailed statistics from the US data (Smith, Zidar, and Zwick, forthcoming). In Table IV, we report statistics about the dispersion of annual and long-run (persistent) individual returns as well as the level of returns at the top (which is critical for the right tail). Overall, the benchmark model compares well with the data for the overall population (top row) from Fagereng et al. (2020), with all statistics on dispersion and top returns within one or two percentage points of their empirical counterparts.[30] For comparison, the last row shows the L-INEQ calibration, which, as expected, displays a smaller dispersion and, more importantly, lower returns at the very top of the distribution of the persistent component of returns than those in the benchmark and the data.

The second row reports statistics that are available for business owners, which are, again, either in line with or higher than those in the benchmark model. The third row shows dispersion statistics of annual returns on investment for private firms from the US tax data (reported recently by Smith, Zidar, and Zwick, forthcoming), which are almost

[29]For example, the top 0.1% share by capital income varied between 30% and 41% since 2000 according to Saez and Zucman (2016, fig. 3); the corresponding figure in the model is 34.4%. Similarly, the top 1% share of capital income is 52% in the model. The closest statistic we are able to find is by Smith, Zidar, and Zwick (2021, Table A5), who report shares sorted by individual components of capital income. The top 1% shares for interest, dividend, and capital gains income have all been above 60% since 2000.

[30]The exception is the extremely high kurtosis reported by these authors, which indicates very long tails in individual returns. We do not find this concerning, both because this statistic is very sensitive to outliers (e.g., Kim and White, 2004, show that a small number of outliers in financial returns data can cause severe overestimation of kurtosis) and also because, even if accurate, it simply says the model does not rely on extreme tail events to match the top wealth inequality and other statistics. We would be more worried if the opposite were true.

Table IV – Distribution of Rates of Return (Untargeted) in the Model and the Data

	Annual Returns			Persistent Component of Returns					
	Std dev	P90-P10	Kurtosis	Std dev	P90-P10	Kurtosis	P90	P99	P99.9
Data (Norway)	8.6	14.2	47.8	6.0	7.7	78.4	4.3	11.6*	23.4*
Data (Norway, bus. own.)	–	–	–	4.8	10.9	14.2	10.1	–	–
Data (US, private firms)	17.3	32.9	8.5	–	–	–	–	–	–
Benchmark Model	8.4	17.1	7.6	4.1	9.2	6.1	5.8	13.9	19.7
L-INEQ Calibration	6.7	13.1	9.2	3.8	9.2	4.3	5.6	11.2	15.8

Notes: Returns on wealth in percentage points. All cross-sectional returns are value weighted. *The statistics for Norway are for individual returns on wealth (net worth) taken from Fagereng et al. (2020), with the exception of P99 and P99.9, which have been kindly provided to us by the authors. The US statistics are from Smith, Zidar, and Zwick (forthcoming, Table B.2) and are for S-corps' returns on investment; the authors also report statistics for partnerships, which are very similar (std dev of 19.0% and P90-P10 of 35.6). For each individual, the persistent component of returns is calculated as the unweighted average of annual, before-tax, returns between ages 25 and 75, after taking out the average return by age following Fagereng et al.

twice as high as in the model. This should not come as a surprise, given that these are returns on a narrower definition of investments with high returns. (Unfortunately, statistics on long-run returns are not available.) Our main takeaway from these results is that the return distribution generated by the model seems quite plausible and in line with the data, even though none of these statistics were targeted.

Finally, how much misallocation does the baseline model generate? A widely used statistic for misallocation (Hsieh and Klenow, 2009) measures how much lower the TFP level is relative to its efficient counterpart—the model once all frictions are eliminated. This statistic is 0.16 for the baseline model and 0.11 for the L-INEQ calibration, indicating lower misallocation than that in the US data reported by Bils, Klenow, and Ruane (2017). Although some authors have argued for an even lower value for this statistic (Midrigan and Xu, 2014), in Section 7, we consider alternative parameterizations and assumptions that deliver a statistic as low as 0.05 and show that our conclusions about wealth taxes are not very sensitive to the precise value of misallocation.

5 Tax Reform

In this section, we analyze the effects of a tax reform that replaces a capital income tax (setting $\tau_k = 0$) with a flat-rate wealth tax, τ_a, while keeping all other taxes fixed. In Section 6, we will conduct an optimal tax analysis in which the labor income tax is chosen jointly with τ_a or τ_k to maximize welfare. Compared with that analysis, the tax reform

Table V – Tax Reform: Change in Macro Variables from Current US Benchmark

	Tax Reforms: Change from US Benchmark									
	Quantities (% Change)						Prices (Change)			
	K	Q	TFP_Q	L	Y	C	$\overline{w}$	$\overline{w}$ (net)	$\Delta r^\dagger$	$\Delta r^\dagger$ (net)
RN reform	16.4	22.6	5.3	1.2	9.2	9.5	8.0	8.0	0.21	–0.36
BB reform	9.2	16.0	6.2	1.2	6.9	7.7	5.6	5.6	0.67	–0.38

Notes: RN and BB refer to the revenue-neutral and balanced-budget reforms, respectively. Percentage changes are computed with respect to the benchmark economy, which has $\tau_k = 25\%$ and $\tau_a = 0\%$. †Changes in the interest rate are reported in percentage points. The net wage is defined as $(1 - \tau_\ell)\, w$, and the net interest rate is defined as $(1 - \tau_k)\, r$ or $r - \tau_a$, depending on the model. The TFP variable is measured in the intermediate goods market.

we study here serves two important purposes. First, it is a simpler experiment in that (i) it does not rely on the social objective function maximized, and (ii) by keeping other taxes fixed, it allows us to focus on the trade-offs between the two capital taxes in isolation of other mechanisms that would become operational when, for example, τ_ℓ were also adjusted. Second, its relative simplicity makes it more appealing from a policy perspective than an optimal policy that requires changes to several tax tools simultaneously.

To make the comparison meaningful, we need to impose a neutrality condition. While revenue neutrality is an obvious choice, it raises a subtle issue: pension payments are anchored to average earnings ($\overline{y}$ in eq. 15), so a reform that changes $\overline{y}$ also changes SSP, violating the budget balance if revenue is kept constant. To deal with this issue, we consider two cases. The first is our main "revenue-neutral" (RN) reform, in which we keep the dollar value of pension income of every individual i at every age fixed at its US benchmark level. The second is the "balanced-budget" (BB) reform, in which we let pensions scale with $\overline{y}$ according to (15), while picking τ_a to balance the government budget. Except where we note explicitly, the results we present below pertain to the RN reform.

5.1 Results

The RN reform requires a wealth tax of $\tau_a = 1.19\%$ to generate the same revenue as the benchmark US economy (with $\tau_k = 25\%$).[31] The BB reform requires a slightly higher rate, $\tau_a = 1.67\%$, mainly because of the added cost of higher pensions. A glance at the left

[31]The corresponding tax rate in the L-INEQ model is $\tau_a = 1.46\%$, largely because there is less extreme wealth to tax at the top.

panel of Table V shows that aggregate quantities increase across the board with the switch to a wealth tax. In the RN reform, K and Q are higher by 16.4% and 22.6%, respectively, and the K/Y ratio rises from 3 to 3.2. The larger increase in Q relative to K reflects the improved reallocation of capital due to the wealth tax. This improvement in efficiency can be expressed as a 5.3% increase in TFP in the intermediate goods sector, which increases aggregate TFP by 2.1%.

Furthermore, L and $\overline{w}$ are higher by 1.2% and 8% respectively, clearly showing that the 9.2% rise in output is accounted for primarily by the higher Q, not L. Finally, the after-tax net interest rate falls by about 36 basis points since wealth taxes erode the principal, and the rise in the before-tax interest rate is too small (21 basis points) to offset the principal loss. The results for the BB reform are qualitatively the same (also in Table V). Quantities increase slightly less than in the RN reform, owing to the slightly higher tax rate, with the exception of TFP.

Turning to distributional outcomes, wealth inequality is higher under the wealth tax (in both the RN and BB cases), as anticipated from the illustrative example in Section 2, with the top 1% share rising in the RN reform from 36% to 43% and the top 10% share rising from 66% to 71%. Inequality in labor income remains virtually unchanged, which is not surprising given the very small hours response to the reform. We will discuss the changes in consumption and leisure inequality later.

We should note that the total revenue raised from capital goes from 6.1% of GDP in the benchmark US economy with capital income taxation to 3.9% under the wealth tax. This is because the level of capital revenues is kept fixed in the reform, while wages (and hence labor tax revenues) are higher and GDP is larger in the reform economy. So, there is no presumption that a wealth tax needs to raise the tax burden on capital.

5.2 Welfare Analysis

To quantify the welfare effects of the reform, we use two measures. The first one, CE_1, is constructed at the individual level, which allows us to quantify the gains/losses experienced by different groups. For an h-year-old in state $\mathcal{S} \equiv (\mathsf{a}, \mathbf{S})$ in the US benchmark, $\mathsf{CE}_1(\mathsf{h}, \mathcal{S})$ is the fixed proportional consumption transfer at all future dates and states that makes her indifferent between the stationary equilibria of the two economies. We can calculate welfare change measures for different groups by integrating $\mathsf{CE}_1(\mathsf{h}, \mathcal{S})$ over the stationary distribution of the group in the US benchmark.[32] We use $\overline{\mathsf{CE}}_1$ to denote the aggregated

[32]Using the stationary distribution of the "reform" economy makes little quantitative difference.

Table VI – Average Welfare Gain from Tax Reform

	RN reform	BB reform	RN reform (L-INEQ)
$\overline{\mathsf{CE}}_1$	6.8	4.8	4.9
$\overline{\mathsf{CE}}_2$	7.2	4.3	4.8
% with welfare gain	67.5	94.4	63.8

Notes: The welfare figures report the percentage gain in consumption-equivalent terms from each tax reform relative to the current US benchmark economy. All numbers reported in the table are in percentage points.

measure for newborn individuals. The second one, $\overline{\mathsf{CE}}_2$, is a macro measure, à la Lucas (1987): it is the fixed proportional consumption transfer (at all dates and states) to all newborn individuals in the US benchmark such that average lifetime utility is equal to that in the tax-reform economy. This measure provides a single figure that is easy to interpret and allows comparison with some previous work. The exact formulas for the calculations described here are in Appendix A.3.

Results. The average welfare gains from the tax reforms are large: newborns in the US benchmark would have to be compensated on average by 6.8% ($\overline{\mathsf{CE}}_1$) of their consumption to be indifferent with the RN reform economy, and by 7.2% using the $\overline{\mathsf{CE}}_2$ measure (Table VI). While these large gains could be partly anticipated from the large increase in average consumption and the little change in labor hours we saw in Table V, welfare calculations also account for the changes in the cross-sectional distributions, which do not seem to dampen the level gains (more on this in the next section).

How are the welfare gains distributed across the population? In Table VII, we divide the population into five age groups—age 20, 21–34, 35–49, 50–64, and retirees—and six bins for entrepreneurial ability that become finer at the top end. Each cell reports the corresponding average welfare change. There are several takeaways. First, all newborn groups gain from the RN reform, and these gains are fairly evenly distributed across ability groups—ranging from 6.7% for the lowest 40% to 13.4% for the top 0.1% $\bar{z}_i$-group.[33] Second, welfare gains decline with age, which is to be expected: since wealth rises and productivity (z_{ih}) falls on average with age, the ratio of capital income to wealth falls, in turn raising the tax burden of wealth taxation relative to capital income taxation. This

[33] Clearly, some subjectivity is involved in judging how even a distribution is. What we have in mind is the comparison—discussed further in the next section—between optimal wealth and capital income taxes, in which the latter generates gains that are much more skewed toward the top end.

Table VII – Welfare Gain/Loss by Age Group and Entrepreneurial Ability

	Entrepreneurial Ability Groups ($\bar{z}_i$ Pctiles)					
	0–40	40–80	80–90	90–99	99–99.9	99.9+
Age groups	RN Reform					
20 (newborn)	6.7	6.3	6.8	8.5	11.5	13.4
21–34	6.3	5.5	5.5	6.5	8.5	9.7
35–49	4.9	3.8	3.3	3.3	3.1	2.8
50–64	2.2	1.5	1.1	0.9	0.4	–0.2
65+	–0.2	–0.3	–0.4	–0.4	–0.7	–1.0
	BB Reform (SS Pensions Adjusted)					
20 (newborn)	4.7	4.2	4.8	6.7	10.3	12.5
21–34	4.5	3.7	3.7	5.2	8.0	9.6
35–49	4.2	3.0	2.6	2.9	3.1	2.9
50–64	4.6	3.8	3.2	2.9	2.0	1.1
65+	6.2	5.8	5.4	4.7	3.4	2.3

Notes: Each entry reports the average welfare gain or loss (CE_1) from the RN and BB wealth tax reforms relative to the current US benchmark for individuals in each age and entrepreneurial ability group. Averages are computed with respect to the US benchmark distribution.

effect partially offsets the income gains from higher wages under wealth taxation, resulting in declining welfare gains by age. Despite this decline, the welfare change is positive for all working-age groups except those with the highest ability (the top 0.1%) whose losses from higher taxes on their large wealth outweigh their gains from higher wages.

That retirees lose from the RN reform is not surprising: by design, their pensions remain fixed at the US benchmark, so they do not share the wage gains experienced by workers, yet the tax obligation on their accumulated wealth is higher after the reform. The BB tax reform alleviates this problem by indexing pensions to average wages. The average welfare gains (Table VI) are lower than in the RN reform—simply because more revenue needs to be raised to pay for higher pensions—but still significant: 4.3% to 4.8%, depending on the welfare measure. On the flip side, now all retiree groups gain significantly from the reform (lower panel of Table VII). Overall, 68% of the population experiences a welfare gain under the RN reform; this fraction jumps to 94% under the BB reform.

How sensitive are these welfare gains to the amount of wealth inequality generated by return heterogeneity? The L-INEQ calibration that targets almost half the top-end inequality in the data (20%) provides an answer. As seen in Table VI, the welfare gain in an RN reform is about 4.8%–4.9%, and about 64% of the population gains, only slightly lower than the benchmark. The welfare change distribution shows the same patterns as the benchmark RN reform with smaller magnitudes but the same substantive conclusions (see Table E.5).

Taking Stock. Let us summarize the main conclusions of the tax reform analysis. First, a wealth tax can raise the same revenue as a capital income tax, with less distortion. In particular, it reduces the misallocation of capital through the use-it-or-lose it effect (and the endogenous savings response it triggers), yielding higher aggregate productivity, average wages, consumption, and welfare. Second, welfare gains are relatively evenly distributed, with all newborn groups preferring the wealth tax economy. Third, allowing pensions to rise with average labor income (BB reform) yields lower average welfare gains but spreads the gains to the vast majority of the population. Fourth, the wealth tax reform delivers smaller but robust welfare gains even when the model is calibrated to generate significantly lower top-end wealth inequality.

These results are silent on whether either tax is desirable at all when the government can adjust the level of other taxes. We address this question of optimal taxation next.

6 Optimal Taxation

In this section, we study the optimal taxation problem of a government that chooses a combination of tax instruments to maximize the ex ante lifetime utility of an individual born into the stationary equilibrium implied by the chosen tax policy, subject to the constraint that it raise enough revenues to pay for $G+SSP$ as before.[34] In the main exercise, the government chooses flat-rate taxes on wealth and labor income. We refer to this as the wealth tax economy (WT) and to its optimum as the optimal wealth tax economy (OWT). For comparison, we consider a second exercise in which the government chooses flat-rate taxes on capital income and labor income—the capital income tax economy (KIT) and its optimum (OKIT).[35] We also consider an extension that introduces progressivity into the wealth tax system through an optimally chosen exemption level, $\underline{a}^*_{\text{ex}}$, and a tax rate τ_a for $a > \underline{a}^*_{\text{ex}}$ (OWT-X).

[34]Specifically, the maximized objective is $\sum_{\mathcal{S}} \left[\Gamma^{\text{OPT}}(1, \mathcal{S}) \times V_1^{\text{OPT}}(c^{\text{OPT}}(1, \mathcal{S}), \ell^{\text{OPT}}(1, \mathcal{S}))\right]$, where Γ^{OPT} is the stationary distribution, and the superscript OPT refers to the relevant optimal tax economy.

[35]A consumption tax is also a tax on labor income. Thus, in line with the literature on optimal taxation, we focus on the trade-offs between capital income (or wealth) and labor income taxes.

TABLE VIII – Optimal Taxation: Tax Rates and Average Welfare Effects

	Benchmark US Economy	RN Reform	OWT	OWT L-INEQ	OWT-X	WTE-X	OKIT
		(1)	(2)	(3)	(4)	(5)	(6)
Tax Rates							
τ_k	25.0	—	—	—	—	—	–13.6%
τ_a	—	1.19	3.03	2.54	3.80[†]	3.30	—
τ_ℓ	22.4	22.4	15.4	18.1	14.4	17.7	31.2
ΔWelfare							
$\overline{CE}_1$	—	6.8	9.0	6.0	9.1	8.4	4.2
$\overline{CE}_2$	—	7.2	8.7	5.2	8.8	8.6	5.1

Notes: Percentage changes are computed with respect to the US benchmark economy calibrated in Section 4. [†]The optimal wealth threshold, $\underline{a}^*_{ex}$—below which $\tau_a = 0$—is equal to $0.3 \times \overline{y}$. In experiment WTE-X, we set the exemption level to 100% of $\overline{y}$. Gains for the whole population are as follows: $CE_2(pop)$: 4.77, 4.31, 2.11, 4.68, 6.18, 4.50.

In Sections 6.1 to 6.3, we compare the stationary equilibria of OWT and OKIT with each other and with the US benchmark. We conduct a full transition analysis in Section 6.4.

Overview of Results. First, OWT combines a high wealth tax with a low labor tax: $\tau_a = 3.03\%$ and $\tau_\ell = 15.4\%$ (Table VIII). Second, the progressive optimal wealth tax system (OWT-X) combines a fairly low threshold—$\underline{a}^*_{ex} = 0.3\overline{y}$—with a higher wealth tax and a lower labor tax: $\tau_a = 3.80\%$ and $\tau_\ell = 14.4\%$. While the average welfare gain is only marginally higher than in OWT, the threshold exempts 32% of the population, which improves distributional outcomes. We also consider higher threshold levels up to 100% of average earnings (WTE-X* in column 5), which yield slightly lower average welfare gains but deliver other distributional benefits, as we discuss below. Third, OWT in the L-INEQ calibration (column 3) is a muted version of the baseline OWT, with a slightly lower wealth tax and higher labor tax: $\tau_a = 2.54\%$ and $\tau_\ell = 18.1\%$. This pattern of quantitative differences with the same substantive conclusions for the baseline and L-INEQ calibrations will be a recurring theme in the results below (as was also the case in the tax reform above).

Turning to OKIT (last column), it provides a *subsidy* to capital income and imposes a high labor tax: $\tau_k = -13.6\%$ and $\tau_\ell = 31.2\%$, which seems surprising at first blush. Whereas OWT shifts the tax burden from labor to capital, OKIT does the opposite—it

Table IX – Optimal Taxation: Changes in Macroeconomic Outcomes

	Change from Benchmark (%)			
	RN Reform	OWT	OWT-X	OKIT
K	16.4	2.6	–3.0	38.6
Q	22.6	10.5	5.4	46.1
L	1.2	3.3	3.3	–1.0
Y	9.2	6.1	4.1	15.7
K/Y	6.7	–3.3	–6.9	19.8
TFP_Q	5.3	7.7	8.7	5.4
C	9.5	7.9	6.3	13.9
$\overline{w}$	8.0	2.8	0.8	16.8
$\overline{w}$ (net)	8.0	12.0	11.2	3.6
r (p.p.)†	0.21	1.23	1.65	–0.60
r (net, p.p)	–0.36	–1.18	–1.53	0.28

Notes: Percentage changes are computed with respect to the benchmark US economy with $\tau_k = 25\%$. †Changes in the interest rate are reported in percentage points. The net wage and net interest rate are defined as $(1-\tau_l)\,w$ and $(1-\tau_k)\,r-\tau_a$, respectively. TFP is measured in the intermediate goods market. The optimal threshold amounts to 25% of the average earnings of the working population in the benchmark economy $(\overline{E})$.

taxes labor more heavily to subsidize capital income. This is a stark contrast between two tax systems that are equivalent without rate of return heterogeneity.

As for welfare, OWT delivers a larger gain (of about 9%) than OKIT. Furthermore, Section 6.4 shows that the welfare gains from OKIT are not robust to considering the transition path: cohorts that are alive at the time of the switch to OKIT would experience large welfare *losses*; this is not the case for OWT, which delivers robust welfare gains to those cohorts.[36] The reason for this asymmetry will become clear shortly.

6.1 Changes in Macro Variables

Table IX reports the percentage changes in aggregates relative to the US benchmark. (RN reform results are reproduced for completeness.) Comparing the two optimal tax

[36]Despite this reversal, we believe it is useful to present and discuss the OKIT stationary state results as a cautionary note and also because they illustrate key mechanisms for capital income taxation that arise with heterogeneous returns.

systems reveals some sharp contrasts. Broadly speaking, OWT results in relatively modest changes, with K barely rising, and Y and $\overline{w}$ rising by smaller amounts compared with the RN reform and OKIT (last column). However, by lowering the tax on labor, OWT boosts after-tax wages significantly (by 12.0%), which in turn incentivizes work, leading to higher labor supply (+3.3%). Notice that OWT delivers a higher TFP_Q gain of 7.7% (and 3% aggregate TFP) than the RN reform and OKIT (5.3% and 5.4%, respectively). Overall, OWT shifts the total tax burden from labor to capital and further shifts the capital tax burden from high-productivity entrepreneurs to low-productivity ones.

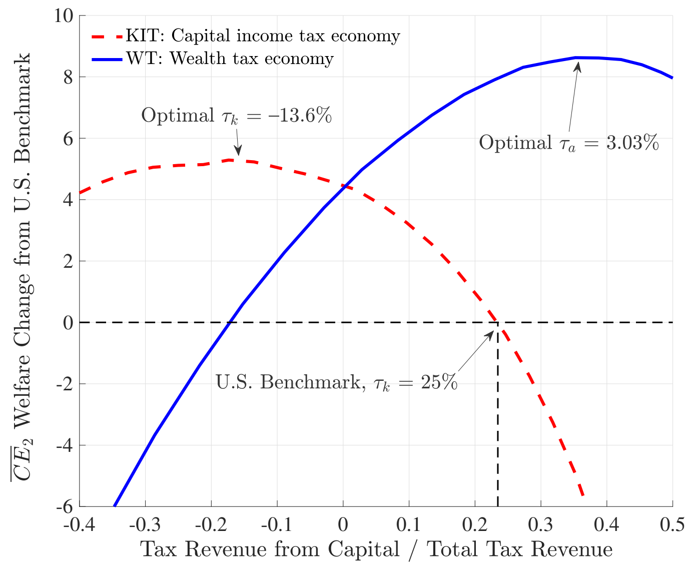
Next, the implications of OWT-X for aggregates are broadly similar to those of OWT, with a slightly smaller rise in output and consumption and a slightly larger rise in TFP. While the low exemption level mainly affects low-wealth individuals, the resulting higher τ_a^* distorts the savings incentive of the wealthy, leading to lower K and, in turn, a smaller rise in C and Y. The exemption might also be benefitting young entrepreneurs with high $\bar{z}_i$, thereby contributing to the higher TFP, although we have not attempted a formal quantification of this effect.

In contrast to OWT, OKIT (last column) causes rather dramatic changes in the economy. The subsidy policy boosts the income of high-productivity entrepreneurs and incentivizes them to save more, resulting in a nearly 40% higher level of K. However, the policy also requires more revenues from labor, leading to very small gains in after-tax wages (3.6%) relative to before-tax wages and output, both of which rise by more than 15%. In this sense, OKIT shifts the tax burden in the opposite direction to OWT—from the wealthy to wage earners—delivering efficiency gains at the expense of large distributional losses. Finally, the contrast between how OWT and OKIT affect the steady state K level has crucial implications for the transition analysis we conduct in Section 6.4. In particular, the transition after a switch to OKIT will involve substantial capital accumulation with significant welfare costs, unlike the transition after a switch to OWT.

6.2 Mechanisms at Play

To better understand the differences between the two tax systems, we plot how the welfare objective ($\overline{\mathsf{CE}}_2$ measure) varies with the share of tax revenue raised from capital in Figure 2. Welfare changes on the y-axis are relative to the US benchmark, which is normalized to zero. Because the total tax revenue is fixed ($= \mathsf{G} + \mathsf{SSP}$), as the revenue share from capital varies along the x-axis, the labor tax adjusts in the background to balance the government budget. The optimal tax rate is found where the objective value is maximized on this graph.

FIGURE 2 – Welfare Gains from Optimal Taxation

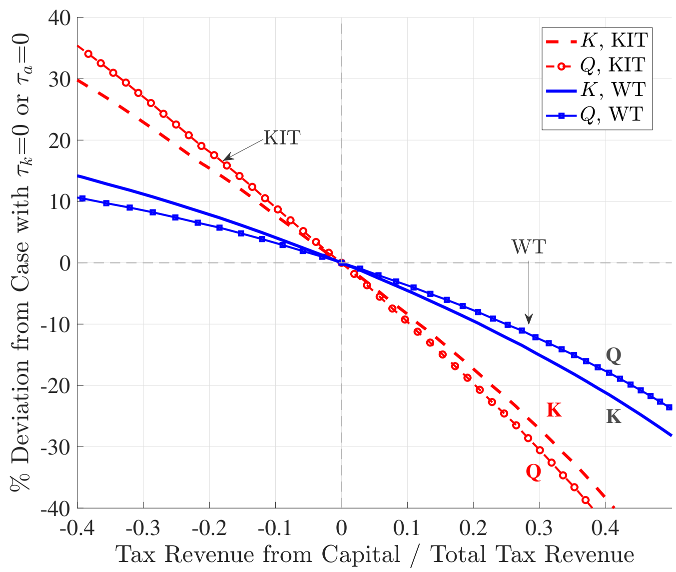


There are a few key takeaways. The first one is the obvious contrast between the slopes of the two lines: whereas welfare declines as more tax revenue is raised from capital under capital income taxation (except at the very low end), it increases under wealth taxation (solid blue line). Second, despite this apparent contrast, both patterns are driven by the same principle: with persistent return heterogeneity, taxing capital has a stronger distorting effect than it does without heterogeneity. For example, under capital income taxation, those who pay the highest taxes are those who are on average the most productive entrepreneurs, and those who are spared are the least productive ones. This asymmetry makes it optimal to flip the tax into a subsidy so as to boost productivity and output. Under wealth taxation, the same asymmetry stemming from return heterogeneity is dealt with by imposing a relatively high tax on wealth, which creates the same type of reallocation toward more productive entrepreneurs. That said, because wealth is still *taxed*, the effects on savings incentives are not as strong as with capital income *subsidies* and therefore do not cause a large rise in K, so the bulk of the gains come from reallocation.

The comparison between the distortions created by each tax system can be seen more clearly in Figure 3. There are two major differences. First, a higher wealth tax reduces K (solid blue line) more gradually than a higher capital income tax (dashed red line). In other words, the same amount of revenue can be raised with the former with a smaller distortion to K than with the latter.[37] Second, and compounding the first effect, Q declines

[37]A higher elasticity of aggregate capital with respect to a capital income tax than with respect to a

Figure 3 – How K and Q Vary with Revenue Raised from Taxing Capital



more gradually than K under wealth taxation, whereas the reverse happens under capital income taxation. Thus, wealth taxation improves the efficiency of capital allocation (Q versus K), whereas capital income taxation does the opposite. The relationship reverses for subsidies, so the capital income subsidy results in both an increase in K and a more efficient allocation.

To sum up, the optimal wealth tax is positive because the bulk of the population gains from the significant rise in average after-tax labor income, while the reduction in capital income affects a smaller group (of capital owners), and the reduction is not as large as it would have been under a comparable capital income tax. The capital income tax is a subsidy because both after-tax labor and capital income increase with the subsidy, although the gains in after-tax wages flatten out because of the rise in τ_ℓ as the subsidy becomes too large (τ_k too negative).[38]

The optimality of a capital income subsidy stands in sharp contrast with some well-understood results in the literature. For example, Conesa, Kitao, and Krueger (2009) find

wealth tax puts relatively more downward pressure on the optimal capital income tax, consistent with Saez and Stantcheva (2018), who show that a higher elasticity reduces the optimal capital tax.

[38]Notice that the objective function of the KIT economy in Figure 2 is relatively flat for $\tau_k < 0$ owing to the flatness of the after-tax labor income just described (and plotted in Figure C.4 in Appendix C). As a result, the magnitude of the subsidy can be sensitive to model details or parameterizations. For example, an earlier version of this paper (Guvenen et al., 2019) had a somewhat different calibration that yielded an optimal τ_k of –35%, while the results for welfare and the wealth tax were very similar to the present version.

an optimal $\tau_k = 36\%$ in a model that shares many features with ours, with the exception of return heterogeneity. Without this heterogeneity, the wealthy are *workers* who earned a high labor income in the past and saved part of it, but they are not any better at investing this wealth than others, so the efficiency losses from capital income taxation are significantly less distorting than shown here. As a result, the distributional benefits can outweigh the costs of distortion and make a positive tax rate optimal. We can obtain the same result here: although there are other differences between our model and theirs, shutting down return heterogeneity alone brings the optimal τ_k from –13.6% to 25.1%.

6.3 Who Gains and Who Loses?

Some of the broad patterns about the distribution of welfare changes parallel those we saw for the tax reform analysis above. For brevity, we mention those patterns briefly and instead focus on results that are unique to the optimal tax analysis. For OWT, the patterns are similar to those from the RN reform discussed earlier: the young gain more than the old; the gains increase with ability above the median for younger individuals, while the opposite happens at older ages—so older wealthy individuals experience welfare losses (Table B.3 in Appendix B.2). The gains and losses are larger relative to the tax reform, which is not too surprising given that in OWT, the optimal τ_a is more than twice as high, and τ_ℓ is not fixed and is lower.

Welfare gains fall with age under OKIT as well, but it contrasts with OWT in that gains rise with ability for all age groups (which is to be expected, given the large subsidy to capital income) and are positive (Table B.3, middle panel). An important difference between OWT and OKIT is that wealth taxes deliver more evenly distributed welfare gains across productivity groups (thanks to the large rise in after-tax wages) than OKIT.

Decomposing Welfare Changes: Levels versus Redistribution

We now decompose the average welfare gain to quantify the contribution of level versus redistributional changes stemming from each policy (reported in Table VIII).[39]

[39]The contribution of level changes is calculated by starting from the US benchmark and scaling up or down the consumption and leisure allocations of all newborns by the percentage change in corresponding aggregates. The average welfare gain corresponding to these hypothetical allocations (call it $\overline{\mathsf{CE}}_2^{\text{level}}$) gives the contribution of level changes, with the remaining portion of $\overline{\mathsf{CE}}_2$ attributable to distributional changes. More specifically, the two components can be written as $1+\overline{\mathsf{CE}}_2 \approxeq (1+\overline{\mathsf{CE}}_2^{\text{level}})(1+\overline{\mathsf{CE}}_2^{\text{redistr.}})$; see Appendix A.3. Decomposing $\overline{\mathsf{CE}}_1$ yields very similar results. We have also further decomposed each component into the part coming from consumption and labor and found that most of the gains are coming from consumption, unsurprising given that the changes in labor supply are small.

Table X – Decomposition of Welfare Gains

	OWT	OWT-X	WTE-X	OWT: L-INEQ	OKIT
Total ($\overline{\mathsf{CE}}_2$)	8.7	8.8	8.6	5.2	5.1
Level	5.9	4.3	5.1	4.4	14.7
Distribution	2.6	4.3	3.3	0.7	–8.3

Notes: The table reports the decomposition of the average welfare gains of newborns as measured by $\overline{\mathsf{CE}}_2$.The optimal exemption threshold, $\underline{a}^*_{\text{ex}}$, is equal to $0.3 \times \overline{y}$. The high exemption threshold is equal to $\overline{y}$. See the text for details.

Table X reports the decomposition results. With OWT, about 2/3 of the average welfare gain is due to a positive level effect (5.9% out of 8.7%), and the remaining 1/3 is due to a positive redistributional effect. Therefore, an optimally designed wealth tax system can improve welfare by both growing the economy *and* improving equity. It also contrasts with the equity-efficiency trade-off present with a capital income tax ($\tau_k > 0$), which implies a level loss due to the savings distortions of a positive tax, which is weighed against distributional gains stemming from lower consumption inequality or idiosyncratic uncertainty.

Adding an exemption level (OWT-X, second column) has a marginal impact on the *average* welfare gain relative to OWT, but now the share accounted for by redistribution rises from one-third to one-half. The level effect is smaller because the higher τ_a limits the rise in K and therefore in C, whereas the redistributional gain is larger because both the exemption level and the lower τ_ℓ benefit individuals at the lower end more than others. The exemption is also a boon to retirees, who lose on average under flat-rate wealth taxation in OWT (for the same reason explained in the RN reform) but gain under OWT-X. For example, whereas only about 1% of retirees in the bottom 90% of the $\bar{z}$ distribution gain from OWT, this figure rises to about 75% in OWT-X (Table B.3 in Appendix B.2). Furthermore, OWT-X benefits the highest-ability newborns (top 1%) more than OWT does (Table B.3) by relieving them of the tax burden at younger ages, while the opposite happens at older ages. In other words, the exemption level strengthens the redistribution from old to young among the most productive entrepreneurs.

Given that the optimal exemption level is fairly low, we also examine the effects of higher values. We find that values up to 100% of $\overline{y}$ deliver only slightly lower average welfare gains than the optimum but exempts about 60% of the population from the wealth tax, delivering redistributional gains between OWT and OWT-X (WTE-X in Table X).

However, further increasing $\underline{a}^*_{\text{ex}}$ (not reported) leads to rapid declines in average as well as redistributional gains, suggesting that thresholds in the millions of dollars, which are often proposed in policy debates, may not improve overall welfare or achieve desirable redistributional objectives.

Turning to capital income taxation, the usual equity-efficiency trade-off just described operates in this case too but is manifested in the opposite direction: because the optimal policy is a subsidy, there are large level gains (14.7%) combined with large distributional losses (–8.3%), adding up to a smaller gain of 5.1% (Table X, OKIT column).

Before concluding this section, two remarks are in order. First, does the optimal wealth tax always deliver distributional gains? Clearly, there is no reason for this to always be true.[40] There are two obvious and plausible cases that deserve mention. The first one is the amount of inequality to begin with. This can be seen in the low-inequality, L-INEQ, calibration in Table X, where a smaller share of welfare gains comes from redistribution (0.7% out of 5.2%), with the bulk coming from level effects, which is not surprising given the much lower top-end inequality in this calibration. This result also underscores the importance of a parameterized model to match the top-end inequality in the data for a sound quantitative analysis of capital taxation.

The second case is when the society does not have strong preferences for redistribution—or aversion to inequality—which corresponds to a lower risk aversion in our setup. In the baseline, we set $\sigma = 4$ for comparability with Conesa, Kitao, and Krueger (2009), who also studied the equity-efficiency trade-off in a related setting. We solve the model for $\sigma = 1$ and repeat the OWT experiment (not reported in the table). This raises the level gains from 5.9% to 9.7% while also turning the distributional gains of 2.6% to a loss of 5.3%, for an overall welfare gain of 4%. The effect is smaller for OKIT, with a negligible decline in level gains and a moderate increase in distributional losses (from 8.3% to 10%) for a total gain that falls from 5.1% to 3.1%

In summary, these results collectively show that a wealth tax is a powerful and flexible policy instrument that can achieve the desired balance between equity and efficiency embedded in a society's preferences. What our baseline analysis has shown is that in a more egalitarian society ($\sigma = 4$), the resulting outcome can be an improvement in both efficiency and equity relative to the current US benchmark.

[40] It is possible, however, to establish that a wealth tax yields efficiency gains under very general conditions, which basically require that entrepreneurial productivity is positively autocorrelated. We prove this result formally in Guvenen, Kambourov, Kuruscu, and Ocampo (2022a).

Does Higher Wealth Inequality Necessarily Imply Distributional Losses?

Wealth inequality increases under both tax systems, with the top 1% share rising from 36% to 45% in OWT and OWT-X and to 47% in OKIT. While these magnitudes are similar, the distributional consequences of the two tax systems are in sharp contrast: wealth taxation delivers solid welfare gains from redistribution (e.g., by 4.3% for OWT-X), while OKIT results in an 8.3% welfare loss from a worsening distribution (Table X). An important conclusion we draw from these results is that a rise in wealth inequality may not necessarily indicate worsening inequality in welfare. What matters more is whether the wealth is held by productive individuals.

6.4 Equilibrium with Transition

We now extend the analysis by modeling the transition path to the new steady state after the switch to a new optimal policy regime. The goal here is not to solve for the optimal path of taxes with transition but rather to solve for an equilibrium that holds throughout the transition and in the new stationary state while minimally deviating from the OKIT and OWT tax rates found above. To this end, we fix one of the two taxes (e.g., τ_a) at its non-transition optimum (from Table VIII), allow the government to run a non-balanced budget during the transition, and choose the other tax (τ_ℓ) such that the budget—which now includes interest payments on the accumulated debt—is balanced in the new stationary equilibrium. Therefore, this tax rate needs to be solved jointly with the equilibrium transition path, defining a new fixed point problem.[41]

Starting with OWT, we fix τ_a at 3.03% (from Table VIII) and choose τ_ℓ as just described, which yields $\tau_\ell = 17.01\%$ (compared with 15.42% without the transition). We first compare the newborns—those who enter the economy in the first year of the reform—with the same cohort without the policy reform. As seen in Table XI, newborns experience an average welfare gain of 6.0%, which is more than two-thirds of the gain we found in the stationary state comparison (8.7%). Broadening the comparison to the entire population alive at the time of the reform shows modest differences, with welfare rising by 3.5% *with* transition versus 4.3% *without* transition (the latter figure was not reported earlier).

These welfare results are driven by two main factors. First, the immediate reduction in τ_ℓ causes after-tax wage income to jump, raising average consumption by 4.5% in the

[41]In Guvenen et al. (2019), we also conducted the reverse exercise—by fixing τ_ℓ and choosing the capital tax rates to balance the budget—and found substantively very similar results, so we did not repeat that experiment here.

Table XI – Extension: Policy Analysis Accounting for the Transition Path

	OWT	OKIT
τ_k	0.00	–13.60*
τ_a	3.03*	0.00
τ_ℓ	17.01	35.90
$\overline{\mathsf{CE}}_2$ (newborn)	6.0 (8.7†)	–8.4 (5.1)
$\overline{\mathsf{CE}}_2$ (all)	3.5 (4.3)	–6.1 (4.5)

Notes: Both the tax rates and welfare figures are reported in percentages. †The numbers in parentheses report the welfare gains from the comparison across stationary equilibria above. *In each experiment, the tax on wealth or capital income is kept at its non-transition optimum shown in Table VIII, while the labor income tax is adjusted to obtain an equilibrium with transition.

first year of reform. The lower τ_ℓ also raises labor supply, raising output by 1% in the first year. Moreover, the rise in consumption is larger at the lower end of the income distribution—because wage income accounts for a larger share of total disposable resources—which leads to a more even consumption distribution.

The second factor is that the OWT transition requires (almost) no capital accumulation at the aggregate level because the levels of K in pre- and post-reform steady states are virtually the same, differing from each other by a few percent (Table IX). Thus, there is no aggregate capital accumulation that reduces average consumption during the transition. Rather, the gains are driven by the reallocation of capital toward more productive individuals through the use-it-or-lose effect and the differential behavioral savings response across the population (with high-productivity entrepreneurs increasing their savings rate and low-productivity entrepreneurs doing the opposite), which reinforce this reallocation.[42]

Under OKIT, the average welfare of newborns *falls* by 8.4%—in contrast to the 4.3% *gain* we found above without transition. Looking at the entire population alive at the time of the reform shows an average loss of 6.1%, overturning the 4.5% gain without transition (not reported earlier).

This stark reversal stems from two sources and can be anticipated from the discussion of the OWT case. The first one is straightforward: the transition analysis makes explicit

[42]That said, K is not constant during the transition but follows a nonmonotonic path, falling for the first 10 years or so and then taking another 30 years to rise back to its pre-reform level. This is because switching to wealth taxes reduces the after-tax return for many older and wealthier individuals, who now find it optimal to spend down their wealth. Although the opposite happens for young and productive individuals, their wealth is a smaller fraction of the aggregate. However, Q rises monotonically—thanks to reallocation—so output rises throughout the transition.

TABLE XII – Robustness: Optimal Wealth Tax

	Baseline	Credit Spread		Public	Corporate	Pure rents	Non-linear OKIT	
	OWT	10.1%	6%	Firms	Sector	model	$\tilde{\tau}_k(y) = y - \psi y^{\eta}$	
τ_a	3.03	2.33	2.46	2.76	3.25	1.40	—	—
τ_ℓ	15.4	13.6	15.5	17.6	16.3	27.0	22.4 (fixed)	32.3
τ_k	—	—	—	—	—	—	$\underset{(\psi,\eta)}{(0.73, 1.022)}$	$\underset{(\psi,\eta)}{(1.20, 0.992)}$
					Change in Welfare (%)			
$\overline{CE}_1$	9.0	6.1	4.3	5.9	5.8	–1.7	0.9	4.2
$\overline{CE}_2$	8.7	5.6	3.5	4.8	5.5	–1.4	0.8	5.4

Notes: The seven robustness experiments are as follows: (1) replacing collateral constraints with unlimited borrowing subject to a credit spread of 10%, generating a debt-to-GDP ratio of 1.5; (2) same as (1) but with a spread of 6%; (3) allowing firms to stochastically transition to relaxed collateral constraints; (4) introducing a corporate sector with Cobb-Douglas production and no borrowing limits; (5) eliminating z_i heterogeneity to focus on pure monopolistic rents; (6) a tax reform that replaces τ_k with a nonlinear capital income tax; and (7) an optimal nonlinear capital income tax experiment (choosing ψ, η, τ_ℓ).

the cost of accumulating the large capital stock of the post-reform steady state (about 48% higher), which requires higher savings and lower consumption early in the transition. Second, τ_ℓ jumps from 22.4% in the benchmark to 35.9% in the first year of transition, lowering after-tax labor income both directly and indirectly (by depressing labor supply, which falls by 5% in the first year of the new policy, driving a 2.6% fall in output upon impact). Both of these costs are borne over the first several decades of the transition, whereas the benefits (higher wages and consumption) are realized only gradually and are thus discounted, adding up to a large welfare loss.

Overall, we find that incorporating the transition path has a modest effect on the implications of optimal wealth taxes found in the baseline analysis, whereas it upends the welfare gains from large capital subsidies that was found in the steady state comparison. In this sense, this analysis strengthens the case for wealth taxes and significantly undermines the case for capital income subsidies.

7 Robustness and Extensions

So far, we have discussed several robustness checks, including a low-inequality calibration (L-INEQ), endogenous entrepreneurial labor supply (see also Appendix F), lower risk aversion, and introducing progressivity through an exemption level, among others. We now present several additional extensions and robustness analyses.

Alternative Modeling of Financial Frictions

We consider several alternative forms and calibrations of financial constraints, three of which we report in Table XII.

Credit Spread. First, we removed the collateral constraint and allowed unlimited borrowing subject to a credit spread between borrowing and saving rates (column 2). We calibrated the spread to match the baseline debt-to-GDP ratio of 1.5, which gives a spread of 10.1%. The optimal wealth tax is 2.33%, delivering an average welfare gain of 5.6%. Although this credit spread is not unreasonable for small and young firms, it is on the high side for larger firms. As an alternative (column 3), we set the spread to 6%, which is in line with empirical estimates.[43] The recalibrated model yields a debt/GDP ratio (1.97) higher than that in the data but still delivers welfare gains ranging from 3.5% to 4.3%.

Firms Going "Public" Stochastically. We consider a second extension in which firms stochastically become "public," by which we mean they gain access to a substantially higher credit limit: $\overline{\vartheta}_{\text{public}} \gg \overline{\vartheta}(z)$. The public/private status of a firm is inherited across generations, and public firms exit ($\mathbb{I}_{ih} = 0$) at a lower rate than private firms. The arrival rate of the "Calvo fairy" is calibrated to generate a target ratio of public to private firms of 0.5%.[44] We set $\overline{\vartheta}_{\text{public}} = 10$, which corresponds to a leverage ratio of 90%—at the very top end of values seen in the data (Asker, Farre-Mensa, and Ljungqvist, 2011).[45] Further details of the model and calibration are in Appendix E.4. The calibration generates a debt/GDP ratio of 2.4, showing that the new structure indeed delivers much more borrowing than the baseline. The optimal wealth tax is slightly lower (2.76%) and generates lower but still robust welfare gains ranging from 4.8% to 5.9%.

Other Extensions Involving Financial Constraints. Because firms become public at a gradual pace, we could ask if the results would be different if we relaxed the constraints starting at age zero. As one way to capture this, we recalibrated the baseline collateral constraint to match a debt-to-GDP ratio of 2.5, which implies a rise in ϑ from 2.8 in the baseline to 11.5 for the top ability group. The optimal wealth tax is 2.3%, and the welfare gain is 4.2% (Table E.10 in Appendix E.5). Finally, we solved the model with a flat collateral constraint (i.e., one that does not depend on ability: $\vartheta(z) = \overline{\vartheta}$) and recalibrated

[43] For example, in a sample of high income countries, Kochen (2022) calculates the mean spread for 10-year-old firms to be 6.6%.

[44] The target is calculated as the ratio of the number of public firms in Compustat to firms with 5+ employees from the Business Dynamics Statistics of the US Census Bureau.

[45] We also solved the $\overline{\vartheta}_{\text{public}} = \infty$ case, which, perhaps somewhat surprisingly, delivers a comparable rise in welfare but this value is harder to defend empirically.

as before (Table E.10). The welfare gain from OWT is 11.2%, higher than the baseline, confirming our assertion in Section 3.3 that the baseline specification is a more conservative assumption. Finally, we have also considered a different collateral constraint of the form $k_i \leqslant a_i + \vartheta \cdot (z_i a_i)^{\mu}$, which has the same form that emerges from imperfect enforceability of collateral constraints—see the discussion in footnote 12. OWT yields welfare gains ranging from 4% to 4.4%. Results are available upon request.

We now turn to other robustness exercises.

Adding a Corporate Sector. We solve a version of the model with an aggregate corporate sector that produces output according to a Cobb-Douglas production function of capital and labor ($Y_c = AK_c^{\alpha}L_c^{1-\alpha}$) and faces no financial constraints. The final good is a Cobb-Douglas aggregate of the output of the two sectors: $Y = Y_c^{\rho}Y_p^{1-\rho}$, with subscripts denoting "corporate" and "private" sectors, respectively, and $L_c + L_p = L$. There is a common capital market for corporate firms and entrepreneurs where they can borrow at rate r. See Appendix E.3 for further details. We set ρ and A to match the corporate sector's share of aggregate sales (50%) and aggregate capital stock (60%). We choose these empirical moments to be on the on the high end of the estimates of the size of the corporate sector so as to provide an upper bound on the potential impact of adding a corporate sector on our results.[46] The optimal wealth tax is 3.25% and delivers a welfare gain of 5.5% for newborn individuals.[47] Because private firms are a smaller fraction of the economy, misallocation is much lower than in the baseline (0.065 here versus 0.16), yet the welfare gain is almost 2/3 of the baseline figure. Most of the gains are due to reallocation *within* private firms, with only a small contribution from a small shift in aggregate capital from the corporate sector to private firms (a 2% shift in the capital share).[48]

Pure Rents Model. A flat-rate wealth tax effectively taxes normal returns at a higher rate than supernormal or excess returns. While this feature is the key driver of the use-it-or-lose-it mechanism, and consequently of the efficiency gains, it goes against the well-understood result in public finance that rents should be taxed at a higher rate. In our model, every entrepreneur is a monopolist for the variety she produces and thus earns

[46] We calibrate the model to the same targets as before, with the exception of $\vartheta(z)$, which we keep fixed. This implies a private debt to asset ratio of 0.75, which is higher than 0.45 reported by Asker, Farre-Mensa, and Ljungqvist (2011).

[47] Asker, Farre-Mensa, and Ljungqvist (2011) report 41% and 47%, respectively for the two moments when the government and agricultural sectors are excluded. A calibration that matches these moments delivers a welfare gain of 8.8%.

[48] We also considered a version of the model in which the corporate and private goods were perfect substitutes. The same mechanisms are also present in this case, and the efficiency and welfare gains are larger, with the reallocation of capital across sectors playing a larger role.

monopoly rents. So, the feature that makes taxing excess returns desirable is already built into our model, and the optimality of the wealth tax happens despite this force.

To better explain this point, we note that return heterogeneity results here from two features: heterogeneity in z_i and heterogeneity in monopoly profits due to borrowing frictions. Thus, if we eliminate heterogeneity in z_i, all firms will continue to earn monopoly rents (albeit with much less return heterogeneity). If we conduct the optimal wealth tax exercise in this setting, what do we find? Column 5 in Table XII shows that the optimal tax rate is $\tau_a = 1.4\%$ (and $\tau_\ell = 27\%$), but it results in a welfare loss of 1.4% to 1.7% relative to the US benchmark—which features capital income taxes.[49] In other words, without heterogeneity in z_i, we can replicate the result that it is better to tax excess returns with a capital income tax.[50] The optimality of a wealth tax and the large welfare gains relative to capital income taxes arise from inherent differences in entrepreneurial ability, z_i, which are powerful enough to overcome the desire to tax rents built into the model.

Nonlinear Capital Income Taxes. The use-it-or-lose-it mechanism reallocates wealth to high-productivity individuals by taxing lower returns more heavily than higher returns, which raises the question of whether a similar effect can be achieved through nonlinear capital income taxes. To investigate this, we generalize τ_k to a log-linear form: $\tilde{y} = y - \tilde{\tau}_k(y) = \psi y^\eta$, where y and $\tilde{y}$ are pre-tax and after-tax income, respectively, η determines the degree of progressivity, and $\psi > 0$ pins down the average tax rate for a given η.[51]

We repeat the tax reform and the optimal tax experiments with this new structure. In the tax reform, we replace $\tau_k = 25\%$ with $\tilde{\tau}_k(y)$ and keep τ_ℓ and revenue fixed and optimize over (ψ, η). The optimal tax schedule is regressive ($\eta = 1.022$) with the average tax rate, $\tilde{\tau}_k(y)/y$, falling from 44.8% for the bottom 10% of the capital income distribution to 24.7% for the top 1% . The welfare gain is much smaller ($\overline{\mathsf{CE}}_2 = 0.8\%$) than the wealth tax reform above (7.2%) (Table XII, column 6). In the second experiment (column 7), we optimize over 3 parameters (ψ, η, τ_ℓ). The optimal tax is a subsidy as before, but now it is progressive, ranging from 33% to 19% from the bottom percentile group to the top. The remaining results are substantively very similar to the OKIT case, with a welfare gain that is only marginally higher (5.2% versus 5.1% in OKIT). Overall, while nonlinear capital income taxes improve over OKIT, the changes are modest.

[49] Changes in macro quantities corresponding to experiments in Table XII are reported in Table E.11.

[50] The optimal wealth tax is positive because it is less distorting than the labor income tax, which has already been raised relative to the US benchmark. So, it is optimal only in the absence of capital income tax as a feasible instrument.

[51] This specification has a long history in public finance and has been used more recently by Benabou (2000) and Heathcote, Storesletten, and Violante (2014).

Adding a Wealth Tax on Top of Existing Capital Income Taxes. The reader will notice that the tax reform and optimal tax analysis in this paper bear no resemblance to the policy proposals recently circulated by policy makers in the United States and widely debated by the public. The crucial difference is that we consider the wealth tax as an *alternative* to the capital income tax, and a significant part of the welfare gains arises from eliminating the highly distorting capital income tax. In addition, the wealth tax we consider is not levied on multimillionaires or billionaires but is much more broad based—even when there is an exemption threshold.

That said, given that we have a model capable of generating billionaires and matching the right tail of the wealth distribution, it is useful to shed light on the potential implications of proposed policies. So, we consider two experiments in which a wealth tax is imposed *on top* of the existing US tax system (therefore keeping the capital income tax at 25%). In the first case, a 2% wealth tax is levied without changing existing taxes. It is not clear how the proceeds from the proposed new tax on wealth are going to be used, so to show a useful range of possibilities, we solve for two special cases. In the first, tax revenue is not used for any purposes that yield utility. In the second experiment, the government puts the new revenues to one of its best uses: it lowers τ_ℓ to generate the same revenue as in the baseline. In the former case, the outcomes range from a 10% welfare loss to no change in welfare (see columns vi–ix of Table E.10 in Appendix E.5). Overall, we conclude that the consequences of *adding* a wealth tax can range from no substantive effect to potentially large welfare losses.

Other Robustness and Extensions. We conducted other robustness exercises, which we briefly summarize here (and report in Table E.10 in Appendix E.5). We (i) allow for higher markups ($\mu = 0.8$), (ii) eliminate the stochastic fluctuations in productivity ($z_{ih} = \bar{z}_i$ for all h), and (iii) assume everybody starts life in the middle lane ($z_{i0} = \bar{z}_i$ for all i) but modify the transition matrix to allow moves up into fast lane. In these three cases, the optimal τ_a varies between 2.16% (case ii) and 2.8% (iii), and the welfare gains range from 5.5% (ii) to 8.2% (iii). The other main substantive conclusions remain intact. While (i) and (iii) can match the targets nearly as well as the benchmark model, (ii) cannot, which is why we introduced stochastic fluctuations in z_{ih} in the first place.

In addition, we address the effect of introducing managerial effort into our framework with an extension in Appendix F (by modifying eq. (8) to include the entrepreneur's labor supply) and discuss the conditions under which it amplifies or dampens the effects of taxes relative to the baseline model.

A final takeaway from these robustness experiments concerns misallocation. Our substantive conclusions about the effectiveness of wealth taxes are broadly robust to the *level* of misallocation of the benchmark economy in the experiments above. For example, in the baseline calibration, the economy is 16% below the efficient TFP frontier (the Hsieh-Klenow measure of misallocation), whereas this number is 6.5% when the corporate sector is added and 5% when the debt-to-GDP ratio is calibrated to 2.5. While the average welfare gain is somewhat lower in the last case (4.2%), it is still robustly positive. Instead, welfare gains are more closely related to the degree to which the wealth tax can reduce misallocation, with a correlation of 0.92 between the welfare gain and the change in misallocation (between benchmark and OWT economies).

8 Discussion and Conclusions

In this paper, we studied the efficiency and distributional implications of wealth taxation and compared them to those of capital income taxation. Under the latter, entrepreneurs who are more productive, and therefore generate more income, pay higher taxes. Under wealth taxation, on the other hand, entrepreneurs who have similar wealth levels pay similar taxes regardless of their productivity, thereby expanding the tax base and shifting the tax burden toward unproductive entrepreneurs. Furthermore, wealth taxes reduce the after-tax returns of high-productivity entrepreneurs less than those of low-productivity ones, which creates a behavioral savings response, which further shifts resources toward the productive ones. Overall, our analysis lends support to the consideration of wealth taxation as a more desirable alternative to capital income taxation, as it has the potential to improve aggregate productivity, grow the economy, generate redistributional gains, and improve welfare.

The wealth tax we propose in this paper differs in crucial ways from those that have been implemented by governments in the past as well as from those that are currently being debated by the public. Specifically, the wealth tax is levied after existing capital income taxes *are repealed*; it is levied on the *book value* of assets rather than the market value; and it is broad-based rather than targeted at the top of the wealth distribution. As a result, its implications are also different in important ways.

For example, one conclusion from our analysis is that capital income taxation is significantly more distorting when returns are heterogeneous than when returns are assumed to be homogeneous, which has been the default assumption in the literature. Repealing the capital income tax eliminates these distortions, and replacing with a wealth tax raises the same (or more) revenues with much less distortion. Furthermore, because it is not an additional tax, there is no presumption that the overall capital tax burden will go up. In fact, in our tax reform experiment, the tax revenue collected from capital went from about 6% of GDP in the US benchmark with capital income taxation to 4% of GDP when the economy moved to the wealth tax. At the same time, if the society's preference for redistribution is strong, then the wealth tax becomes an effective tool for raising higher revenues than the capital income tax with much less distortion (as in our baseline OWT experiment, given $\sigma = 4$).

Imposing the wealth tax on the book value is important for the efficiency gains from the use-it-or-lose-it mechanism. Because the market value incorporates the entrepreneur's productivity, basing the tax on the market value would raise the tax burden of the highly productive entrepreneurs, weakening the positive reallocation and efficiency gains. Incidentally, levying the wealth tax on the book value obviates the need to assess the market value of private firms, which has been one of the most important practical challenges governments have faced when implementing wealth taxes in the past.

Because of the broad-based nature of the wealth tax we study, it raises substantially more revenues than the wealth tax policies implemented by many governments that have been targeted at the very top. For example, with the exception of Luxembourg, Norway, and Switzerland, all OECD countries that have implemented a wealth tax designed them to be sufficiently narrow that they raised minimal revenues from it (0.1% to 0.2% of GDP since 1980, OECD, 2018).[52] These low revenues, coupled with the high cost of enforcement (e.g., the cost of assessing market values) was an important factor in the declining popularity of wealth taxes (Kopczuk, 2013; OECD, 2018). The broad-based wealth tax we study can alleviate this problem. Moreover, eliminating the capital income tax would free up the resources of tax agencies that can then be directed towards enforcing the wealth tax.

Having said that, there are additional practical issues that are likely to arise when a wealth tax is implemented, which we do not address in this paper. As mentioned in the introduction, these include concerns about capital flight as wealthy households choose to relocate their assets to jurisdictions without a wealth tax, the tax treatment of unrealized capital gains, the extent to which private business owners may choose to reclassify their labor income as capital income to benefit from the reduced taxation on capital income, among others. We hope that the results of our paper provide an impetus for exploring these and other issues that we have not addressed in this paper. These questions are on our current and future research agenda.

[52]The revenues remained less than 0.5% of GDP in Norway and below 1% in Switzerland.

References

Aiyagari, S. Rao. 1995. "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting." *Journal of Political Economy* 103 (6):1158–1175. URL https://doi.org/10.1086/601445.

Allais, Maurice. 1977. *L'impôt Sur Le Capital et la Réforme Monétaire*. Hermann.

Asker, John, Joan Farre-Mensa, and Alexander Ljungqvist. 2011. "Comparing the Investment Behavior of Public and Private Firms." Working Paper 17394, National Bureau of Economic Research. URL http://www.nber.org/papers/w17394.

Auerbach, Alan J, Jagadeesh Gokhale, Laurence J Kotlikoff, John Sabelhaus, and David N Weil. 1995. "The Annuitization of Americans' Resources: A Cohort Analysis." Working Paper 5089, National Bureau of Economic Research. URL http://www.nber.org/papers/w5089.

Bach, Laurent, Laurent E. Calvet, and Paolo Sodini. 2020. "Rich Pickings? Risk, Return, and Skill in Household Wealth." *American Economic Review* 110 (9):2703–2747. URL https://www.aeaweb.org/articles?id=10.1257/aer.20170666.

Bell, Felicite C. and Michael L. Miller. 2002. "Life Tables for the United States Social Security Area: 1900-2100"." Actuarial Study 116, Office of the Actuary, Social Security Administration. URL https://www.ssa.gov/oact/NOTES/as120/LifeTables_Body.html.

Benabou, Roland. 2000. "Unequal Societies: Income Distribution and the Social Contract." *American Economic Review* 90 (1):96–129. URL https://www.aeaweb.org/articles?id=10.1257/aer.90.1.96.

Benhabib, Jess and Alberto Bisin. 2018. "Skewed Wealth Distributions: Theory and Empirics." *Journal of Economic Literature* 56 (4):1261–1291. URL https://www.aeaweb.org/articles?id=10.1257/jel.20161390.

Benhabib, Jess, Alberto Bisin, and Mi Luo. 2017. "Earnings Inequality and Other Determinants of Wealth Inequality." *American Economic Review* 107 (5):593–97. URL https://www.aeaweb.org/articles?id=10.1257/aer.p20171005.

Benhabib, Jess, Alberto Bisin, and Shenghao Zhu. 2011. "The Distribution of Wealth and Fiscal Policy in Economies With Finitely Lived Agents." *Econometrica* 79 (1):123–157. URL https://doi.org/10.3982/ECTA8416.

Bils, Mark, Peter J. Klenow, and Ciane Ruane. 2017. "Misallocation and Mismeasurement?" Working paper, Stanford University.

Buera, F., J. Kaboski, and Yongseok Shin. 2011. "Finance and Development: A Tale of Two Sectors." *American Economic Review* :1964–2002URL https://www.aeaweb.org/articles?id=10.1257/aer.101.5.1964.

Cagetti, Marco and Mariacristina De Nardi. 2006. "Entrepreneurship, Frictions, and Wealth." *Journal of Political Economy* 114 (5):835–870. URL https://doi.org/10.1086/508032.

Castañeda, Ana, Javier Díaz-Giménez, and José-Víctor Ríos-Rull. 2003. "Accounting for the

U.S. Earnings and Wealth Inequality." *Journal of Political Economy* 111 (4):818–857. URL https://doi.org/10.1086/375382.

Chari, V. V. and Patrick J. Kehoe. 1999. "Optimal Fiscal and Monetary Policy." In *Handbook of Macroeconomics*, vol. 1, edited by John B. Taylor and Michael Woodford, chap. 26. Elsevier, 1671–1745. URL https://www.sciencedirect.com/science/article/pii/S1574004899100399.

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger. 2009. "Taxing Capital? Not a Bad Idea After All!" *American Economic Review* 99 (1):25–48. URL https://www.aeaweb.org/articles?id=10.1257/aer.99.1.25.

De Nardi, Mariacristina and Fang Yang. 2016. "Wealth inequality, family background, and estate taxation." *Journal of Monetary Economics* 77:130 – 145. URL https://doi.org/10.1016/j.jmoneco.2015.10.005.

Erosa, Andres and Martin Gervais. 2002. "Optimal Taxation in Life-Cycle Economies." *Journal of Economic Theory* 105 (2):338–369. URL https://doi.org/10.1006/jeth.2001.2877.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri. 2020. "Heterogeneity and Persistence in Returns to Wealth." *Econometrica* 88 (1):115–170. URL https://doi.org/10.3982/ECTA14835.

Federal Reserve Statistical Release. 2015Q3. "Z.1 Financial Accounts of the United States: Flow of funds, balance sheets and integrated macroeconomic accounts." Tech. rep., Board of Governors of the Federal Reserve System. URL https://www.federalreserve.gov/releases/z1/20151210/z1.pdf.

Gabaix, Xavier. 2009. "Power Laws in Economics and Finance." *Annual Review of Economics* 1:255–93. URL https://doi.org/10.1146/annurev.economics.050708.142940.

Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll. 2016. "The Dynamics of Inequality." *Econometrica* 84 (6):2071–2111. URL https://doi.org/10.3982/ECTA13569.

Garriga, Carlos. 2003. "Optimal Fiscal Policy in Overlapping Generations Models." *mimeo* .

Gentry, William M and R. Glenn Hubbard. 2000. "Entrepreneurship and Household Saving." Working Paper 7894, National Bureau of Economic Research. URL http://www.nber.org/papers/w7894.

Golosov, Mikhail, Aleh Tsyvinski, and Ivan Werning. 2006. "New Dynamic Public Finance: A User's Guide." In *NBER Macroeconomic Annual 2006*. MIT Press. URL http://www.nber.org/chapters/c11181.

Guvenen, Fatih, Gueorgui Kambourov, Burhanettin Kuruscu, and Sergio Ocampo. 2022a. "Taxing Wealth and Capital Income when Returns are Heterogeneous." Working paper. URL https://ocamp020.github.io/GKKO-THEORY.pdf.

Guvenen, Fatih, Gueorgui Kambourov, Burhanettin Kuruscu, Sergio Ocampo-Diaz, and Daphne Chen. 2019. "Use It or Lose It: Efficiency Gains from Wealth Taxation." Working Paper 26284, National Bureau of Economic Research. URL http://www.nber.org/papers/w26284.

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song. 2021. "What Do Data on Millions of U.S. Workers Say About Labor Income Risk?" *Econometrica* 89 (5):2303–2339. URL https://doi.org/10.3982/ECTA14603.

Guvenen, Fatih, Jr. Mataloni, Raymond J, Dylan G Rassier, and Kim J Ruhl. 2022b. "Offshore Profit Shifting and Aggregate Measurement: Balance of Payments, Foreign Investment, Productivity, and the Labor Share." *American Economic Review* 112 (6):1848–1884. URL https://www.aeaweb.org/articles?id=10.1257/aer.20190285.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante. 2014. "Consumption and Labor Supply with Partial Insurance: An Analytical Framework." *American Economic Review* 104 (7):2075–2126. URL https://www.aeaweb.org/articles?id=10.1257/aer.104.7.2075.

Hendricks, Lutz. 2001. "Bequests and Retirement Wealth in the United States." Working paper, Arizona State University. URL https://lhendricks.org/Research/wps/bequdata_paper.pdf.

Hopenhayn, Hugo A. 2014. "Firms, Misallocation, and Aggregate Productivity: A Review." *Annual Review of Economics* 6 (1):735–770. URL https://doi.org/10.1146/annurev-economics-082912-110223.

Hsieh, Chang-Tai and Peter J. Klenow. 2009. "Misallocation and Maunfacturing TFP in China and India." *The Quarterly Journal of Economics* 124 (4):1403–1448. URL https://doi.org/10.1162/qjec.2009.124.4.1403.

Hubbard, R. Glenn, Kenneth L. Judd, Robert E. Hall, and Lawrence Summers. 1986. "Liquidity Constraints, Fiscal Policy, and Consumption." *Brookings Papers on Economic Activity* 1986 (1):1–59. URL https://doi.org/10.2307/2534413.

Imrohoroglu, Selahattin. 1998. "A Quantitative Analysis of Capital Income Taxation." *International Economic Review* 39:307–328. URL https://doi.org/10.2307/2527295.

Itskhoki, Oleg and Benjamin Moll. 2019. "Optimal Development Policies With Financial Frictions." *Econometrica* 87 (1):139–173. URL https://doi.org/10.3982/ECTA13761.

Jones, Charles. 2022. "Taxing Top Incomes in a World of Ideas." *Journal of Political Economy* 130 (9):2227–2274.

Jones, Charles I. and Jihee Kim. 2018. "A Schumpeterian Model of Top Income Inequality." *Journal of Political Economy* 126 (5):1785–1826. URL https://doi.org/10.1086/699190.

Judd, Kenneth L. 1985. "Redistributive Taxation in a Simple Perfect Foresight Model." *Journal of Public Economics* 28 (1):59–83. URL https://doi.org/10.1016/0047-2727(85)90020-9.

Kelley, Donna, Slavica Singer, and Mike Herrington. 2011. "Global Entrepreneurship Monitor, 2011 Global Report." Report, Global Entrrepeneurship Research Association. URL https://gemconsortium.org/report/gem-2011-global-report.

Kim, Tae-Hwan and Halbert White. 2004. "On more robust estimation of skewness and kurtosis." *Finance Research Letters* 1 (1):56–73. URL https://doi.org/10.1016/S1544-6123(03)00003-5.

Kitao, Sagiri. 2010. "Labor-Dependent Capital Income Taxation." *Journal of Monetary Economics* 57 (8):959–974. URL https://www.sciencedirect.com/science/article/pii/S0304393210001030.

Kochen, Federico. 2022. "Finance Over the Life Cycle of Firms." Working paper, New York University. URL https://www.ecb.europa.eu/pub/conferences/ecbforum/shared/pdf/2022/kochen_paper.en.pdf.

Kopczuk, Wojciech. 2013. "Taxation of Intergenerational Transfers and Wealth." In *Handbook of Public Economics, Volume 5*, *Handbook of Public Economics*, vol. 5, edited by Alan J. Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez, chap. 6. Elsevier, 329–390.

Krusell, Per and Anthony A. Smith, Jr. 1998. "Income and Wealth Heterogeneity in the Macroeconomy." *Journal of Political Economy* 106 (5):867–896. URL https://doi.org/10.1086/250034.

Liang, James, Hui Wang, and Edward P. Lazear. 2018. "Demographics and Entrepreneurship." *Journal of Political Economy* 126 (S1):S140–S196. URL https://doi.org/10.1086/698750.

Lucas, Robert E. 1978. "On the Size Distribution of Business Firms." *The Bell Journal of Economics* 9 (2):508–523. URL http://www.jstor.org/stable/3003596.

———. 1987. *Models of Business Cycles*. New York: Basil Blackwell.

McDaniel, Cara. 2007. "Average Tax Rates on Consumption, Investment, Labor and Capital in the OECD: 1950-2003." Arizona state university mimeo. URL https://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.716.7953.

Midrigan, Virgiliu and Daniel Yi Xu. 2014. "Finance and Misallocation: Evidence from Plant-Level Data." *American Economic Review* 104 (2):422–458. URL https://www.aeaweb.org/articles?id=10.1257/aer.104.2.422.

Moll, Benjamin. 2014. "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?" *American Economic Review* 104 (10):3186–3221. URL http://www.aeaweb.org/articles?id=10.1257/aer.104.10.3186.

OECD. 2018. "The Role and Design of Net Wealth Taxes in the OECD." *OECD Tax Policy Studies* (26). URL https://doi.org/10.1787/9789264290303-en.

———. 2021. "Revenue Statistics - the United States." URL http://oe.cd/revenue-statistics.

Piketty, Thomas. 2014. *Capital in the Twenty-First Century*. Cambridge: Belknap of Harvard UP.

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman. 2018. "Distributional National Accounts: Methods and Estimates for the United States." *Quarterly Journal of Economics* 133 (2):553–609. URL https://doi.org/10.1093/qje/qjx043.

Quadrini, Vincenzo. 2000. "Entrepreneurship, Saving, and Social Mobility." *Review of Economic Dynamics* 3 (1):1–40. URL https://doi.org/10.1006/redy.1999.0077.

Restuccia, Diego and Richard Rogerson. 2008. "Policy distortions and aggregate productivity with heterogeneous establishments." *Review of Economic Dynamics* 11 (4):707 – 720. URL http://www.sciencedirect.com/science/article/pii/S1094202508000203.

Saez, Emmanuel and Stefanie Stantcheva. 2018. "A simpler theory of optimal capital taxation." *Journal of Public Economics* 162:120–142. URL https://www.sciencedirect.com/science/article/pii/S0047272717301688. In Honor of Sir Tony Atkinson (1944-2017).

Saez, Emmanuel and Gabriel Zucman. 2016. "Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data." *Quarterly Journal of Economics* 131 (2):519–578. URL https://doi.org/10.1093/qje/qjw004.

Salgado, Sergio. 2020. "Technical Change and Entrepreneurship." Working paper, Wharton School at University of Pennsylvania. URL https://rodneywhitecenter.wharton.upenn.edu/wp-content/uploads/2021/03/10-21.Salgado.pdf.

Scheuer, Florian and Joel Slemrod. 2021. "Taxing Our Wealth." *Journal of Economic Perspectives* 35 (1):207–30. URL https://www.aeaweb.org/articles?id=10.1257/jep.35.1.207.

Smith, Matthew, Owen Zidar, and Eric Zwick. 2021. "Top Wealth in America: New Estimates and Implications for Taxing the Rich." Working paper, Princeton University. URL https://economics.princeton.edu/working-papers/top-wealth-in-america-new-estimates-and-implications-for-taxing-the-rich/.

———. forthcoming. "Top Wealth in America: New Estimates under Heterogeneous Returns." *Quarterly Journal of Economics* .

Solon, Gary. 1999. "Intergenerational Mobility in the Labor Market." In *Handbook of Labor Economics, Vol 3A*, vol. 3A, edited by Orley Ashenfelter and David Card. North-Holland, 1761–1800. URL https://doi.org/10.1016/S1573-4463(99)03010-2.

Stantcheva, Stefanie. 2020. "Dynamic Taxation." *Annual Review of Economics* 12:801–831. URL https://doi.org/10.1146/annurev-economics-100119-013035.

Straub, Ludwig and Iván Werning. 2020. "Positive Long-Run Capital Taxation: Chamley-Judd Revisited." *American Economic Review* 110 (1):86–119. URL https://www.aeaweb.org/articles?id=10.1257/aer.20150210.

Tørsløv, Thomas, Ludvig Wier, and Gabriel Zucman. 2022. "The Missing Profits of Nations." *The Review of Economic Studies* URL https://doi.org/10.1093/restud/rdac049.

US Department of the Treasury. 2017. "A Financial System that Creates Economic Opportunities: Capital Markets." URL https://www.treasury.gov/press-center/press-releases/documents/a-financial-system-capital-markets-final-final.pdf.

Vermeulen, Philip. 2018. "How Fat is the Top tail of the Wealth Distribution?" *Reivew of Income and Wealth* 64 (2):357–387. URL https://doi.org/10.1111/roiw.12279.

Wolff, Edward N. 2006. "Changes in Household Wealth in the 1980s and 1990s in the U.S." In *International Perspectives on Household Wealth*, edited by Edward N. Wolff, chap. 4. Eduaward Elgar Publishing, 107,150. URL https://www.e-elgar.com/shop/international-perspectives-on-household-wealth?___website=uk_warehouse.

ONLINE APPENDIX

A Model Details and Additional Equations

A.1 Social Security Pension System

When an individual retires at age R, she starts receiving social security income $y^R(\kappa, e)$ that depends on her type κ in the following way:

$$y^R(\kappa, e) = \Phi(\kappa, e)\,\overline{y}, \tag{A.1}$$

where Φ is the replacement ratio. The replacement ratio is progressive and given by

$$\Phi(\kappa, e) = \begin{cases} 0.9\frac{y_1^R(\kappa,e)}{\overline{y}_1^R} & \text{if } \frac{y_1^R(\kappa,e)}{\overline{y}_1^R} \leqslant 0.3 \\ 0.27 + 0.32\left(\frac{y_1^R(\kappa,e)}{\overline{y}_1^R} - 0.3\right) & \text{if } 0.3 < \frac{y_1^R(\kappa,e)}{\overline{y}_1^R} \leqslant 2 \\ 0.81 + 0.15\left(\frac{y_1^R(\kappa,e)}{\overline{y}_1^R} - 2\right) & \text{if } 2 < \frac{y_1^R(\kappa,e)}{\overline{y}_1^R} \leqslant 4.1 \\ 1.13 & \text{if } 4.1 < \frac{y_1^R(\kappa,e)}{\overline{y}_1^R} \end{cases} \tag{A.2}$$

where $y_1^R(\kappa, e)$ is the average efficiency units over lifetime that an individual of type κ gets conditional on having a given $e_R = e$:

$$y_1^R(\kappa, e_R) = \frac{1}{R}\int_{h<R,a,\mathbf{S}} y_h(\kappa, e)\, d\Gamma(h, a, \mathbf{S}). \tag{A.3}$$

The vector $\mathbf{S} = (\overline{z}, \mathbb{I}, \kappa, e)$ is the vector of exogenous states of an individual, and the integral is taken with respect to the stationary distribution (Γ) of individuals so that e_R is the one given on the left-hand side. Finally, $\overline{y}_1^R$ is the average of $y_1^R(\kappa, e)$ across κ and e. The term SSP denotes the aggregate value of "social security pension" payments:

$$SSP \equiv \int_{h\geqslant R,a,\mathbf{S}} y^R(\kappa, e)\, d\Gamma(h, a, \mathbf{S}). \tag{A.4}$$

A.2 Recursive Competitive Equilibrium

Definition. *Let* $c_h(a, \mathbf{S})$, $\ell_h(a, \mathbf{S})$, $a_{h+1}(a, \mathbf{S})$, *and* $k(a, z)$ *denote the optimal decision rules and* $\Gamma(h, a, \mathbf{S})$ *denote the stationary distribution of individuals. A recursive competitive equilibrium is given by the following conditions:*

1. *Consumers maximize utility given* $p(x)$, $\overline{w}$, r, *and taxes.*

2. *The solution to the final goods producer gives the pricing function,* $p(x)$, *and wage rate,* $\overline{w}$.

3. $Q = \left(\int_{h,a,\mathbf{S}} (z \times k(a, z))^{\mu}\, d\Gamma(h, a, \mathbf{S})\right)^{1/\mu}$ *and* $L = \int_{h,a,\mathbf{S}} (w_h(\kappa, e)\,\ell_h(a, \mathbf{S}))\, d\Gamma(h, a, \mathbf{S})$, *where* $\log w_h(\kappa, e) = \kappa + g(h) + e$.

4. *The bond market clears:*

$$0 = \int_{h,a,\mathbf{S}} (a - k(a,z))\, d\Gamma(h, a, \mathbf{S}). \tag{A.5}$$

5. *The government budget balances. The revenue raised by taxes on labor, consumption, bequests, and capital income or wealth equals government consumption,* G*, plus pension payments to retirees, SSP:*

$$\begin{aligned} G + SSP = \tau_k &\int_{h,a,\mathbf{S}} (ra + \pi(a,z))\, d\Gamma(h, a, \mathbf{S}) \\ + \tau_a &\int_{h,a,\mathbf{S}} (a)\, d\Gamma(h, a, \mathbf{S}) \\ + \tau_\ell &\int_{h<R,a,\mathbf{S}} (\overline{w} w_h(\kappa, e)\, \ell_h(a, \mathbf{S}))\, d\Gamma(h, a, \mathbf{S}) \\ + \tau_c &\int_{h,a,\mathbf{S}} c_h(a, \mathbf{S})\, d\Gamma(h, a, \mathbf{S}), \\ + \tau_b &\int_{h,a,\mathbf{S}} (1 - s_{h+1})\, a_{h+1}(a, \mathbf{S})\, d\Gamma(h, a, \mathbf{S}), \end{aligned} \tag{A.6}$$

where $\tau_a \equiv 0$ *in the capital income tax economy and* $\tau_k \equiv 0$ *in the wealth tax economy, and SSP is given in equation (A.4).*

A.3 Formulas for Welfare Analyses

A.3.1 Formulas for Section 5.2

The formulas that define CE_1 and $\overline{CE}_2$, introduced in Section 5.2, are as follows. We compute CE_1 for an h-year-old individual in state $\mathcal{S} = (a, \mathbf{S})$ as the percentage change in consumption at all future dates and states required to make her indifferent between the stationary equilibria of the two economies:

$$V_h^{US}\left((1 + CE_1(h, \mathcal{S})) \times c_h^{US}(\mathcal{S}), \ell_h^{US}(\mathcal{S})\right) = V_h^{RN}\left(c_h^{RN}(\mathcal{S}), \ell_h^{RN}(\mathcal{S})\right), \tag{A.7}$$

where V_h is the lifetime value function and (c, ℓ) are the consumption and leisure allocations starting from state $(h, \mathcal{S})$, and the superscripts indicate the relevant economy (e.g., US versus RN). At the aggregate level, the main measure we will look at is the welfare change for newborns, which is obtained by integrating over the stationary distribution in the benchmark economy $(\Gamma_{h=1}^{US}(\mathcal{S}))$:

$$\overline{CE}_1 \equiv \sum_{\mathcal{S}} \Gamma^{US}(1, \mathcal{S}) \times CE_1(1, \mathcal{S}). \tag{A.8}$$

Using equation 1 and the fact that $u(c, \ell) = \frac{\left(c^\gamma (1-\ell)^{1-\gamma}\right)^{1-\sigma}}{1-\sigma}$, we can compute $CE_1(h, \mathcal{S})$ directly from the value functions: $1 + CE_1(h, \mathcal{S}) = \left(\frac{V_h^R(\mathcal{S}) - B_h^{US}(\mathcal{S})}{V_h^{US}(\mathcal{S}) - B_h^{US}(\mathcal{S})}\right)^{1/\gamma(1-\sigma)}$, where $V_h^{US}(\mathcal{S})$ is the value function of an agent of age h at state $\mathcal{S}$ and $B_h(\mathcal{S})$ is the expected discounted value of the utility payoff of bequests (thus, $V_h^{US}(\mathcal{S}) - B_h^{US}(\mathcal{S})$ gives the value coming just from consumption and leisure).

$\overline{CE}_2$ measures the fixed proportional consumption transfer to all newborn individuals in the US benchmark economy such that average utility is equal to that in the tax-reform economy. For the RN reform, it reads

$$\sum_{\mathcal{S}} \Gamma^{US}(1,\mathcal{S}) \cdot V_1^{US}\left(\left(1+\overline{CE}_2\right) c_1^{US}(\mathcal{S}), \ell_1^{US}(\mathcal{S})\right) = \sum_{\mathcal{S}} \Gamma^{RN}(1,\mathcal{S}) \cdot V_1^{RN}\left(c_1^{RN}(\mathcal{S}), \ell_1^{RN}(\mathcal{S})\right). \quad \text{(A.9)}$$

We can get a closed form expression for the welfare gain by first defining the average expected discounted value of the utility payoff of bequests for newborn agents:

$$\overline{B}_1^{US}(\mathcal{S}) \equiv \sum_{\mathcal{S}} \Gamma^{US}(1,\mathcal{S}) \cdot B_1^{US}(\mathcal{S}). \quad \text{(A.10)}$$

This allows us to get an expression for the welfare gain:

$$1+\overline{CE}_2 = \left(\frac{\sum_{\mathcal{S}} \Gamma^{RN}(1,\mathcal{S}) \cdot V_1^{RN}(\mathcal{S}) - \overline{B}_1^{US}(\mathcal{S})}{\sum_{\mathcal{S}} \Gamma^{US}(1,\mathcal{S}) \cdot V_1^{US}(\mathcal{S}) - \overline{B}_1^{US}(\mathcal{S})}\right)^{\frac{1}{\gamma(1-\sigma)}}. \quad \text{(A.11)}$$

A.3.2 Formulas for the Welfare Decomposition in Section 6.3

We derive in this section the formulas for decomposing $\overline{CE}_2$. The formulas for CE_1 are analogous.

Level-Distribution Decomposition

The welfare gain from changes in consumption and leisure can be jointly decomposed into gains from changes in levels and gains from changes in the distribution.

Consumption and Leisure Level To construct this measure, define first an alternative consumption policy that takes into account just the change in the level of aggregate consumption:

$$\hat{c}_h(\mathcal{S}) \equiv \frac{C^{RN}}{C^{US}} c_h^{US}(\mathcal{S}) \qquad \text{where } C^x \equiv \sum_{h,\mathcal{S}} c_h^x(\mathcal{S})\,\Gamma^x(h,\mathcal{S}) \text{ for } x \in \{\text{US,RN}\}. \quad \text{(A.12)}$$

Similarly, define the alternative policy for leisure as

$$\hat{\ell}_h(\mathcal{S}) \equiv \frac{L^{RN}}{L^{US}} \ell_h^{US}(\mathcal{S}) \qquad \text{where } L^x \equiv \sum_{h,\mathcal{S}} \ell_h^x(\mathcal{S})\,\Gamma^x(h,\mathcal{S}) \text{ for } x \in \{\text{US,RN}\}. \quad \text{(A.13)}$$

The level gain is obtained by equating the welfare under the benchmark policies and the alternative policies defined above, while keeping constant bequests in the two economies:

$$\sum_{\mathcal{S}} \Gamma^{US}(1,\mathcal{S}) \cdot V_1^{US}\left(\left(1+\overline{CE}_2^{L}\right) c_1^{US}(\mathcal{S}), \ell_1^{US}(\mathcal{S}), b_1^{US}(\mathcal{S})\right)$$
$$= \sum_{\mathcal{S}} \Gamma^{RN}(1,\mathcal{S}) \cdot V_1^{RN}\left(\hat{c}_1(\mathcal{S}), \hat{\ell}_1(\mathcal{S}), b_1^{US}(\mathcal{S})\right). \tag{A.14}$$

Given the preferences, we assume this gives

$$\begin{aligned}
1+\overline{CE}_2^{L} &= \left(\frac{\sum_{\mathcal{S}} \Gamma^{US}(1,\mathcal{S}) \cdot V_1\left(\hat{c}_1(\mathcal{S}), \hat{\ell}_1(\mathcal{S}), b_1^{US}(\mathcal{S})\right) - \overline{B_1^{US}(\mathcal{S})}}{\sum_{\mathcal{S}} \Gamma^{US}(1,\mathcal{S}) \cdot V_1^{US}(\mathcal{S}) - \overline{B_1^{US}(\mathcal{S})}}\right)^{\frac{1}{\gamma(1-\sigma)}} \\
&= \frac{C^{RN}}{C^{US}}\left(\frac{L^{RN}}{L^{US}}\right)^{\frac{1-\gamma}{\gamma}} \left(\frac{\sum_{\mathcal{S}} \Gamma^{US}(1,\mathcal{S}) \cdot V_1\left(c_1^{US}(\mathcal{S}), \ell_1^{US}(\mathcal{S}), b_1^{US}(\mathcal{S})\right) - \overline{B}_1^{US}(\mathcal{S})}{\sum_{\mathcal{S}} \Gamma^{US}(1,\mathcal{S}) \cdot V_1^{US}(\mathcal{S}) - \overline{B}_1^{US}(\mathcal{S})}\right)^{\frac{1}{\gamma(1-\sigma)}} \\
&= \frac{C^{RN}}{C^{US}}\left(\frac{L^{RN}}{L^{US}}\right)^{\frac{1-\gamma}{\gamma}}.
\end{aligned} \tag{A.15}$$

Consumption and Leisure Distribution The distributional gains correspond to the change in the value of agents from adjusting the policy functions while keeping the level comparable. Once again we keep the value of bequests fixed:

$$\sum_{\mathcal{S}} \Gamma^{US}(1,\mathcal{S}) \cdot V_1^{US}\left(\left(1+\overline{CE}_2^{D}\right) \hat{c}_1(\mathcal{S}), \hat{\ell}_1(\mathcal{S}), b_1^{US}(\mathcal{S})\right)$$
$$= \sum_{\mathcal{S}} \Gamma^{RN}(1,\mathcal{S}) \cdot V_1^{RN}\left(c_1^{RN}(\mathcal{S}), \ell_1^{RN}(\mathcal{S}), b_1^{US}(\mathcal{S})\right). \tag{A.16}$$

Given the preferences we assume, this gives

$$\begin{aligned}
1+\overline{CE}_2^{D} &= \left(\frac{\sum_{\mathcal{S}} \Gamma^{US}(1,\mathcal{S}) \cdot V_1\left(c_1^{RN}(\mathcal{S}), \ell_1^{RN}(\mathcal{S}), b_1^{US}(\mathcal{S})\right) - \overline{B}_1^{US}(\mathcal{S})}{\sum_{\mathcal{S}} \Gamma^{US}(1,\mathcal{S}) \cdot V_1\left(\hat{c}_1(\mathcal{S}), \hat{\ell}_1(\mathcal{S}), b_1^{US}(\mathcal{S})\right) - \overline{B}_1^{US}(\mathcal{S})}\right)^{\frac{1}{\gamma(1-\sigma)}} \\
&= \left[\frac{C^{RN}}{C^{US}}\left(\frac{L^{RN}}{L^{US}}\right)^{\frac{1-\gamma}{\gamma}}\right]^{-1} \times \left(\frac{\sum_{\mathcal{S}} \Gamma^{US}(1,\mathcal{S}) \cdot V_1\left(c_1^{RN}(\mathcal{S}), \ell_1^{RN}(\mathcal{S}), b_1^{US}(\mathcal{S})\right) - \overline{B}_1^{US}}{\sum_{\mathcal{S}} \Gamma^{US}(1,\mathcal{S}) \cdot V_1^{US}(\mathcal{S}) - \overline{B}_1^{US}}\right) \\
&= \frac{1}{1+\overline{CE}_2^{L}}\left(\frac{\sum_{\mathcal{S}} \Gamma^{US}(1,\mathcal{S}) \cdot V_1\left(c_1^{RN}(\mathcal{S}), \ell_1^{RN}(\mathcal{S}), b_1^{US}(\mathcal{S})\right) - \overline{B}_1^{US}}{\sum_{\mathcal{S}} \Gamma^{US}(1,\mathcal{S}) \cdot V_1^{US}(\mathcal{S}) - \overline{B}_1^{US}}\right)^{\frac{1}{\gamma(1-\sigma)}} \\
&= \frac{1+\overline{CE}_2^{c\ell}}{1+\overline{CE}_2^{L}},
\end{aligned} \tag{A.17}$$

where we define

$$1+\overline{\mathsf{CE}}_2^{c\ell} \equiv \left(\frac{\sum_{\mathcal{S}} \Gamma^{\mathrm{US}}(1,\mathcal{S}) \cdot V_1\left(c_1^{RN}(\mathcal{S}), \ell_1^{RN}(\mathcal{S}), b_1^{US}(\mathcal{S})\right) - \overline{B}_1^{\mathrm{US}}(\mathcal{S})}{\sum_{\mathcal{S}} \Gamma^{\mathrm{US}}(1,\mathcal{S}) \cdot V_1^{\mathrm{US}}(\mathcal{S}) - \overline{B}_1^{\mathrm{US}}(\mathcal{S})}\right)^{\frac{1}{\gamma(1-\sigma)}} \tag{A.18}$$

as the joint gain from consumption and leisure. By construction we can decompose the value form consumption and leisure into the level and distributional changes:

$$1+\overline{\mathsf{CE}}_2^{c\ell} = \left(1+\overline{\mathsf{CE}}_2^{L}\right)\left(1+\overline{\mathsf{CE}}_2^{D}\right). \tag{A.19}$$

Complete Decomposition

To totally decompose the level of the consumption-equivalent welfare gain, we need to take into account the change in bequests. This is

$$\begin{aligned}
1+\overline{\mathrm{CE}}_2 &= \left(\frac{\sum_{\mathcal{S}} \Gamma^{\mathrm{RN}}(1,\mathcal{S}) \cdot V_1^{\mathrm{RN}}(\mathcal{S}) - \overline{B}_1^{\mathrm{US}}(\mathcal{S})}{\sum_{\mathcal{S}} \Gamma^{\mathrm{US}}(1,\mathcal{S}) \cdot V_1^{\mathrm{US}}(\mathcal{S}) - \overline{B}_1^{\mathrm{US}}(\mathcal{S})}\right)^{\frac{1}{\gamma(1-\sigma)}} \\
1+\overline{\mathrm{CE}}_2 &= \left(\frac{\sum_{\mathcal{S}} \Gamma^{\mathrm{US}}(1,\mathcal{S}) \cdot V_1\left(c^{RN}(1,\mathcal{S}), \ell^{RN}(1,\mathcal{S}), b^{US}(1,\mathcal{S})\right) - \overline{B}_1^{\mathrm{US}}(\mathcal{S})}{\sum_{\mathcal{S}} \Gamma^{\mathrm{US}}(1,\mathcal{S}) \cdot V_1^{\mathrm{US}}(\mathcal{S}) - \overline{B}_1^{\mathrm{US}}(\mathcal{S})}\right)^{\frac{1}{\gamma(1-\sigma)}} \times \\
&\quad \left(\frac{\sum_{\mathcal{S}} \Gamma^{\mathrm{RN}}(1,\mathcal{S}) \cdot V_1^{\mathrm{RN}}(\mathcal{S}) - \overline{B}_1^{\mathrm{US}}(\mathcal{S})}{\sum_{\mathcal{S}} \Gamma^{\mathrm{US}}(1,\mathcal{S}) \cdot V_1\left(c_1^{RN}(\mathcal{S}), \ell_1^{RN}(\mathcal{S}), b_1^{US}(\mathcal{S})\right) - \overline{B}_1^{\mathrm{US}}(\mathcal{S})}\right)^{\frac{1}{\gamma(1-\sigma)}} \\
1+\overline{\mathrm{CE}}_2 &= \left(1+\overline{\mathrm{CE}}_2^{c\ell}\right)\left(1+\overline{\mathrm{CE}}_2^{b}\right).
\end{aligned} \tag{A.20}$$

B Additional Tables

Table B.1 – *Forbes* Self-Made Index

	Description	Fraction
1	Inherited fortune but not working to increase it	7.00
2	Inherited fortune and has a role managing it	4.75
3	Inherited fortune and helping to increase it marginally	5.50
4	Inherited fortune and increasing it in a meaningful way	5.25
5	Inherited small or medium-size business and made it into a 10-digit fortune	8.50
6	Hired or hands-off investor who didn't create the business	2.25
7	Self-made who got a head start from wealthy parents & moneyed background	10.00
8	Self-made who came from a middle- or upper-middle-class background	**32.00**
9	Self-made who came from a largely working-class background; rose from little to nothing	**14.50**
10	Self-made who not only grew up poor but also overcame significant obstacles	**7.75**
	Forbes's definition of self-made: Groups 8 to 10	**54.25**

Notes: Table reports *Forbes*'s categories for classifying individuals in its top-400 list, along with their share among the individuals in the list. Self-made individuals correspond to categories 8, 9, and 10.

B.1 Additional Results on the Distribution of Capital Income and Wealth

The benchmark model is consistent with the high concentration of capital income in the economy. Table B.2 shows that the concentration of capital income is higher than the concentration of wealth in the model. For instance, those in the top 1% of the wealth distribution hold 35.1% of all the wealth and 48.2% of all the capital income in the economy. Those in the top 1% of the capital income distribution hold 51.9% of all the capital income in the economy. Furthermore, the Gini coefficients of wealth and capital income are 0.78 and 0.87, respectively.

Simultaneously, capital income and wealth are highly correlated in the model. The correlation coefficient is 0.85. This is consistent with Table B.2, which shows that capital income is concentrated among the wealthiest individuals in the economy. It is also consistent with the correlation of the returns to wealth and wealth levels. In particular, for the 35-49 age group, returns are in the 5%-6.1% range in the bottom half of the wealth distribution but increase to 6.5% at the 60th percentile, 7.3% at the 95th, 8.4% at 99th, 11.4% at the 99.9th, and 12.7% at the 100th. The same patterns arise later in the life cycle of individuals but with lower levels of returns.

Table B.2 – Concentration of Capital Income and Wealth

Top x% of Wealth Dist.	Wealth Share (%)	Capital Income Share (%)	Top x% of Capital Income Dist.	Capital Income Share (%)
0.1	22.3	32.0	0.1	34.3
0.5	30.5	43.0	0.5	45.7
1	35.1	48.2	1	51.9
10	64.9	73.1	10	78.9
50	96.4	97.0	50	98.1

Notes: The table describes the concentration of the wealth and capital income distribution. The left panel reports top wealth shares and the corresponding capital income shares of individuals in the respective group of top-wealth holders. The right panel reports the capital income shares of top-capital-income earners. All numbers in percentage points.

B.2 Additional Results on the Distribution of Welfare Gains/Losses

TABLE B.3 – Optimal Tax Experiments: Distribution of Welfare Gains and Losses

(A) Optimal Capital Income Taxes

	Distribution of Welfare Gains and Losses						Fraction with Positive Welfare Gain					
	Ability Groups ($\bar{z}_i$ *Percentiles)*						*Ability Groups* ($\bar{z}_i$ *Percentiles)*					
	0-40	40-80	80-90	90-99	99-99.9	99.9+	0-40	40-80	80-90	90-99	99-99.9	99.9+
20	3.4	3.8	5.1	7.5	11.4	13.8	99.6	98.8	99.8	99.9	100.0	100.0
21–34	3.3	3.6	4.7	7.0	11.2	13.9	99.7	99.1	99.8	99.9	100.0	100.0
35–49	2.9	2.8	3.5	4.8	7.1	8.7	99.4	98.0	99.6	99.8	100.0	100.0
50–64	1.6	1.5	1.9	2.7	3.8	4.6	97.8	94.9	99.3	99.6	99.9	99.9
65+	0.1	0.2	0.4	0.9	1.6	1.9	94.5	96.3	99.8	99.9	99.9	99.8

(B) Optimal Wealth Taxes

	Distribution of Welfare Gains and Losses						Fraction with Positive Welfare Gain					
	Ability Groups ($\bar{z}_i$ *Percentiles)*						*Ability Groups* ($\bar{z}_i$ *Percentiles)*					
	0-40	40-80	80-90	90-99	99-99.9	99.9+	0-40	40-80	80-90	90-99	99-99.9	99.9+
20	9.4	8.3	8.3	10.1	13.9	16.3	97.5	94.6	94.3	95.8	97.9	98.8
21–34	8.7	6.8	5.8	6.4	8.0	8.6	97.6	92.9	90.0	90.2	89.7	87.0
35–49	6.3	4.1	2.4	1.6	–0.4	–2.3	93.6	80.4	71.0	64.5	52.6	42.4
50–64	2.5	1.0	–0.1	–1.2	–3.4	–5.2	74.9	62.5	52.9	45.3	34.5	27.6
65+	–0.5	–0.9	–1.3	–1.9	–3.1	–4.3	1.4	0.8	0.7	0.7	0.6	0.4

(C) Optimal Wealth Taxes with Exemption Threshold

	Distribution of Welfare Gains and Losses						Fraction with Positive Welfare Gain					
	Ability Groups ($\bar{z}_i$ *Percentiles)*						*Ability Groups* ($\bar{z}_i$ *Percentiles)*					
	0-40	40-80	80-90	90-99	99-99.9	99.9+	0-40	40-80	80-90	90-99	99-99.9	99.9+
20	9.4	8.5	8.6	10.4	14.3	17.0	97.2	93.1	92.4	95.0	97.5	98.3
21–34	8.7	6.8	5.8	6.3	7.7	8.1	97.3	91.3	86.9	87.4	87.0	83.7
35–49	6.3	3.9	2.0	0.9	–1.7	–4.3	92.4	78.7	67.6	60.5	48.2	38.4
50–64	2.6	1.1	–0.3	–1.7	–4.6	–7.0	78.7	66.3	56.4	48.0	36.2	28.9
65+	–0.3	–0.7	–1.1	–2.0	–3.7	–5.3	79.8	73.3	65.1	56.6	43.8	35.4

Notes: Each panel reports the average welfare gain (CE_1) and the share of individuals who experience a positive welfare gain (CE_1) in a given age and entrepreneurial productivity group (ranked based on the permanent component of entrepreneurial productivity $\bar{z}$) from the corresponding optimal tax experiment. The average and shares are computed with respect to the benchmark distribution. All numbers are in percentage points.

Table B.4 – Welfare Change *with Transition*: Switch to Optimal Tax System with Transition

(a) Optimal Capital Income Taxes

	Distribution of Welfare Gains and Losses						Fraction with Positive Welfare Gain					
	Ability Groups ($\bar{z}_i$ Percentiles)						*Ability Groups ($\bar{z}_i$ Percentiles)*					
	0-40	40-80	80-90	90-99	99-99.9	99.9+	0-40	40-80	80-90	90-99	99-99.9	99.9+
20	-8.8	-7.5	-4.8	0.2	8.7	13.8	4.1	8.2	17.0	27.3	71.8	99.6
21–34	-8.2	-5.9	-1.9	5.7	19.8	30.2	3.5	10.7	28.4	58.9	81.0	84.8
35–49	-6.3	-3.9	0.0	6.5	18.5	27.1	8.6	20.2	47.8	58.9	69.9	75.0
50–64	-3.1	-1.3	1.3	5.2	12.2	17.0	26.5	37.9	54.4	60.7	69.6	75.3
65+	0.6	1.2	2.2	4.0	7.0	9.1	99.6	100	100	100	100	100

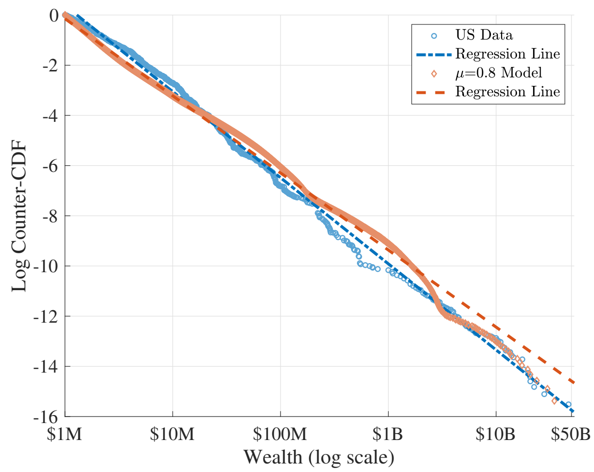
(b) Optimal Wealth Taxes

	Distribution of Welfare Gains and Losses						Fraction with Positive Welfare Gain					
	Ability Groups ($\bar{z}_i$ Percentiles)						*Ability Groups ($\bar{z}_i$ Percentiles)*					
	0-40	40-80	80-90	90-99	99-99.9	99.9+	0-40	40-80	80-90	90-99	99-99.9	99.9+
20	5.4	4.9	5.6	8.4	13.5	16.7	95.7	93.8	95.0	97.7	99.5	99.7
21–34	4.8	3.8	3.9	6.0	10.0	12.1	95.6	90.6	90.5	93.5	94.9	94.2
35–49	2.9	1.7	1.1	1.5	1.6	1.0	84.6	72.8	67.3	69.4	67.8	64.5
50–64	0.5	-0.3	-0.8	-1.1	-2.2	-3.4	59.8	50.6	44.1	42.4	38.6	35.9
65+	-0.7	-0.9	-1.1	-1.4	-2.5	-3.7	3.2	5.5	6.9	9.0	10.9	11.5

Notes: Each panel reports the average welfare gain (CE_1) and the share of individuals who experience a positive welfare gain (CE_1) in a given age and entrepreneurial productivity group (ranked based on the permanent component of entrepreneurial productivity $\bar{z}$) who are alive at the time of the corresponding tax experiment with transition. The average and shares are computed with respect to the benchmark distribution. All numbers are in percentage points.

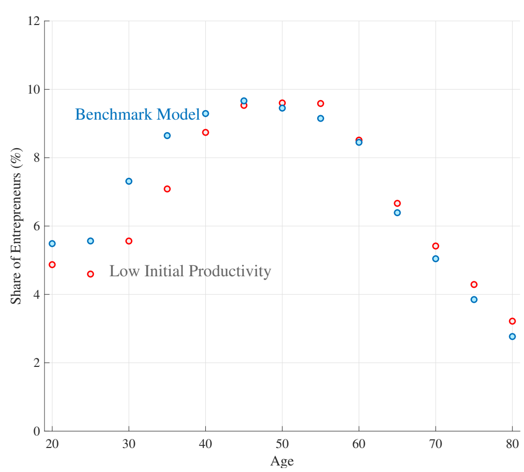
C Additional Figures

Figure C.1 – Stronger Diminishing Returns in Entrepreneurial Production, $\mu = 0.8$



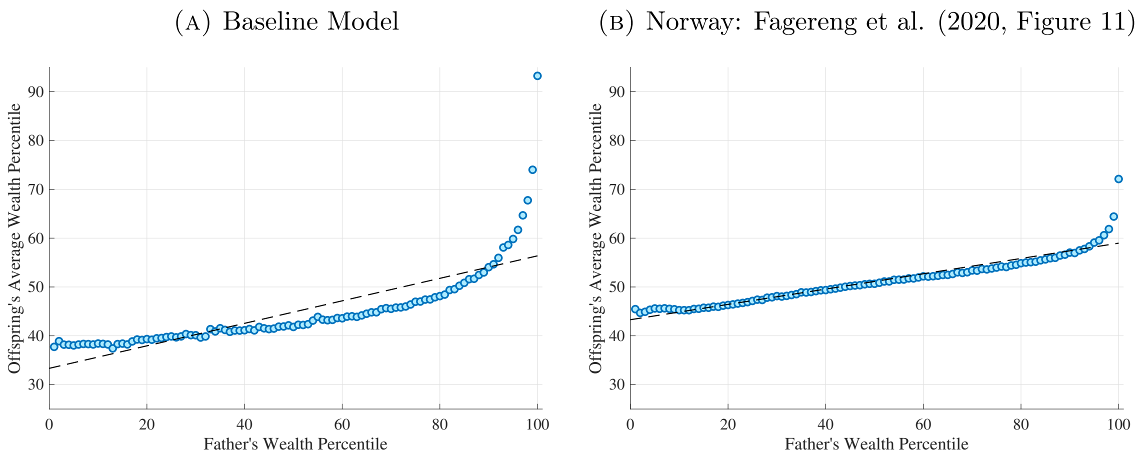
Notes: The figure reports the natural logarithm of the counter-CDF of wealth above one million dollars, corresponding to the right tail of the distribution. The data for the US (in blue) come from Vermeulen (2018). The orange diamonds correspond an alternative calibration of the model with $\mu = 0.8$. Both axes are in natural logs. The horizontal axis ticks are placed at powers of 10 for readability.

Figure C.2 – Fraction of Entrepreneurs over the Life Cycle



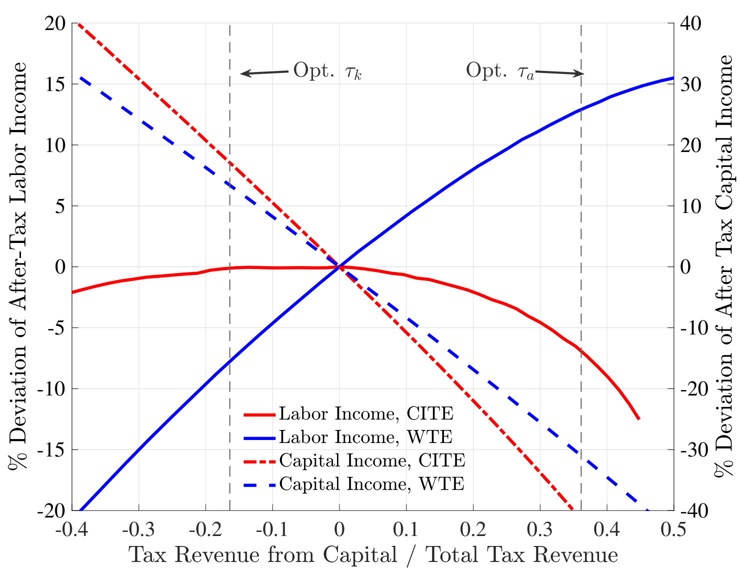
Notes: The figure plots the fraction of entrepreneurs, in percent, over the life cycle. An entrepreneur is defined as someone who earns more than 50% of their income from their business. The low-initial-productivity model has the same productivity shock process as in the benchmark, except that nobody starts in the fast lane ($z_{i0} = \bar{z}_i$) but those with $\bar{z}_i$ above median have a 3% probability of entering the fast lane each period.

Figure C.3 – Intergenerational Rank-Rank Correlation in Wealth: Model vs. Data



Notes: The figures report the rank-rank plots of wealth between fathers and their offspring for the baseline model and Norwegian data. The blue circles mark the average percentile within the cohort of the offspring of fathers in a given percentile (rank) of the wealth distribution. The dashed line corresponds to the trend line. The Norwegian data come from Fagereng et al. (2020, Figure 11).

Figure C.4 – Average After-Tax Labor and Capital Income vs. Capital Tax Revenues



Notes: The figure reports the percentage change of total labor (solid line) and capital income (dashed line) with respect to the benchmark economy for different levels capital income tax (CITE in red) or wealth tax (WTE in blue). For each level of the tax the labor income tax adjusts to balance the government's budget. Welfare gains are in percentages. Each economy is indexed by its ratio of tax revenue from capital income or wealth taxes to total revenue. Total revenue is constant across economies.

D Misallocation in the Economy

Our model economy is distorted because of the existence of financial frictions in the form of borrowing constraints. We can measure the effects of these distortions on aggregate TFP and output, following a large and growing literature that frames the discussion on misallocation in terms of various wedges, such as capital, labor, and output wedges. In particular, we follow Hsieh and Klenow (2009) and compute measures of misallocation for our model economy.

Instead of modeling and capturing the effect of a particular distortion, or distortions, we infer the underlying distortions and wedges in the economy by studying the extent to which the marginal revenue products of capital and labor differ across firms. This is based on the insight that without any distortions, the marginal revenue products of capital and labor have to be equalized across all firms.[53]

In our model, competitive final goods producers use effective capital, Q, and labor, L, in production as in (9), where Q is produced using intermediate goods as in (10). Each intermediate goods producer i produces $x_i = z_i k_i$, where z_i is i's entrepreneurial ability and k_i is capital.

TFP in the Q sector. We will first focus on the intermediate goods sector. Under the alternative capital-wedge approach, the problem of each intermediate goods producer is

$$\pi_i = \max_{k_i} p\left(z_i k_i\right) z_i k_i - \left(1+\tau_i\right)\left(R+\delta\right) k_i \ , \tag{D.1}$$

where τ_i is a firm-specific wedge. There are no collateral constraints. There is only one input and, as a result, only one wedge can be identified.

The revenue TFP in sector Q for each firm i is

$$\mathsf{TFPR}_{Q,i} \equiv \frac{p\left(x_i\right) x_i}{k_i} = \frac{1}{\mu}\left(1+\tau_i\right)\left(R+\delta\right). \tag{D.2}$$

The aggregate TFP in sector Q can be expressed as

$$\mathsf{TFP}_Q \equiv \frac{Q}{K} = \left(\int_i \left(z_i \frac{\overline{\mathsf{TFPR}_Q}}{\mathsf{TFPR}_{Q,i}}\right)^{\frac{\mu}{1-\mu}} di\right)^{\frac{1-\mu}{\mu}} , \tag{D.3}$$

where the average TFPR_Q is

$$\overline{\mathsf{TFPR}_Q} = \left(\int \frac{1}{\mathsf{TFPR}_{Q,i}} \frac{p\left(x_i\right) x_i}{p_q Q} di\right)^{-1} . \tag{D.4}$$

[53] This is the case in the monopolistic competition models, such as in Hsieh and Klenow (2009). Alternatively, in environments like the ones in Lucas (1978) and Restuccia and Rogerson (2008), in which firms feature decreasing returns to scale but produce the same homogeneous good, the marginal products of capital and labor have to be equalized in the non-distorted economy. See Hopenhayn (2014) for a review.

In the non-distorted economy, without capital wedges, the level of TFP in the Q sector is

$$\mathsf{TFP}_Q^\star = \left(\int_i (z_i)^{\frac{\mu}{1-\mu}}\, di\right)^{\frac{1-\mu}{\mu}} \equiv \bar{z}. \tag{D.5}$$

Therefore, we can measure the improvement in TFP in the Q sector, Ω_Q, as a result of eliminating the capital wedges, or equivalently, as a result of eliminating the collateral constraints:

$$\Omega_Q \equiv 1 - \frac{\mathsf{TFP}_Q}{\mathsf{TFP}_Q^\star} = 1 - \left(\int_i \left(\frac{\bar{z}}{z_i}\frac{\mathsf{TFPR}_{Q,i}}{\overline{\mathsf{TFPR}_Q}}\right)^{\frac{\mu}{1-\mu}} di\right)^{\frac{\mu-1}{\mu}}. \tag{D.6}$$

This measure does not capture the aggregate effect on the economy because (i) it applies only to the Q sector and not to the production of the final good, and (ii) it does not take into account changes in aggregate capital in the efficient economy with respect to the equilibrium of the distorted economy. In our benchmark economy, we obtain a value of $\Omega_Q = 0.35$, implying TFP gains of 35% in the Q sector coming from eliminating the collateral constraints.

Aggregate TFP. The final goods producers operate competitively and face no constraints or distortions, so there is no labor misallocation in the model. Because of this, the only source of misallocation and TFP losses is the Q sector. We can therefore write output as

$$Y = \mathrm{TFP} \cdot K^\alpha L^{1-\alpha}, \tag{D.7}$$

where $\mathrm{TFP} \equiv \mathrm{TFP}_Q^\alpha$ captures the aggregate TFP of the model. Similarly, we can define the efficient TFP level of the economy as $\mathrm{TFP}^\star \equiv (\mathrm{TFP}^\star)^\alpha$ and the aggregate TFP gain from eliminating distortions in the economy as

$$\Omega_Y \equiv 1 - \frac{\mathsf{TFP}}{\mathsf{TFP}^\star} = 1 - \left(\frac{\mathsf{TFP}_Q}{\mathsf{TFP}_Q^\star}\right)^\alpha. \tag{D.8}$$

In our benchmark, the total productivity gain from eliminating the collateral constraints in the Q sector amounts to 16% higher TFP.

Finally, we can use (D.7) to decompose the aggregate effect of tax reforms (say, the revenue-neutral reform) on output into the individual effects on TFP, the level of capital and the level of labor. We can write

$$\frac{Y^{\mathrm{RN}}}{Y^{\mathrm{US}}} = \frac{\mathrm{TFP}^{\mathrm{RN}}}{\mathrm{TFP}^{\mathrm{US}}}\left(\frac{K^{\mathrm{RN}}}{K^{\mathrm{US}}}\right)^\alpha \left(\frac{L^{\mathrm{RN}}}{L^{\mathrm{US}}}\right)^{1-\alpha}. \tag{D.9}$$

See Table E.8 for an application.

E Extensions and Robustness

E.1 Low-Inequality Calibration

Table E.5 – Welfare Change: L-INEQ Calibration

(a) Tax Reform

	Distribution of Welfare Gains and Losses						Fraction with Positive Welfare Gain					
	Ability Groups ($\bar{z}_i$ Percentiles)						*Ability Groups ($\bar{z}_i$ Percentiles)*					
	0-40	40-80	80-90	90-99	99-99.9	99.9+	0-40	40-80	80-90	90-99	99-99.9	99.9+
20	4.0	4.7	5.9	7.7	10.8	12.9	96.3	96.3	98.5	99.2	99.8	99.9
21–34	3.7	3.9	4.5	5.6	7.3	8.2	97.2	96.2	96.7	96.7	95.7	94.1
35–49	2.6	2.4	2.3	2.2	1.7	1.0	92.0	87.9	85.9	82.1	73.7	66.2
50–64	0.8	0.6	0.4	0.1	-0.7	-1.5	66.5	63.0	59.9	54.7	45.7	39.4
65+	-0.7	-0.6	-0.6	-0.8	-1.2	-1.6	2.5	9.9	11.7	11.5	10.6	9.5

(b) Optimal Wealth Taxes

	Distribution of Welfare Gains and Losses						Fraction with Positive Welfare Gain					
	Ability Groups ($\bar{z}_i$ Percentiles)						*Ability Groups ($\bar{z}_i$ Percentiles)*					
	0-40	40-80	80-90	90-99	99-99.9	99.9+	0-40	40-80	80-90	90-99	99-99.9	99.9+
20	5.4	5.7	6.8	8.7	12.1	14.5	94.5	93.2	96.1	97.5	99.0	99.5
21–34	4.9	4.5	4.7	5.5	6.9	7.2	94.9	91.7	92.6	91.9	89.8	86.7
35–49	3.2	2.2	1.7	1.0	-0.5	-2.1	84.7	75.6	72.2	65.1	53.2	43.6
50–64	0.5	-0.2	-0.6	-1.4	-3.1	-4.5	58.1	50.4	45.0	39.0	30.2	24.5
65+	-1.4	-1.5	-1.7	-2.0	-2.9	-3.8	0.6	2.9	3.9	3.8	3.3	2.8

Notes: Each panel reports the average welfare gain (CE_1) and the share of individuals who experience a positive welfare gain (CE_1) in a given age and entrepreneurial productivity group (ranked according to the permanent component of entrepreneurial productivity $\bar{z}$) from the corresponding tax experiment under the low inequality (L-INEQ) calibration. The average and shares are computed with respect to the benchmark distribution of the L-INEQ calibration. All numbers are in percentage points.

Table E.6 – Tax Reform: Change in Macro Variables from Low Inequality Calibration

	Change from the Benchmark of L-INEQ Calibration										
	Quantities (% Change)							*Prices (Change)*			
	K	Q	K/Y	TFP_Q	L	Y	C	$\overline{w}$	$\overline{w}$ (net)	$\Delta r^{\dagger}$	$\Delta r^{\dagger}$ (net)
Revenue-neutral reform	11.2	15.0	3.4	3.4	0.9	6.3	6.4	5.4	5.4	0.0	0.5
Opt. Wealth Taxes	4.0	8.3	4.1	4.1	2.5	4.8	5.9	2.2	7.9	–0.5	–1.1

Notes: RN refers to the revenue-neutral reform and OWT to the optimal wealth tax reform. Percentage changes are computed with respect to the low-inequality calibration, which has $\tau_k = 25\%$ and $\tau_a = 0\%$. †Changes in the interest rate are reported in percentage points. The net wage is defined as $(1-\tau_\ell)\,w$, and the net interest rate is defined as $(1-\tau_k)\,r$ or $r-\tau_a$, depending on the model. The TFP variable is measured in the intermediate goods market.

E.2 Incomplete Markets Model with "Awesome-State" Labor Income Shocks

We consider a version of our benchmark model without return heterogeneity and life-cycle demographics, which essentially becomes a perpetual-youth Aiyagari-style model. We introduce "awesome-state" idiosyncratic income shocks à la Castañeda, Díaz-Giménez, and Ríos-Rull (2003) and try to match their calibration and parameter choices as closely as we can.

In contrast to our benchmark model, there is no individual production of intermediate goods, and all output is produced by the competitive final goods producers that operate a technology

$$Y = K^{\alpha} L^{1-\alpha}, \tag{E.1}$$

where $K \equiv \int a_i di$ is the total amount of capital (or wealth) in the economy. Final good producers rent capital at a rate r and labor at a wage w. In equilibrium it holds that

$$r = \alpha \frac{Y}{K}; \qquad w = (1-\alpha)\,\frac{Y}{L}. \tag{E.2}$$

This production setup is equivalent to the one in our benchmark model when $z_i = \overline{z}\,(=1)$ for all individuals and $\mu = 1$, so there are no monopolistic rents in the production of intermediate goods. All individuals are therefore workers and have a common rate of return r.

We also change the life cycle of individuals to match that in Castañeda, Díaz-Giménez, and Ríos-Rull (2003). Workers are only subject to idiosyncratic labor efficiency shocks. In terms of our benchmark model, only e_i differs across individuals, with no type-dependent (κ_i) or age-dependent variation in labor income. The labor income of an individual is therefore $\overline{w} e_i \ell_i$, where ℓ_i is the endogenously determined labor supply. Workers retire with a constant probability p_{ret}. While retired, workers earn a retirement income ω_{ret} and die with a constant probability p_{death}. Only retirees can die. Upon death, individuals are replaced by a new worker (their descendant) that inherits their assets. In contrast to Castañeda, Díaz-Giménez, and Ríos-Rull (2003), there is no direct correlation between the worker's labor efficiency before retirement and that of their descendant.

We parametrize the model assuming that the labor efficiency shocks follow a discrete Markov process taking n_e values with a transition matrix Π^e. Newborn workers draw their initial labor efficiency from a distribution G^e. We take the number of states ($n_e = 4$) and the transition matrix between the states (Π^e) from Castañeda, Díaz-Giménez, and Ríos-Rull (2003, Table 4). We take the values of e_1, e_2 and e_3 from Castañeda, Díaz-Giménez, and Ríos-Rull (2003, Table 5). The value of e_{n_e} corresponds to the "awesome-state," and we set it to match a share of wealth held by the top 1% of 30%. We set $e_{n_e} = 265$. We take the values for $p_{ret} = 0.022$ and $p_{death} = 0.066$ from Castañeda, Díaz-Giménez, and Ríos-Rull (2003, Table 3). We set the value of ω_{ret} so as to obtain a ratio of transfers to GDP of 4.9% Castañeda, Díaz-Giménez, and Ríos-Rull (2003, pg. 837).

Finally, we set G^e according to *Step 2* of the procedure described in Castañeda, Díaz-Giménez, and Ríos-Rull (2003, Appendix A), overweighting the stationary distribution of labor efficiency γ^e (Castañeda, Díaz-Giménez, and Ríos-Rull, 2003, Table 5) with $\phi_2 = 0.525$ (Castañeda, Díaz-Giménez, and Ríos-Rull, 2003, Table 3) which controls the intergenerational earnings persistence. Accordingly, we set

$$G_1^e = \gamma_1^e + \phi_2\gamma_2^e + \phi_2^2\gamma_3^e + \phi_2^3\gamma_4^e, \tag{E.3}$$

$$G_2^e = (1 - \phi_2)\left(\gamma_2^e + \phi_2\gamma_3^e + \phi_2^2\gamma_4^e\right), \tag{E.4}$$

$$G_3^e = (1 - \phi_2)\left(\gamma_3^e + \phi_2\gamma_4^e\right), \tag{E.5}$$

$$G_4^e = (1 - \phi_2)\,\gamma_4^e. \tag{E.6}$$

To further facilitate comparison, we set the discount factor to $\beta = 0.924$ (Castañeda, Díaz-Giménez, and Ríos-Rull, 2003, Table 3). The remaining parameters of the model are left unchanged with respect to our benchmark.

E.3 Equilibrium with a Corporate Sector

Consider a model with two sectors: corporate and private. The goods of the two sectors are imperfect substitutes and are aggregated into a final good using a Cobb-Douglas technology:

$$Y = Y_c^{\rho} Y_p^{1-\rho}. \tag{E.7}$$

The corporate and private goods are also produced using Cobb-Douglas technologies:

$$Y_c = AK_c^{\alpha} L_c^{1-\alpha}, \qquad Y_p = Q^{\alpha} L_p^{1-\alpha}, \tag{E.8}$$

where $Q = \left(\int x_i^{\mu} di\right)^{1/\mu}$.

The corporate sector firms operate in perfect competition and face no financial constraints. There is a common market for labor with a price $\overline{w}$ per efficiency unit. There is a common capital market for corporate firms and private intermediate goods producers with an interest rate r.

The intermediate private goods x_i are sold by individual monopolists at a price $p(x_i)$ as described in the main text.

Equilibrium Conditions. The first order conditions of the final good aggregator are

$$p_c = \rho \left(\frac{Y_c}{Y_p}\right)^{-(1-\rho)} \qquad p_p = (1-\rho)\left(\frac{Y_c}{Y_p}\right)^{\rho}.$$

From these conditions, we get the following conditions for the expenditure shares across sectors:

$$\frac{p_c Y_c}{Y} = \rho \qquad \frac{p_p Y_p}{Y} = 1-\rho.$$

The first-order conditions of the corporate sector are

$$r + \delta = p_c \alpha A \left(\frac{K_c}{L_c}\right)^{-(1-\alpha)} \qquad \overline{w} = p_c\,(1-\alpha)\,A\left(\frac{K_c}{L_c}\right)^{\alpha}.$$

The first-order conditions of the private sector are

$$p\,(x_i) = p_p \alpha \left(\frac{x_i}{Q}\right)^{\mu-1}\left(\frac{Q}{L_p}\right)^{-(1-\alpha)} \qquad \overline{w} = p_p\,(1-\alpha)\left(\frac{Q}{L_p}\right)^{\alpha}.$$

The first-order conditions imply a relationship between corporate and private labor: $^{L_c}/_{L_p} = {}^{\rho}/_{1-\rho}$. This in turn implies a constant share of labor in the corporate sector of $^{L_c}/_{L} = {}^{L_c}/_{L_c+L_p} = \rho$.

Calibration. There are only two additional parameters. The share of the corporate sector in the production of the final good, ρ, and the productivity of the corporate sector, A.

Asker, Farre-Mensa, and Ljungqvist (2011) estimate that privately held US firms account for 69% of private sector employment, 59% of aggregate sales, and 53% of aggregate nonresidential fixed investment. In keeping with these estimates, we set $\gamma = 0.4$, which matches the share of the private sector in aggregate sales. Our calibration also implies a share of capital in the private sector of 50%, also in line with the data.

We keep the borrowing limit as in the benchmark, which gives a debt-to-asset ratio of 0.52. This is slightly higher than 0.45 reported by Asker, Farre-Mensa, and Ljungqvist (2011). Finally, we calibrate the remaining parameters to match the same moments as in our baseline in particular, a capital-to-GDP ratio of 3 and a top 1% share of 36%.

Benchmark Outcomes. It is instructive to discuss how the introduction of the corporate sector affects equilibrium outcomes and the calibration. The corporate sector increases the demand for capital, increasing the equilibrium interest rate. The higher interest rate reduces wealth inequality because low productivity entrepreneurs earn higher returns by lending in the bond market, while high productivity entrepreneurs earn lower returns because they face higher borrowing costs. Thus, the calibration produces a higher dispersion in entrepreneurial productivity to match the same top wealth concentration as in the data. These changes imply a slightly higher TFP_Q loss within the private sector (relative to that in our baseline). However, the TFP loss is substantially smaller at the aggregate level. As we illustrate next, the recalibrated model with the corporate sector gives very similar outcomes as our baseline model.

Table E.7 – Robustness: Optimal Wealth Tax

	Baseline		Corporate Sector	
	TR	OWT	TR	OWT
	(1)	(2)	(3)	(4)
τ_a	1.19	3.03	1.24	3.85
τ_ℓ	22.4	15.4	22.4	12.8
τ_k	—	—	—	—
	Change in Welfare (%)			
$\overline{\mathsf{CE}}_1$	6.8	9.0	6.1	9.5
$\overline{\mathsf{CE}}_2$	7.2	8.7	6.3	8.8
	Productivity			
$\frac{\text{TFP}^\star - \text{TFP}}{\text{TFP}^\star}$	0.14	0.13	0.09	0.08
	(Reference - US Economy)			
	0.16		0.11	

Notes: Columns (1) and (2) correspond to the tax reform and optimal wealth tax experiments in the baseline model. Columns (3) and (4) correspond to the tax reform and optimal wealth tax experiments for an extension with a corporate sector that operates a constant returns to scale technology and faces no financial constraints. The outputs of the corporate and private sectors are imperfect substitutes. The "Reference - US Economy" TFP corresponds to the level in the calibrated economy under the baseline US tax system, with $\tau_k = 0.25$ and $\tau_a = 0$.

Tax Reform and Optimal Wealth Taxes. The mechanisms that operate in our baseline model are also present in the extension with a corporate sector. An increase in wealth taxes favors the accumulation of capital by high-return individuals, who are themselves entrepreneurs producing in the private sector. Because of this reallocation of capital, productivity in the private sector improves, resulting in higher overall productivity and an increase in output in the private sector relative to that of the corporate sector. The outcome of the tax reform and optimal tax experiments is similar to those in the baseline model, in terms of both the level of taxes and the magnitude of the welfare gains. See Table E.7.

The optimal wealth tax is 3.8% and implies a welfare gain of 8.8% for newborn agents. The gains are carried by improvements in productivity that raise wages. Total factor productivity increases so that the distance from the efficient productivity (i.e., $(\text{TFP}^\star - \text{TFP})/\text{TFP}^\star$) falls from 0.11 to 0.08 (compared with a fall from 0.16 to 0.13 in our baseline model).

This rise is explained by the reallocation of capital among private businesses and between the corporate and the private sector. TFP_Q increases by 14%, while the share of capital in the corporate sector decreases from 52% to 51%. This reallocation of capital happens as total capital decreases by 2.5%. The decrease is higher in the corporate sector (3.6%) than in the private sector (1.3%). Despite the decrease in the level of capital, both corporate and private output increase, by 1% and 7% respectively, resulting in a 4.7% increase in total output.

How to Compute TFP? The aggregate production function can be written as

$$Y = Z\hat{K}^{\alpha}\hat{L}^{1-\alpha}, \tag{E.9}$$

where $\hat{K} \equiv K_c^{\rho} K_p^{1-\rho}$ and $\hat{L} \equiv L_c^{\rho} L_p^{1-\rho}$ are aggregated inputs. Total factor productivity is

$$Z \equiv \frac{Y}{\hat{K}^{\alpha}\hat{L}^{1-\alpha}} = A^{\rho}\left(\mathrm{TFP}_Q\right)^{\alpha(1-\rho)}. \tag{E.10}$$

To decompose the change in total output between two allocations, Y and Y', write in logs

$$\log\left({}^{Y'}/_{Y}\right) = \underbrace{\alpha \log\left(\frac{K'}{K}\right)}_{\text{Total Capital}} + \underbrace{(1-\alpha)\log\left(\frac{L'}{L}\right)}_{\text{Total Labor}} + \underbrace{\log\left(\frac{Z'}{Z}\right)}_{\text{Productivity}} \tag{E.11}$$

$$+\,\alpha \underbrace{\left(\rho \log\left(\frac{K_c'/K'}{K_c/K}\right) + (1-\rho)\log\left(\frac{K_p'/K'}{K_p/K}\right)\right)}_{\text{Realocation of Capital Across Sectors}}.$$

We implement this decomposition in Table E.8. Most of the change in output comes from increases in productivity, carried by reallocation within the private sector.

E.4 Extension with Public Firms

We consider another extension in which firms stochastically become "public," by which we mean they face a substantial increase in their access to credit. In this version, entrepreneurial productivity is heterogeneous but fixed (i.e., $z_{ih} = \bar{z}_i$ for all i and h, unless $\mathbb{I}_{ih} = 0$ so $z_{ih} = 0$), but each period, a firm exogenously transitions to become public with probability p_{public} and sees a jump in its collateral ratio to $\bar{\vartheta}_{\text{public}} \gg \bar{\vartheta}(z)$. Public and private firms also differ in the probability with which they exit ($\mathbb{I}_{ih} = 0$), with private firms exiting at a higher rate.

Entrepreneurial Ability and Productivity. The entrepreneurial *productivity* of individual i at age h, denoted z_{ih}, has two components: their entrepreneurial *ability*, $\bar{z}_i$, which is a fixed characteristic of the individual, and a second component that determines whether the individual's firm is active and, if so, whether it operates as a "private" or "public" firm. The ability component is transmitted imperfectly from a parent to her child just as in the benchmark model:

$$\log\left(\bar{z}_i^{\text{child}}\right) = \rho_z \log\left(\bar{z}_i^{\text{parent}}\right) + \varepsilon_{\bar{z}_i}, \tag{E.12}$$

where $\varepsilon_{\bar{z}_i} \sim \mathcal{N}\left(0, \sigma^2_{\bar{z}_i}\right)$.

TABLE E.8 – Decomposition of the Change in Output

		Contributions of				
		ΔK	ΔL	ΔTFP	Reallocation of capital	
$\log\left(\frac{Y'}{Y}\right) =$		$\alpha \log\left(\frac{K'}{K}\right)$	$(1-\alpha)\log\left(\frac{L'}{L}\right)$	$\log\left(\frac{Z'}{Z}\right)$	$\alpha\rho \log\left(\frac{K_c'/K'}{K_c/K}\right)$	$\alpha(1-\rho)\log\left(\frac{K_p'/K'}{K_p/K}\right)$
Baseline						
Tax Reform	8.84	6.08	0.7	2.06	—	—
OWT	5.94	1.04	1.94	2.97	—	—
Corp. Model						
Tax Reform	8.43	5.63	0.79	1.94	–0.14	0.21
OWT	4.60	–1.01	2.38	3.12	–0.19	0.29

Notes: The contribution of TFP is computed from the change of TFP_Q, and it corresponds to $\alpha(1-\rho)\log\left(\mathrm{TFP}_Q'/\mathrm{TFP}_Q\right)$, with $\rho = 0$ in the baseline model. There is no reallocation of capital across sectors in the baseline model, because all output is produced by the private sector.

There are three states for the firm $\mathbb{I}_{ih} \in \{\mathcal{P}r, \mathcal{P}u, 0\}$, corresponding to private, public, and inactive, respectively. Private and public firms operate with a productivity equal to the owner's entrepreneurial ability ($z_{ih} = \bar{z}_i$), while inactive firms have no productivity ($z_{ih} = 0$) and hence do not operate. The private and public status of firms is inherited across generations, capturing firms being inherited upon death of the previous owners. New firms are all private, so that if an individual with $\mathbb{I}_{ih} = 0$ dies, their offspring will operate a private firm.

High productivity private firms (those with $\bar{z}_i > \bar{z}_{\text{median}}$) have a probability p_{public} of becoming public, and active firms have a probability $p_0^{\mathcal{P}r}$ and $p_0^{\mathcal{P}u}$ of becoming inactive, which depends on their private/public status. Inactive firms remain so. The evolution of z_{ih} can be summarized by the following three-state Markov chain:

$$z_{ih} = \begin{cases} \bar{z}_i & \text{if } \mathbb{I}_{ih} = \mathcal{P}r \\ \bar{z}_i & \text{if } \mathbb{I}_{ih} = \mathcal{P}u \\ 0 & \text{if } \mathbb{I}_{ih} = 0 \end{cases} \quad \text{and} \quad \Pi_{\mathbb{I}} = \begin{bmatrix} 1 - p_{\text{public}} - p_0^{\mathcal{P}u} & p_{\text{public}} & p_0^{\mathcal{P}r} \\ 0 & 1 - p_0^{\mathcal{P}u} & p_0^{\mathcal{P}u} \\ 0 & 0 & 1 \end{bmatrix}. \tag{E.13}$$

Financial Markets. There is a bond market in which intra-period borrowing and lending take place at interest rate, r. The market works in the same way as in our benchmark. The access to the market depends on the entrepreneur's ability and the private/public status of the firm. Borrowing is collateralized and is subject to a limit indexed to individuals' assets:

$$k_{ih} \leqslant \vartheta(\bar{z}_i, \mathbb{I}_{ih}) \times a_{ih}. \tag{E.14}$$

For private firms, we keep the same properties as in the benchmark, with $\vartheta(\bar{z}_i, \mathcal{P}r) \geqslant 1$ and $\vartheta'(\bar{z}_i, \mathcal{P}r) > 0$. Public firms have more access to credit and so $\vartheta(\bar{z}_i, \mathcal{P}r) = \bar{\vartheta}_{\text{public}} \gg \max \vartheta(\bar{z}_i, \mathcal{P}r)$.

Table E.9 – Robustness: Optimal Wealth Tax

	Tax Reform		OWT	
	Baseline	Public Firms	Baseline	Public Firms
τ_a	1.13	1.52	3.03	2.76
τ_ℓ	22.4		15.4	17.6
	Change in Welfare (%)			
$\overline{CE}_1$	6.8	4.4	9.0	5.9
$\overline{CE}_2$	7.2	4.1	8.7	4.8

Notes: The table reports taxes, welfare gain for the revenue-neutral tax reform and optimal wealth tax economies for the benchmark model and the alternative model with firms with increased credit access. All numbers are in percentage points.

Parameterization. We set all parameters as in the benchmark, with the exception of the discount factor β, the consumption share in utility γ, the strength of the bequest motive χ, the dispersion of the labor fixed effect $(\sigma_{\varepsilon_\kappa})$ and of the entrepreneurial ability $(\sigma_{\varepsilon_{\bar{z}}})$, and the new parameters $\left\{p_0^{\mathcal{P}r}, p_0^{\mathcal{P}u}, p_{\text{public}}, \overline{\vartheta}_{\text{public}}\right\}$. We set these parameters to jointly match a capital-to-output ratio of 3.0, an average number of labor hours of 0.4, a bequest-to-wealth ratio of 1.2 percent, a standard deviation of log earnings of 0.8, and a top 1% wealth share of 37% as in the benchmark. We also target a share of public firms of 0.5% and a leverage ratio of 90% for public firms.[54]

The calibration implies high levels of debt, with a debt-to-output ratio of 2.43, carried by public firms that account for 88% of debt. We keep the borrowing of private firms as in the benchmark, with the lowest-ability group, $\bar{z}_0$, not being able to borrow at all $(\vartheta(\bar{z}_0, \mathcal{P}r) = 1)$, and the borrowing limit increasing linearly with ability from there on: $\vartheta(\bar{z}) = 1 + \varphi(\bar{z} - \bar{z}_0)$ with $\varphi = 0.225$. Wealth concentration is also higher than in the benchmark, with a top 0.1% wealth share of 28%.

Tax Reform and Optimal Wealth Tax. We conduct the same tax reform and optimal wealth tax experiments as we did in our benchmark. The substantive results in terms of efficiency and welfare gains from replacing capital income with wealth taxes remain unchanged; however, the size of the gains in both TFP and welfare are lower. The TFP gains are about one-half of what they are in our benchmark and welfare gains are between one-half and two-thirds, depending on the welfare measure. The general pattern across aggregates is the same as before. The level of the optimal wealth tax is lower (2.76%).

[54] We target the ratio of public firms in the US from Compustat relative to the number of firms with at least five employees from the Business Dynamics Statistics of the US Census Bureau.

E.5 Additional Robustness and Extensions

Table E.10 reports the results of nine additional robustness experiments to complement those reported in Table XII: (i) calibrating to looser constraints by targeting a debt-to-GDP ratio of 2.5, (ii) making borrowing constraints independent of productivity, $\vartheta(z) = \vartheta$; (iii) reducing the CES curvature to $\mu = 0.8$; (iv) removing life-cycle stochastic variation in productivity, $z_i = \bar{z}_i$; (v) having all individuals be born in the middle lane and transition to fast lane with probability $p_3 = 3\%$; and (vi-ix) adding wealth taxes *on top of* the current tax system with revenues used for wasteful spending or rebated by reducing the labor income tax.

TABLE E.10 – Additional Robustness and Extensions: Optimal Wealth Tax

	Looser Constraints	Constant ϑ	Higher Markups	Constant Productivity	No Start in Fast Lane	Add τ_a to Benchmark			
						2% Wealth Tax		OWT Wealth Tax	
	debt/GDP = 2.5	$\vartheta(z) = \overline{\vartheta}$	$\mu = 0.8$	$z_{ih} = \overline{z}_i$	$z_{ih} = \overline{z}_i$	τ_ℓ fixed	Adjust τ_ℓ	τ_ℓ fixed	Adjust τ_ℓ
	(i)	(ii)	(iii)	(iv)	(v)	(vi)	(vii)	(viii)	(ix)
τ_a	2.34	3.66	2.45	2.16	2.8	2.00		3.03	
τ_ℓ	19.5	12.4	18.0	19.4	16.1	22.4	14.9	22.4	12.0
	Welfare Change								
$\overline{CE}_1$	4.4	11.8	8.2	6.0	8.5	–8.3	0.9	–11.9	0.3
$\overline{CE}_2$	4.2	11.2	7.6	5.5	8.2	–9.9	0.0	–14.2	–1.0
	Change in Macro Variables (%)								
K	4.0	0.8	6.5	–2.3	2.8	–20.5	–15.6	–28.7	–22.6
Q	6.6	12.0	12.4	17.0	11.8	–17.5	–13.0	–24.7	–19.1
Y	3.6	7.3	6.1	8.0	6.5	–6.5	–4.0	–9.5	–6.2
L	1.6	4.3	2.0	2.4	3.2	1.6	2.5	2.3	3.5
C	4.1	10.1	7.3	11.1	8.5	–9.6	–2.5	–13.6	–4.2
TFP_Q	2.5	11.1	5.6	19.7	8.8	3.8	3.0	5.6	4.6
$\overline{w}$	2.0	2.9	3.9	5.5	3.3	–8.0	–6.4	–11.5	–9.4
$\overline{w}$ (net)	5.7	16.1	9.9	9.6	11.6	–8.0	2.7	–11.5	2.8
	Benchmark Economy's Debt and Productivity ($\tau_k = 0.25$, $\tau_a = 0.0$, $\tau_\ell = 0.224$)								
debt/GDP	2.5	1.5	1.5	1.5	1.5	1.5			
$\frac{TFP^\star - TFP}{TFP^\star}$	0.05	0.21	0.14	0.21	—	0.16			

Notes: The nine additional robustness experiments are as follows: (i) calibrating to looser constraints by targeting debt/GDP ratio of 2.5, (ii) making borrowing constraints independent of productivity, $\vartheta(z) = \vartheta$; (iii) reducing the CES curvature to $\mu = 0.8$; (iv) removing lifecycle stochastic variation in productivity, $z_i = \overline{z}_i$; (v) having all individuals born in the middle lane and transition to fast lane with probability $p_3 = 3\%$; (vi-ix) adding wealth taxes on top of the current tax system with revenues used for wasteful spending or rebated by reducing the labor income tax.

Table E.11 – Robustness Additional Results: Optimal Wealth Tax

	Baseline OWT	Credit Spread 10.1%	Credit Spread 6%	Public Firms	Corporate Sector	Pure Rents Model	Non-linear OKIT $\tilde{\tau}_k(y) = y - \psi y^{\eta}$	
		(1)	(2)	(3)	(4)	(5)	(6)	(7)
τ_a	3.03	2.33	2.46	2.76	3.25	1.40	—	—
τ_ℓ	15.4	13.6	15.5	17.6	16.3	27.0	22.4 (fixed)	32.3
τ_k	—	—	—	—	—	—	$(0.73, 1.022)$ (ψ,η)	$(1.20, 0.992)$ (ψ,η)
				Change in Welfare (%)				
$\overline{CE}_1$	9.0	6.1	4.3	5.9	5.8	–1.7	0.9	4.2
$\overline{CE}_2$	8.7	5.6	3.5	4.8	5.5	–1.4	0.8	5.4
				Change in Macro Variables (%)				
K	2.6	4.4	–0.8	1.5	–2.1	–2.5	5.4	41.3
Q	10.5	2.9	–1.1	4.8	11.0	–2.4	6.3	49.2
Y	6.1	3.0	0.9	3.3	2.5	–2.4	3.1	16.2
L	1.2	3.0	2.4	2.2	2.5	–2.4	1.0	–1.6
C	9.5	5.3	3.0	4.3	3.5	–3.0	3.1	14.3
TFP_Q	7.7	–1.5	–0.3	3.2	7.7	0.1	0.8	5.6
$\overline{w}$	2.8	–0.1	–1.4	1.0	0.0	–0.0	2.1	18.1
$\overline{w}$ (net)	12.0	11.3	7.3	7.3	7.9	–5.9	2.1	3.0
		Benchmark Economy's Debt and Productivity ($\tau_k = 0.25$, $\tau_a = 0.0$, $\tau_\ell = 0.224$)						
debt/GDP	1.5	1.5	2.0	2.4	0.76	1.1	1.5	
$\frac{\text{TFP}^{\star}-\text{TFP}}{\text{TFP}^{\star}}$	0.16	0.077	0.027	0.14	0.065	0.004	0.16	

Notes: The seven robustness experiments are as follows: (1) replacing collateral constraints with unlimited borrowing, subject to a credit spread of 10% generating a debt-to-GDP ratio of 1.5; (2) same as (1) but with a spread of 6%; (3) allowing firms to stochastically transition to relaxed collateral constraints; (4) introducing a corporate sector with Cobb-Douglas production and no borrowing limits; (5) eliminating z_i heterogeneity to focus on pure monopolistic rents; (6) tax reform that replaces τ_k with a nonlinear capital income tax; and (7) optimal nonlinear capital income tax experiment (choosing ψ, η, τ_ℓ).

F Endogenous Entrepreneurial Hours

In the baseline formulation, entrepreneurs' labor supply does not enter their production function. This was a deliberate choice to avoid introducing another (potentially interesting) channel through which wealth and capital income taxes can operate, which would add another layer to the analysis. Leaving a full analysis to future research, we show in this section how a plausible extension that introduces labor supply would interact with wealth taxes. The main result is that the labor supply of entrepreneurs would rise under wealth taxes, relative to the supply under capital income taxes, as long as their initial labor hours are not too high, and vice versa when they are. We give a sketch of this result here and provide more details and derivations in the following subsection.

F.1 Overview of Result

The main new channel results from a standard income versus substitution effect. To see this, consider the modified production function, $x = z(k\ell)^{\mu}$, replacing (8), so the entrepreneurs' problem (16) becomes

$$\max_{\ell, k \leqslant \vartheta(z) a} \left((1-\tau_a) a + [R(zk\ell)^{\mu} - (r+\delta) k + ra](1-\tau) - a'\right)^{\gamma} (1-\ell)^{1-\gamma},$$

where $\tau \in \{\tau_a, \tau_k\}$ and $\tau_a = 0$ if $\tau_k > 0$. The first order condition for hours is given as

$$(1-\tau)\mu R(zk)^{\mu} \ell^{\mu-1}(1-\ell) = \frac{1-\gamma}{\gamma}\left((1-\tau_a)a + [R(zk\ell)^{\mu} - (r+\delta)k + ra](1-\tau) - a'\right).$$

The left-hand side corresponds to the marginal benefit of extra work, which is the marginal utility of consuming extra output. The marginal utility depends on leisure, since consumption and leisure are complements in the utility function. So, when ℓ is high, that is, when leisure is low, the marginal benefit (MB) of extra work is lower. Switching to a wealth tax increases MB, because τ_a is a much smaller tax than τ_k on output. But if ℓ is high, the increase in MB will be small. Now, consider the marginal cost (MC): it is the utility loss due to extra work, which is proportional to consumption due to complementarity. If a switch to a wealth tax reduces consumption, it is obvious that ℓ increases. But if the wealth tax raises her consumption, what happens to ℓ depends on how much MB increases relative to MC. We can show that for our benchmark parameterization, a sufficient condition for hours to increase is $\ell \leqslant 0.43$ for the capital-constrained entrepreneur and $\ell \leqslant 0.88$ for the unconstrained entrepreneur.

To see this, consider the problem of an entrepreneur who chooses hours of work ℓ in her own firm and capital:

$$\max_{\ell, k \leqslant \vartheta(z) a} \left((1-\tau_a) a + [R(zk\ell)^{\mu} - (r+\delta) k + ra](1-\tau) - a'\right)^{\gamma} (1-\ell)^{1-\gamma},$$

where $\tau \in \{\tau_a, \tau_k\}$ and $\tau_a = 0$ if $\tau_k > 0$. The first-order condition with respect to ℓ gives

$$\frac{dC}{d\ell} C^{\gamma-1}(1-\ell)^{1-\gamma} = \left(\frac{1-\gamma}{\gamma}\right) C^{\gamma}(1-\ell)^{-\gamma}.$$

The left-hand side is the marginal benefit, and the right-hand side is the marginal cost of extra hours of work in one's firm. Simplifying this expression and substituting consumption gives

$$(1-\tau)\mu R(zk)^{\mu} \ell^{\mu-1}(1-\ell) = \tfrac{1-\gamma}{\gamma}\left((1-\tau_a)a + [R(zk\ell)^{\mu} - (r+\delta)k + ra](1-\tau) - a'\right).$$

F.2 Details and Derivations

A. Capital-Constrained Entrepreneur $(k = \vartheta(z) a)$

In this case, $k = \vartheta(z) a$ is fixed, and the first order condition is given by the following:

$$(1-\tau)\,\mu R\,(z\vartheta(z)\,a\ell)^{\mu}\,\tfrac{1-\ell}{\ell} =$$
$$\tfrac{1-\gamma}{\gamma}\,((1-\tau_a)\,a + [R\,(z\vartheta(z)\,a\ell)^{\mu} - (r+\delta)\,\vartheta(z)\,a + ra]\,(1-\tau) - a')\,.$$

The left-hand side decreases with ℓ and the right-hand side increases with ℓ; thus, there is a unique solution. Consider what happens to the left-hand side and right-hand side for a given ℓ if we switch from a capital income tax to a wealth tax:

$$\Delta LHS = (\tau_k - \tau_a)\,R\,(z\vartheta(z)\,a\ell)^{\mu}\,\mu\frac{1-\ell}{\ell}$$
$$\Delta RHS = \frac{1-\gamma}{\gamma}\left(-\tau_a a + (\tau_k - \tau_a)\,[R\,(z\vartheta(z)\,a\ell)^{\mu} - (r+\delta)\,\vartheta(z)\,a + ra] - \Delta a'\right).$$

If $\Delta LHS > \Delta RHS$, then ℓ would increase. To see the conditions under which this would happen, note that the same term $(\tau_k - \tau_a)\,R\,(z\vartheta(z)\,a\ell)^{\mu}$ appears on both sides. However, there are some additional negative terms on the right-hand side:

1. $-(r+\delta)\,\vartheta(z)\,a + ra < 0$,
2. $-\Delta a' < 0$ if $\Delta C > 0$ (the case where $\Delta C < 0$ obviously gives an increase in ℓ), and
3. $-\tau_a a < 0$.

So, definitely $(\tau_k - \tau_a)\,R\,(z\vartheta(z)\,a\ell)^{\mu} > \Delta C$. Thus, if $\mu\frac{1-\ell}{\ell} > \frac{1-\gamma}{\gamma}$, we definitely know that $\Delta LHS > \Delta RHS$. Using our benchmark parameterization $\mu = 0.9$ and $\gamma = 0.46$, we have

$$\begin{aligned}\frac{1-\ell}{\ell} &\geqslant 1.3\\ \frac{1}{\ell} &\geqslant 2.3\\ \ell &\leqslant 0.43.\end{aligned}$$

Of course, this is a sufficient condition. So, if the entrepreneur were not working too much initially (i.e. $\ell \leqslant 0.43$), then switching to a wealth tax would increase her entrepreneurial hours. Otherwise, the income effect would be greater than the substitution effect, and she would reduce her entrepreneurial hours. If we used $\mu = 0.45$ and $\gamma = 0.46$ instead, the entrepreneurial hours would increase if

$$\ell \leqslant 0.28.$$

B. Capital-Unconstrained Entrepreneur

When the entrepreneur is not capital constrained, we have the same first-order condition for labor supply:

$$(1-\tau)\,\mu R\,(zk)^{\mu}\,\ell^{\mu-1}\,(1-\ell) = \tfrac{1-\gamma}{\gamma}\,((1-\tau_a)\,a + [R\,(zk\ell)^{\mu} - (r+\delta)\,k + ra]\,(1-\tau) - a')\,.$$

The first-order condition for k is given as

$$\begin{aligned}\mu k^{\mu-1} R (z\ell)^{\mu} &= r+\delta \\ k &= \left(\frac{\mu R (z\ell)^{\mu}}{r+\delta}\right)^{1/(1-\mu)}.\end{aligned}$$

Inserting the latter into consumption, we obtain

$$C = (1-\tau_a)\,a + \left[\left(\frac{\mu R z^{\mu}}{r+\delta}\right)^{1/(1-\mu)} \ell^{\mu/(1-\mu)} (r+\delta)\,\frac{1-\mu}{\mu} + ra\right](1-\tau) - a',$$

and inserting it into $\mu R (zk)^{\mu} \ell^{\mu-1}$ on the left-hand side of the first-order condition for labor supply gives

$$\begin{aligned}\mu R (zk)^{\mu} \ell^{\mu-1} &= \mu R z^{\mu} \ell^{\mu-1} \left(\frac{\mu R (z\ell)^{\mu}}{r+\delta}\right)^{\mu/(1-\mu)} \\ &= \left(\frac{\mu R z^{\mu}}{(r+\delta)^{\mu}}\right)^{1/(1-\mu)} \ell^{(2\mu-1)/(1-\mu)}.\end{aligned}$$

Using the expression for C and $\mu R (zk)^{\mu} \ell^{\mu-1}$, we can write the first-order condition for labor supply as

$$\begin{aligned}(1-\tau)\left(\frac{\mu R z^{\mu}}{(r+\delta)^{\mu}}\right)^{1/(1-\mu)} \ell^{(2\mu-1)/(1-\mu)} (1-\ell) &= \\ \frac{1-\gamma}{\gamma}\left((1-\tau_a)\,a + \left[\left(\frac{\mu R z^{\mu}}{(r+\delta)^{\mu}}\right)^{1/(1-\mu)} \ell^{\mu/(1-\mu)} \frac{1-\mu}{\mu} + ra\right](1-\tau) - a'\right).\end{aligned}$$

The left-hand side of this equation corresponds to the marginal benefit, and the right-hand side corresponds to the marginal cost of extra hours of work by the entrepreneur. A switch to a wealth tax increases the left-hand side (since $\tau_a << \tau_k$). At an interior ℓ, that will increase hours of work. The right-hand side might increase or decrease with such a switch. If it decreases, then optimal hours of work increase unambiguously. For example, for wealth-rich entrepreneurs with relatively modest productivity, a wealth tax might reduce their after-tax wealth and consumption, leading them to work more.[55] Consider what happens to the left-hand and the right-hand sides

[55] When $\mu < 0.5$, the left-hand side is strictly decreasing and the right-hand side is strictly increasing and strictly concave in ℓ. Thus, the increase the left-hand side increases hours of work, and the increase in right-hand side reduces hours of work. When $\mu > 0.5$, the right-hand side would be strictly increasing and convex in ℓ. The left-hand side is strictly concave and has a maximum at $\ell = \frac{2\mu-1}{\mu}$. To see this, take the derivative of the left-hand side to obtain

$$\frac{dLHS}{d\ell} = \text{a } (+) \text{ constant} \times \ell^{\left(\frac{2\mu-1}{1-\mu}\right)} \times \frac{2\mu-1-\mu\ell}{(1-\mu)\,\ell}.$$

Note that $LHS = 0$ and $RHS > 0$ for $\ell = 0$, so the net benefit (MB-MC) of extra hours of work at $\ell = 0$

for a given ℓ if we switch from a capital income tax to a wealth tax:

$$\Delta \mathsf{LHS} = (\tau_k - \tau_a)\left(\frac{\mu R z^{\mu}}{(r+\delta)^{\mu}}\right)^{1/(1-\mu)} \ell^{(2\mu-1)/(1-\mu)}(1-\ell)$$

and

$$\Delta \mathsf{RHS} = \frac{1-\gamma}{\gamma}\left(\tau_k r a - \tau_a(1+r)a + (\tau_k - \tau_a)\left(\frac{\mu R z^{\mu}}{(r+\delta)^{\mu}}\right)^{1/(1-\mu)} \frac{(1-\mu)\ell^{\mu/(1-\mu)}}{\mu} - \Delta a'\right).$$

Note that if the $\Delta\mathsf{RHS} < 0$, the switch to a wealth tax definitely increases entrepreneurial hours. So, we will focus on the case in which $\Delta\mathsf{RHS} > 0$. In this case, $\Delta a' > 0$ because of monotonicity. We also know from all our experiments that a wealth tax puts a higher tax burden on the majority of the population and those who earn the market interest rate. So, we will work with the assumption that $\tau_k r a - \tau_a(1+r)a < 0$. Then, a sufficient condition for $\Delta\mathsf{LHS} > \Delta\mathsf{RHS}$ is that

$$\ell^{(2\mu-1)/(1-\mu)}(1-\ell) \quad \geqslant \quad \frac{1-\gamma}{\gamma}\frac{(1-\mu)\ell^{\mu/(1-\mu)}}{\mu},$$

which implies

$$\begin{aligned} \frac{1}{\ell} &\geqslant \frac{1-\gamma}{\gamma}\frac{1-\mu}{\mu}+1 \\ \frac{1}{\ell} &\geqslant \frac{(1-\gamma)(1-\mu)+\gamma\mu}{\gamma\mu} \\ \ell &\leqslant \frac{\gamma\mu}{(1-\gamma)(1-\mu)+\gamma\mu}. \end{aligned}$$

In our calibration, $\gamma = 0.46$ and $\mu = 0.9$, which gives $\ell < 0.88$. If we set $\mu = 0.45$, then $\ell < 0.41$.

is negative. If there is an optimal interior $\ell^* > 0$, then the left-hand side should be above the right-hand side for $\ell < \ell^*$, and the slope of the left-hand side should be smaller than the slope of the right-hand side at $\ell = \ell^*$. Thus, again the increase in the left-hand side increases hours of work, and the increase in the right-hand side reduces hours of work.