University of Toronto
Department of Economics

Working Paper 648

Use It or Lose It: Efficiency Gains from Wealth Taxation

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

September 25, 2019

Use It or Lose It: Efficiency Gains from 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 heterogeneity in rates of return across individuals, we revisit this question. With such heterogeneity, the two tax systems 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 lifetime utility to a newborn (about 7.5% in consumption-equivalent terms). Turning to optimal taxation, the optimal wealth tax (OWT) in a stationary equilibrium is positive and yields even larger welfare gains. In contrast, the optimal capital income tax (OCIT) is negative—a subsidy—and large, and it delivers lower welfare gains than the wealth tax. Furthermore, the subsidy policy increases consumption inequality, whereas the wealth tax reduces it slightly. We also consider an extension that models the transition path and find that individuals who are alive at the time of the policy change, on average, would incur large welfare losses if the new policy is OCIT but would experience large welfare gains if the new policy is an OWT. We conclude that wealth taxation has the potential to raise productivity while simultaneously reducing consumption inequality.

JEL Codes: E21, E22, E62, H21.
Keywords: Wealth taxation, Capital income tax, Rate of return heterogeneity, Power law models, Wealth inequality.

*This draft: September 6, 2019. First draft: February 15, 2017. 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, Mike Golosov, Joao Gomes, Pierre-Olivier Gourinchas, Luigi Guiso, Bob Hall, Hugo Hopenhayn, Nobu Kiyotaki, Dirk Krueger, Per Krusell, Guido Menzio, Claudio Michelacci, Ben Moll, 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. For funding, Guvenen thanks the National Science Foundation; Kuruscu thanks the Marie Sklodowska-Curie programme under Horizon 2020; and Kambourov and Kuruscu thank the Social Sciences and Humanities Research Council. 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: U of Minnesota; ocamp020@umn.edu; Chen: Econ One

1 Introduction

In this paper, we ask a simple question: How does wealth taxation differ from capital income taxation?[1] To fix ideas, let a denote wealth, r denote the rate of return on wealth, and τ_k and τ_a denote the flat 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

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

In a variety of benchmark economic models, the answer to the question above is not very interesting: the two tax systems are equivalent, with $\tau_a = \frac{r\tau_k}{1+r}$. Partly because of this equivalence, the academic literature on capital taxes most often focuses on capital income taxes, with the understanding that they can be reinterpreted as wealth taxes. However, the equivalence result relies on the assumption that all individuals face the same rate of return on wealth, which we also made implicitly above by not indexing r with a subscript i. But, what happens if this assumption does not hold—that is, if rates of return vary across individuals (as the empirical evidence we review below indicates)?

To see some of the implications for capital taxation, consider the following illustrative example. Two entrepreneurs start out with the same wealth level—say, \$1,000 each—but earn different returns on their wealth, 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 because he generates positive capital income. Under wealth taxation, on the other hand, both entrepreneurs will pay the same amount of tax on wealth regardless of their productivity, which will expand the tax base, shift the tax burden toward the unproductive entrepreneur, and reduce (potential) tax distortions on the productive entrepreneur.[2] 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 the aggregate capital stock, in turn raising productivity and output. In this sense, wealth taxation has a "use-it-or-lose-it" effect that is not present in capital income taxation.

[1] We use *capital* and *wealth* interchangeably throughout the paper.

[2] Table I provides some illustrative calculations for this example.

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. The main contribution of this paper is to study these implications in a full-blown quantitative overlapping-generations model with return heterogeneity. As we elaborate in a moment, we find that the two taxes have very different—and sometimes opposite—implications.

Three more considerations motivate us to take return heterogeneity seriously for studying capital taxation. First, a growing number of recent empirical studies cast strong doubt on the assumption of homogeneous returns across households. Using administrative panel datasets that track millions of individuals over long periods of time, these studies document large and persistent differences in rates of return across individuals, even after adjusting for risk and other factors (e.g., Fagereng, Guiso, Malacrino and Pistaferri, 2016a; Bach, Calvet and Sodini, 2018; and Smith, Yagan, Zidar and Zwick, 2017).[3] These new pieces of evidence make studying the tax implications of return heterogeneity more than a theoretical curiosity.

Second, an active literature on power law models shows that rate of return heterogeneity is a powerful modeling tool that can generate key features of inequality that have proved challenging to explain through other mechanisms.[4] This is an important consideration for the purposes of this paper: because the wealth distribution is extremely concentrated, not only in the United States but also in many other countries, the bulk of the capital tax burden falls on a small fraction of wealthy households. This makes capital taxation much more about the "right tail" than taxes on consumption and labor income, which are more evenly distributed than wealth. Thus, it is important for our model not only to generate the extreme wealth concentration at the top but also to be consistent with other features that are relevant for capturing the key trade-offs that wealthy individuals face.

One such feature is the thick Pareto tail of the wealth distribution seen in many countries (Vermeulen, 2016), which is challenging to generate (even by some models of inequality that match the share of wealth held by the top 1%) but emerges naturally in

[3]Among these, Fagereng et al. (2016a) study a 20-year-long panel that covers all households in Norway and contains extensive details about their portfolios and investments during this time. They find large differences across individuals in their (risk-adjusted) rates of return averaged over 20 years. Bach et al. (2018) analyze a similar panel dataset from Sweden and conclude that the main driver of wealth inequality at the top is heterogeneity in rates of return. Finally, for the United States, Smith et al. (2017) use a unique panel dataset from the US Department of the Treasury that contains information on 10 million firms and their owners; they document persistent heterogeneity in firm profitability even after adjusting for risk and size.

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

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). Another important feature that determines the trade-offs that the wealthy face is the extent to which their wealth is dynastic (inherited) or self-made (accumulated). In the United States, a significant fraction of the very wealthy are self-made and accumulate wealth very rapidly during their lifetime. For example, about 53% of the individuals on the 2017 US Forbes 400 list were self-made billionaires, which implies a conservative lower bound of a 1,000-fold increase in their wealth over the life cycle. A calibrated model featuring return heterogeneity can generate this pattern, as we show in this paper.

Third, studying wealth taxation also has a practical motivation: it is a policy tool that has long been used by governments around the world. Until the last decade or so, many of the richest OECD countries (e.g., France, Germany, Spain, Italy, the Netherlands, and Nordic countries, among others) had wealth taxation. Although its popularity has waned in recent decades, wealth taxation is still being used in France, Spain, the Netherlands, Switzerland, and Norway.[5] In light of this reality, studying the effects of wealth taxation (and how it differs from capital income taxation) is an important step toward providing better guidance to policy makers.

For the quantitative analysis, we study an overlapping-generations model in which individuals derive utility from consumption and leisure. 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 invest in their firm over and above their own assets. 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 low relative to the wealth they hold.

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. Individuals also face idiosyncratic labor income risk, mortality risk, and various intergenerational links, although plausible variations in these details do not change the

[5] As of 2017. See OECD (2018) for a recent review of the use of wealth taxes across OECD countries.

substantive conclusions. The calibrated model is consistent with key features of the US wealth distribution mentioned above (including the Pareto right tail and the rapid wealth growth of the very wealthy), as well as the amount of borrowing by US businesses and the degree of capital misallocation, among other features.

Our analysis produces three sets of results. First, we study a revenue-neutral tax reform that replaces the current US tax system of capital income taxation with a flat-rate wealth tax, keeping taxes on labor and consumption unchanged. Comparing across stationary equilibria, we find that this reform raises average welfare significantly—equivalent to about 7%–8% of consumption (per person per year) for newborn individuals in our baseline calibration. The gains come from a combination of the more efficient allocation of capital and the higher capital levels generated by the use-it-or-lose-it mechanism inherent in wealth taxation.[6]

Second, we move to 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. We repeat the same analysis but this time have the government choose linear taxes on labor income and capital income. As before, we start by comparing outcomes across stationary equilibria (and later consider the transition analysis). In the first case, we find that 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 (from 22.5% to 14.5%), which is more distorting than the wealth tax in this environment. The bulk of the welfare gain comes from the reduced misallocation of capital (as in the tax reform), a smaller part from higher labor supply, and almost none of it from a change in the capital stock level—which remains almost unchanged—in the new stationary equilibrium. In other words, the benefits of optimal wealth taxes in this experiment do not require more capital accumulation at the aggregate level.

Turning to optimal capital income taxation (hereafter, OCIT), we find that the optimal tax rate turns out to be negative and large, about –35%, implying a large subsidy to capital income. This finding may seem surprising in light of earlier results in the literature, which found a high positive tax rate (of about +35%) using models that share many similarities with ours (e.g., Conesa, Kitao and Krueger, 2009). The main source

[6]Compared with the optimal tax analysis that follows, this tax reform experiment is useful for two main reasons. First, by keeping other policy tools fixed and not relying on a particular objective to be maximized, it allows for a clearer comparison of the effects of capital income versus wealth taxes in isolation from the rest. Second, the same simplicity also makes it appealing for practical and implementation purposes.

of difference is return heterogeneity: shutting down return heterogeneity in our model restores the high positive tax rate found in previous work. To understand why this happens, note that in Aiyagari-style models, the wealthy are simply *workers* who have 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 not especially large. In contrast, with return heterogeneity, entrepreneurs who earn high capital income (per unit of wealth) *today* are precisely the productive ones *today*, which makes taxing capital income much more distorting and the efficiency losses especially large (and makes a subsidy an especially effective policy). These opposite results show how accounting for return heterogeneity can matter not only for wealth taxation but also for capital income taxation.

The same reasons also explain why wealth taxation is less distorting: because returns are not permanent, the wealthy today include many entrepreneurs whose best years are behind them. A wealth tax shifts the burden from less wealthy but more productive entrepreneurs toward these wealthier but less productive ones because the former generate more capital income per unit of wealth (and the opposite for the latter group).

Third, the OWT economy has higher average welfare (9.6% more than the US benchmark) than the OCIT economy (6.3% more). A decomposition of the welfare gains from wealth taxation shows that the gains come from a large rise in the *level* of consumption (driven by higher after-tax wages) and a decline in the *inequality* of (i.e., the marginal utility of) consumption. Thus, optimal wealth taxes yield both first- and second-order gains. This is not the case with OCIT: although they deliver an even larger rise in output, providing capital subsidies requires higher taxes on labor income, resulting in only a small rise in *after-tax* wages. Furthermore, subsidies on capital income lead to a significant rise in inequality—not only in wealth but also, and more importantly, in consumption—resulting in distributional losses, which offset some of the gains from levels—unlike what happens under optimal wealth taxes.

We consider an extension that models the transition path to understand how the individuals who are alive at the time of the policy switch fare from this change. The answer is far from clear because the welfare gains just discussed materialize once the economy reaches a new stationary equilibrium, whereas these individuals may have to incur additional costs (e.g., lower consumption to accumulate capital) during the transition. To keep the analysis computationally feasible, we do not solve for the full optimal policy with transition. Rather, we keep one policy instrument (e.g., the labor tax rate) at its non-transition optimum, allow the government to run a budget deficit or surplus dur-

ing the transition, and choose the other policy instrument (e.g., the wealth tax rate) so that the budget will be balanced in the new stationary equilibrium, which now includes interest payments on the accumulated debt. We also consider the opposite case—fixing the wealth tax rate and choosing the labor income tax. We find that implementing the OCIT policy leads to widespread welfare losses for existing individuals, including the newborns at the time of the policy change. The same is not true for the OWT policy, which delivers significant welfare gains for both the newborns and the overall population that is still alive. For the latter group, in fact, the gains are virtually the same as what arises from a steady state comparison. We discuss these results further in Section 7.1.

We conduct various sensitivity checks and extensions, including (optimally chosen) progressive labor income taxes, optimal wealth taxes with an exemption level, estate taxation, and relaxing or eliminating borrowing constraints, among others. As would be expected, these changes affect the various magnitudes of welfare gains but they do not overturn the main substantive conclusions of our analysis. Across these different specifications, the optimal wealth tax rate ranges from 2% to slightly above 3%.

The rest of the paper is organized as follows. Section 2 elaborates on the simple static example described above. Section 3 lays out the full-blown model, and Section 4 describes the parameterization and model fit. Sections 5 and 6 present the quantitative results from the tax reform and optimal taxation, respectively. Section 7 discusses sensitivity analyses, and Section 8 concludes.

Related Literature

Although the use-it-or-lose-it feature of wealth taxes has been discussed by a few authors, we are not aware of prior academic analyses of its effects, as we do in this paper.[7] Maurice Allais was among the best-known proponents of wealth taxes who spelled out the use-it-or-lose-it rationale in his book on wealth taxation.[8] 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 mostly focused on equity considerations but also described the use-it-or-lose-it mechanism without providing a formal analysis.[9]

[7] This paper was first presented at the 2014 NBER Summer Institute and was widely presented at seminars and conferences starting in 2015. It therefore predates the current public debate on wealth taxes that rose to prominence during the 2020 presidential election campaign.

[8] Allais observed that "[a] tax on the capital stock represents a bonus to production and penalizes the inefficient owner, passive, for whom income taxes encourage inaction" (Allais, 1977, p. 501, translated).

[9] The work of Shourideh (2013) shares some broad similarities with ours. He provides a theoretical analysis of the Mirrleesian taxation problem of wealthy individuals who face a risk-return trade-off in their investment choice. He finds a progressive saving tax to be the optimal policy.

The broader literature on capital taxation is vast, and we will not attempt to review it here (but see Chari and Kehoe (1999) and Golosov, Tsyvinski and Werning (2006) for 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 may be positive and large. The two main differences between our analysis and these studies are (i) the presence of heterogeneous returns and (ii) 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.

As noted above, this paper is also related to the power law models of inequality.[10] This literature points out that the thick Pareto tail of wealth cannot be generated through precautionary savings in response to idiosyncratic income shocks (as in Aiyagari-style models). This is because the wealth distribution inherits the Pareto tail of the income distribution—as shown by Benhabib, Bisin and Luo (2017) theoretically and by Hubmer, Krusell and Smith (2017) via simulations—which is significantly thinner than the tail for wealth. Furthermore, when the idiosyncratic income process is estimated to match micro evidence on income dynamics, these models significantly understate inequality at the top—for example, they generate one-third of the wealth holdings for the top 1% and fail to generate individuals with more than $20 million in wealth, among others.[11]

The power law literature identifies various plausible mechanisms (e.g., birth and death processes, creative destruction, stochastic-beta, heterogeneity in returns, and others) that can generate a Pareto tail in steady state (Benhabib et al., 2011, 2013, 2014; Stachurski and Toda, 2019). Furthermore, as Gabaix, Lasry, Lions and Moll (2016) show, when the heterogeneity in returns is persistent, these models generate behavior that is also consistent with the *dynamics* of inequality. Our model shares similarities with Jones and Kim

[10] Here, we discuss here the most recent strand that focuses on inequality in income and wealth. Earlier important contributions include Gabaix (1999) on Zipf's law in city size distribution, Gabaix (2011) on whether idiosyncratic shocks to firms can cause aggregate fluctuations, Luttmer (2007, 2011) on the dynamics of firm growth and the Pareto tail in the firm size distribution, as well as the much earlier literature in the 1950s that these papers build upon and extend. See Gabaix (2009) and Benhabib and Bisin (2018) for detailed surveys.

[11] See De Nardi, Fella and Pardo (2016), Guvenen, Karahan, Ozkan and Song (2016), and Carroll, Slacalek, Tokuoka and White (2017).

(2018), who emphasize the creative destruction process in entrepreneurial production to explain the Pareto tail of the income distribution. Despite the rapid growth in this literature, the implications of capital taxation in these models have not been explored, and our paper fills this gap.

Finally, our paper has some useful points of contact with different strands of literature that feature (entrepreneurial) firms with heterogeneous productivity facing financial frictions, leading to misallocation of capital, lower productivity, and underdevelopment, among other results. 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 (2009) 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 becomes pro-worker. In our framework, wealth taxation plays a similar role.

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 \$1,000 of wealth at time zero. 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 all their wealth in their business and make no other decisions. To introduce taxation, suppose that there is a government that needs to finance an expenditure of $G = \$50$ through tax revenues collected at the end of the period. The example is summarized in Table I.

Suppose that the government taxes capital income at a flat rate. To raise \$50, the required tax rate is 25% on income and is paid entirely by Michael, since he is the only one earning any capital income. The after-tax rate of return is 0% for Fredo and 15% for Michael. By the end of the period, Fredo's wealth remained unchanged, whereas Michael experienced an increase from \$1,000 to \$1,150 after paying his taxes.

Now instead suppose that the government decides to raise the same revenue with a wealth tax. The tax base is now broader because Fredo does have wealth and cannot

TABLE I – Summary of the Illustrative Example

	Capital Income Tax		Wealth Tax	
	$r_F = 0\%$	$r_M = 20\%$	$r_F = 0\%$	$r_M = 20\%$
Wealth	\$1,000	\$1,000	\$1,000	\$1,000
Pre-tax income	\$0	\$200	\$0	\$200
Tax rate	$\tau_k = \frac{\$50}{\$200} = 0.25$		$\tau_a = \frac{\$50}{\$2{,}200} = 2.27\%$	
Tax liability	\$0	\$50	$\$1,000 \times \tau_a \approx \$23$	$\$1,200 \times \tau_a \approx \$27$
After-tax rate of return	0%	$\frac{\$200-\$50}{\$1{,}000} = 15\%$	$-\frac{\$23}{\$1{,}000} = -2.3\%$	$\frac{\$200-\$27}{\$1{,}000} = 17.3\%$
After-tax wealth ratio	$\frac{W_M}{W_F} = \frac{\$1{,}150}{\$1{,}000} = 1.15$		$\frac{W_M}{W_F} = \frac{\$1{,}173}{\$977} = 1.20$	

Notes: The subscripts F and M refer to Fredo and Michael's variables, respectively. See the text for further details.

avoid taxation, as he did under the capital income tax system. Specifically, the tax base covers the entire wealth stock, or \$2,200, at the end of the period. The tax rate on wealth is $\$50/\$2,200 \approx 2.27\%$. More importantly, Fredo's tax bill is now \$23, up from zero, whereas Michael's tax bill is cut by almost half, going down to \$27 from \$50 before. The after-tax rate of return is, respectively, $(\$0 - \$23)/\$1000 \approx -2.3\%$ for Fredo and $(\$200 - \$27)/\$1,000 \approx 17.3\%$ for Michael. Notice that the dispersion in after-tax returns is higher under wealth taxes as is the end-of-period wealth inequality: $\$1,173/\$977 \approx 1.20$ versus $\$1,150/\$1,000 = 1.15$ before. Most crucially, the more productive entrepreneur (Michael) ends up with a larger fraction of aggregate wealth: 54.6% versus 53.5% under capital income taxes.

Notice that the reallocation process also increases wealth concentration. However, as we will see in the quantitative analysis, distributional losses are not a robust feature of wealth taxes and are mitigated or reversed (into gains) when a proper production function is introduced and wage income is added to the model. In that case, wealth taxes can yield both efficiency *and* distributional gains.

To sum up, wealth taxation has two main effects that are the opposite of capital income taxes. *First*, by shifting some of the tax burden to the less productive entrepreneur, it allows the more productive one to keep more of his wealth, thereby reallocating the aggregate capital stock toward the more productive individual. *Second*, wealth taxes do not compress the after-tax return distribution nearly as much as capital income taxes do, which effectively punish the successful entrepreneur and reward the inefficient one. In a

(more realistic) dynamic setting, this feature will deliver an endogenous savings response, further increasing the reallocation of capital toward the more productive entrepreneur. Going forward, we will refer to the first, direct, effect as the "use-it-or-lose-it channel" and to the second, endogenous, one as the "behavioral response channel." In addition, when the dynamic model is embedded in general equilibrium, the equilibrium response of prices to tax policies will constitute a *third* and important effect, which we will refer to as the "price channel."

A final remark is in order. If this one-period example were repeated for many periods, all aggregate wealth would eventually be owned by the more productive entrepreneur (Michael), thereby eliminating the initial misallocation of capital in the long run. While this is correct, it depends critically on the return gap between entrepreneurs being permanent, which is not a realistic assumption: the fortunes of entrepreneurs vary both over time and from one generation to the next (as entrepreneurial ability is at best imperfectly inherited), so capital misallocation will persist even in the long run. These features will be incorporated into the dynamic model we describe next.

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. Let ϕ_h be the unconditional probability of survival up to age h, and let $s_h \equiv \phi_h/\phi_{h-1}$ be the conditional probability of surviving from age $h-1$ to h. When an individual dies, she is replaced by an offspring that inherits her wealth. Individuals derive utility from consumption, c, and leisure, $1-\ell$ (where ℓ denotes market hours), and maximize expected lifetime utility without any bequest motives: $\mathbb{E}_0\left(\sum_{h=1}^{H}\beta^{h-1}\phi_h u(c_h, 1-\ell_h)\right)$, where β is the time discount factor.

Individuals make four decisions every period: (i) leisure time versus labor supply to the market (until retirement age, $R < H$), (ii) consumption versus savings, (iii) portfolio choice: how much of her own wealth to invest in her business versus how much to lend to others in the bond market, and (iv) 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 of the four decisions 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*, $\overline{z}_i$, 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.

A newborn inherits her entrepreneurial ability imperfectly from her parent:

$$\log(\overline{z}_i^{\text{child}}) = \rho_z \log(\overline{z}_i^{\text{parent}}) + \varepsilon_{\overline{z}_i}, \tag{1}$$

where $\varepsilon_{\overline{z}_i} \sim \mathcal{N}(0, \sigma^2_{\overline{z}_i})$. Because of this imperfect transmission, some low-ability children will inherit large fortunes from their high-ability parents, and vice versa, 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, we assume that high-ability entrepreneurs (specifically, those with $\overline{z}_i > \overline{z}_{\text{median}} = 1$) start life in the *fast lane*, with positive factors amplifying their productivity above their base level, $z_{ih} = \overline{z}_i^{\lambda}$ with $\lambda > 1$. 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: $z_{ih} = \overline{z}_i$, with annual probability p_1. With another probability p_2, all entrepreneurs (regardless of $\overline{z}_i$) face the risk of losing their entrepreneurial productivity completely, $z_{ih} = 0$, and they "retire" from entrepreneurial work. 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:

$$z_{ih} = \begin{cases} \overline{z}_i^{\lambda} & \text{if } \mathbb{I}_{ih} = \mathcal{H} \\ \overline{z}_i & \text{if } \mathbb{I}_{ih} = \mathcal{L} \\ 0 & \text{if } \mathbb{I}_{ih} = 0 \end{cases} \qquad \text{and} \qquad \Pi_{\mathbb{I}} = \begin{bmatrix} 1-p_1-p_2 & p_1 & p_2 \\ 0 & 1-p_2 & p_2 \\ 0 & 0 & 1 \end{bmatrix} \tag{2}$$

denoting the transition matrix, with $\mathbb{I}_{i0} = \mathcal{H}$ if $\overline{z}_i > \overline{z}_{\text{median}}$ and $\mathbb{I}_{i0} = \mathcal{L}$ otherwise.[12]

[12]We prefer this parsimonious specification with only two parameters to more general transition

Modeling this stochastic variation in productivity serves three purposes. First, and most importantly, it allows 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 source of capital misallocation (in addition to the intergenerational channel in eq. (1)). Third, it provides an additional precautionary savings motive for individuals. The main substantive conclusions continue to hold without the stochastic part (i.e., $z_{ih} = \overline{z}_i$), but the model falls short in matching some key wealth moments, as we discuss in Section 7.2.

Labor Market Productivity

The specification of labor market productivity, w_{ih}, is fairly standard:

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

where κ_i is an individual fixed effect, $g(h)$ is a polynomial in age that captures the life cycle component that is common to all individuals, and e_{ih} follows an AR(1) process during the working years ($h < R$):

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

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

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

where $\varepsilon_\kappa \sim \mathcal{N}(0, \sigma_{\varepsilon_\kappa}^2)$. Individuals supply their labor services directly to the final good producer. The aggregate effective labor supply is given by

$$L = \int (w_{ih}\ell_{ih})\, didh, \tag{3}$$

where $w_{ih}\ell_{ih}$ is a worker's efficiency-adjusted labor hours. Therefore, for a given market wage rate per efficiency units of labor, $\overline{w}$, an individual's labor income is $y_{ih} = \overline{w} w_{ih}\ell_{ih}$.

matrices one could write, especially given the difficulty of pinning down these parameters from available micro data.

3.2 Production Technology

Intermediate Goods Producers

There is a continuum of intermediate goods, each produced by a different individual according to a linear technology,

$$x_{ih} = z_{ih} k_{ih}, \tag{4}$$

where k_{ih} is the final good (consumption/capital) used in production by entrepreneur i, and z_{ih} is her stochastic and idiosyncratic entrepreneurial productivity at age h.[13]

The Final Good Producer

The final good, Y, is produced according to a Cobb-Douglas technology,

$$Y = Q^{\alpha} L^{1-\alpha}, \tag{5}$$

where L is the aggregate labor input defined in (3), and Q is the CES composite of intermediate inputs, x_i:

$$Q = \left(\int x_{ih}^{\mu} didh \right)^{1/\mu}. \tag{6}$$

To distinguish Q from the unadjusted capital stock, $K = \int k_{ih} didh$, we will refer to the former as the "quality-adjusted capital stock" since its level depends on the allocation of capital across entrepreneurs (and reflects the extent of misallocation). Each x_i is produced by a different individual in a way that will be specified in a moment. The final goods producing sector is competitive, so the profit maximization problem is

$$\max_{\{x_{ih}\}, L} \left(\int x_{ih}^{\mu} didh \right)^{\alpha/\mu} L^{1-\alpha} - \int p_{ih} x_{ih} didh - \overline{w} L,$$

where p_i is the price of the intermediate good i. The first-order optimality conditions

[13] The entrepreneur's labor supply does not appear in (4) but in (3) instead. We made this choice because the natural starting point for a first analysis of wealth taxation is to focus on entrepreneurs' saving and production decisions and abstract from other endogenous choices that introduce new channels and confound the analysis. This approach is in keeping with the bulk of the previous literature that has employed the entrepreneur- or capitalist-worker framework to study capital taxation (see, e.g., Judd (1985) and Straub and Werning (2018)). As we shall see, several mechanisms will already be operating in the current framework, and we do not want to add more at this stage. However, given that this is an interesting question, in Appendix E, we sketch an extension that introduces labor supply into (4) and discuss the conditions under which this extension amplifies or dampens the effects of taxes that we analyze in the current model.

yield the inverse demand (price) function for each intermediate input and the wage rate:

$$p\left(x_{ih}\right)=\alpha x_{ih}^{\mu-1}Q^{\alpha-\mu}L^{1-\alpha}\qquad \overline{w}=(1-\alpha)Q^{\alpha}L^{-\alpha}. \tag{7}$$

3.3 Markets and the Government

Financial Markets

There is a bond market where intraperiod borrowing and lending take place at a risk-free rate of r. This market has three important features. First, individuals can only use borrowed funds only as capital in production. In particular, they cannot use it to finance consumption. Second, borrowing and lending decisions take place before production *but after* z_{ih} is observed, so there is no uncertainty about whether an entrepreneur can repay at the end of the period. As can be anticipated, individuals with sufficiently 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:

$$k_{ih}\leq\vartheta(\overline{z}_i)\times a_{ih},$$

where $\vartheta(\overline{z}_i)\in[1,\infty]$ and is nondecreasing in z_i. When $\vartheta=1$, individuals can only use their own assets in production; when $\vartheta=\infty$, they can borrow without a limit. We assume that $\vartheta^{'}(\overline{z}_i)>0$ to capture the plausible idea that an entrepreneur's productivity may be partly observable by the market, allowing more productive entrepreneurs to borrow more.[14] In Section 7.2, we consider alternative forms of borrowing constraints, including the limitless case, $\vartheta=\infty$.[15]

The Government and Tax Systems

In the benchmark economy that aims to represent the current US tax system, the government imposes flat taxes at rates τ_k on capital income, τ_ℓ on labor income, and τ_c on

[14] Li (2016) provides evidence of looser constraints for more productive firms in a sample of young, unlisted firms in Japan. Notice that with $\vartheta(z_i)\equiv\overline{\vartheta}$ for all i, financial constraints would have more bite, by making constraints tighter for more productive entrepreneurs, which increases capital misallocation and, consequently, the positive effects of wealth taxes. So, the specification we choose here provides a more conservative benchmark.

[15] One theoretical foundation for collateral constraints is agency problems, which we do not model here. In recent work, Boar and Knowles (2018) study Ramsey taxation when entrepreneurs have an incentive to hide/misreport their output to the financial intermediary and the government. They find that capital income taxes exacerbate these agency problems as well as tax avoidance relative to wealth taxes.

consumption. We will 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 $\tau_k \equiv 0$) but imposes a flat-rate tax, τ_a, on post-production wealth. We will later consider formulations with progressive taxes. We denote a tax system as $\mathcal{T} \equiv (\tau_{\text{cap}}, \tau_\ell)$, where "cap" refers to the tax rate on capital, which can be on either capital income or wealth depending on the context.[16] The government runs a balanced budget every period 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.

A Remark about the Timing of Taxes. The timing choice for the wealth tax—on post-production wealth—was made for keeping the timing consistent with the capital income tax economy. Yet it raises a question about the interpretation of τ_a as a wealth tax because post-production wealth includes current-period entrepreneurial profits (eq. 12 below) in addition to beginning-of-period wealth, a. To the extent that z_{ih} is viewed as partly capturing the effort or time that entrepreneurs spend running their business, entrepreneurial profits can be viewed as including a return on their labor, which suggests that τ_a may be better interpreted as a combined tax on the entrepreneur's wealth *and* labor input. Although, strictly speaking, this interpretation would be correct, it arises from this particular timing choice, which is not an essential element in the model. An alternative formulation that makes this point clear is a timing where taxes are levied on beginning-of-period wealth, a—which clearly does not include entrepreneurial income. As we show in Section 7.2, this formulation delivers the same substantive conclusions as the baseline model, and if anything, it strengthens them. Therefore, this timing choice is not a critical feature of the model, and interpreting τ_a as a wealth tax is justified.

Social Security Pension System

Individuals retire at age R and start receiving pension income, y^R, which is determined according to the following formula:

$$y^R(\kappa, e_{R-1}) = \Phi(\kappa, e_{R-1})\,\overline{y}, \tag{8}$$

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 transitory component of her labor productivity at age $R-1$. The functional form of Φ is taken from the US Social Security Administration's OASDI system, as described in Appendix A.1.

[16] Consumption taxes are not included in $\mathcal{T}$ because they will be fixed throughout the analysis.

3.4 Individuals: Two Decision Problems

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 decision. The first problem is static and does not interact with the second one in a given period, so it can be solved separately from the latter. The only dependence is through the appearance of a—which is predetermined—in the entrepreneur's borrowing constraint.

I. Individual/Entrepreneur's Problem

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

$$\pi\left(a,z\right)=\max_{k\leq\vartheta(z)a}\left\{p\left(zk\right)\times zk-\left(r+\delta\right)k\right\}, \tag{9}$$

where δ is the depreciation rate. The price of the differentiated good in (7) can be written as: $p\left(zk\right)=\mathcal{R}\times\left(zk\right)^{\mu-1}$, where $\mathcal{R}\equiv\alpha Q^{\alpha-\mu}L^{1-\alpha}$, yielding the solution

$$k\left(a,z\right)=\min\left\{\left(\frac{\mu\mathcal{R}z^{\mu}}{r+\delta}\right)^{\frac{1}{1-\mu}},\vartheta(z)a\right\}, \tag{10}$$

with the associated maximized profit function

$$\pi\left(a,z\right)\equiv\begin{cases}\mathcal{R}\left(z\vartheta(z)a\right)^{\mu}-\left(r+\delta\right)\vartheta(z)a & \text{if } k\left(a,z\right)=\vartheta(z)a\\ \left(1-\mu\right)\mathcal{R}z^{\mu}\left(\frac{\mu\mathcal{R}z^{\mu}}{r+\delta}\right)^{\frac{\mu}{1-\mu}} & \text{if } k\left(a,z\right)<\vartheta(z)a\end{cases}. \tag{11}$$

II. Individual's Dynamic Programming Problem

Let ω denote the post-production, after-tax wealth of the individual, which includes current-period profits plus interest income from the bond market:

$$\omega(a,z;\tau_{\text{cap}})\equiv\begin{cases}a+\left(\pi\left(a,z\right)+ra\right)\left(1-\tau_{k}\right) & \text{if } \tau_{\text{cap}}=\tau_{k}\\ \left[a+\left(\pi\left(a,z\right)+ra\right)\right]\left(1-\tau_{a}\right), & \text{if } \tau_{\text{cap}}=\tau_{a}\end{cases}, \tag{12}$$

where τ_{cap} is used used as generic notation for a tax on capital, which can be on wealth or capital income. It will also be convenient to define the *total disposable resources* of

the individual (after production and taxation):

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

which the individual optimally splits between c and a'. Finally, let $\mathbf{S} \equiv (\overline{z}, \mathbb{I}, e, \kappa)$ denote the vector of exogenous individual states. The individual's dynamic problem is given by

$$\begin{aligned} V_h(a; \mathbf{S}) = \max_{c,\ell,a'} & \, u\left(c, 1 - \ell\right) + \beta s_{h+1} E\left[V_{h+1}(a', \mathbf{S}') \mid \mathbf{S}\right] \\ \text{s.t.} \quad & \left(1 + \tau_c\right) c + a' = \mathcal{Y}(a, \ell; z, e, \kappa; \mathcal{T}) \\ & a' \geq 0. \end{aligned} \tag{13}$$

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

4 Quantitative Analysis

4.1 Model Parameterization

The benchmark model is calibrated to the US data. The model period is one year. Individuals enter the economy at age 20, retire from the labor market 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) for the US data, and individuals die by age 100 (81 periods) with certainty.

Preferences. In the baseline analysis, we consider a Cobb-Douglas utility function:

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

We set $\sigma = 4$ following Conesa et al. (2009) and choose γ and β to generate 40 hours of market work per week for the working-age population (i.e., $\ell = 0.4$, assuming 100 hours of discretionary time per week) and a K/Y ratio of 3, which requires $\gamma = 0.46$ and $\beta = 0.9475$.

Tax System. The current US tax system is modeled as a triplet of tax rates: $(\tau_k, \tau_\ell, \tau_c)$. We set $\tau_k = 25\%$, $\tau_\ell = 22.4\%$, and $\tau_c = 7.5\%$ based on McDaniel (2007), who measures these tax rates for the US economy.

Labor market productivity. The deterministic life cycle profile, $g(h)$, is a quadratic polynomial that generates a 50% rise in average labor income from age 21 to age 51.[17] The AR(1) process has a persistence of $\rho_e = 0.9$ and an innovation standard deviation of $\sigma_e = 0.2$, broadly consistent with the estimates in the literature.[18] The intergenerational correlation of the fixed effect is set to $\rho_\kappa = 0.5$ (see Solon (1999)). With these parameters fixed, we set $\sigma_{\varepsilon_\kappa} = 0.305$ to match a cross-sectional standard deviation of log labor earnings of 0.80 (Guvenen, Karahan, Ozkan and Song, 2015).

Entrepreneurial productivity. As discussed in the introduction, empirical evidence on individual rates of return (which would provide natural moments to pin down ρ_z and $\sigma_{\varepsilon_{\bar{z}}}$) has become available only very recently. The most relevant evidence for our parameterization comes from Fagereng et al. (2018), who document the dispersion and intergenerational correlation of individual-level average rates of return for Norwegian households. Although the lack of US evidence is not ideal, Fagereng et al.'s evidence is noteworthy for two reasons. First, it is based on a 20-year-long population panel, with detailed information on investments (including private businesses) and debts, which is hard to come by and enables them to measure the *persistent* component of return heterogeneity. Second, Fagereng et al. show that, despite the much lower income inequality in Norway, wealth inequality is actually very high, with the Gini coefficient and top wealth shares very similar to those for the United States. This suggests that Norwegian evidence on returns may be more relevant than it first appears. With this caveat noted, we set $\rho_z = 0.1$ to match the intergenerational correlation of individual fixed effects in rates of return that the authors report.[19] As for $\sigma_{\varepsilon_{\bar{z}}}$, we choose it—together with the parameters of the stochastic component—to match the wealth share of the top 1%.

The stochastic component, z_{ih}, has three parameters: λ, p_1, and p_2. The amplification in returns generated by λ is critical for reproducing the very fast speed at which many of the super-wealthy individuals reach that status over the life cycle. We match the fraction of billionaires in the Forbes 400 list who are self-made, defined as a billionaire who came from at most an upper-middle-class family.[20] The model counterpart is defined as a billionaire whose current wealth is at least 100 times (alternatively, 1,000 times) larger than her initial wealth. We set $\lambda = 5$, $p_1 = 0.05$, and $p_2 = 0.03$, which generates a self-

[17] $g(h) = -(h-1)^2/1800 + (h-1)/30$

[18] The choice of ρ_e is on the lower side of estimates to account for the fact that we do not model transitory shocks separately.

[19] Because this value seems quite low, we conducted a sensitivity check using $\rho_z = 0.5$ instead but did not find any substantive differences. Results are available upon request.

[20] Details of the Forbes classification of billionaires are in Appendix B, Table B.1.

TABLE II – Benchmark Parameters

Parameters Calibrated Outside of the Model		
Parameter		Value
Capital income tax rate	τ_k	25.0%
Labor income tax rate	τ_ℓ	22.4%
Consumption tax rate	τ_c	7.5%
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
Interg. correlation of entrepreneurial ability	$\rho_{\overline{z}}$	0.1
Intermediate goods aggregate share in production	α	0.4
Curvature parameter of CES production func.	μ	0.9
Depreciation rate	δ	0.05
Curvature of utility function	σ	4.0
Annual transition rate into low entrepr. product. state	p_1	0.05
Annual transition rate into "retirement" from entrepr.	p_2	0.03
Maximum age	H	81
Retirement age	R	45
Parameters Calibrated (Jointly) Inside the Model		
Discount factor	β	0.9475
Consumption share in utility	γ	0.460
Std. dev. of interg. transm. of labor fixed effect	$\sigma_{\varepsilon_\kappa}$	0.305
Std. dev. of interg. transm. of entrepr. ability	$\sigma_{\varepsilon_{\overline{z}}}$	0.072
Productivity boost while in fast lane	λ	5.0

Note: The parameters $\sigma_{\varepsilon_{\overline{z}}}$ and λ are jointly chosen to match the top 1% share and fraction of self-made billionaires. In addition, survival probabilities, ϕ_h, are taken from Bell and Miller (2002) (omitted from the table).

made fraction of billionaires of 50% by the first definition (and 47.5% by the second), compared with 54% in Forbes. For individuals with an above-median value of $\overline{z}_i$, it also implies a mean duration of 12.5 years in the fast lane ($\mathbb{I} = \mathcal{H}$) and an additional 17.9 years as a normal entrepreneur ($\mathbb{I} = \mathcal{L}$).

Production. We target 0.60 as the labor share of output by setting $\alpha = 0.4$. The curvature parameter of the CES aggregator, μ, is set to 0.9, which corresponds to a 10%

TABLE III – Key Moments: Model versus US Data

	US Data	Benchmark
Targeted Moments		
Capital-to-output ratio	3.00	3.00
Std. dev. of log earnings	0.80	0.80
Average labor hours	0.40	0.40
Top 1% wealth share	0.36	0.36
Self-made billionaires (fraction)	54%	50%
Selected Untargeted Moments		
Total tax revenue/GDP	0.295	0.25
Revenue from capital taxes/Total tax revenue	0.280	0.25
Bequest/Wealth	~1%	0.99%
Business Debt/GDP	1.29	1.27

Note: The five parameters chosen to match these moments are: $\beta, \gamma, \sigma_{\varepsilon_\kappa}, \sigma_{\varepsilon_{\bar{z}}}$, and λ. The last two parameters are especially important for matching the top 1% share and the fraction of self-made billionaires in the data.

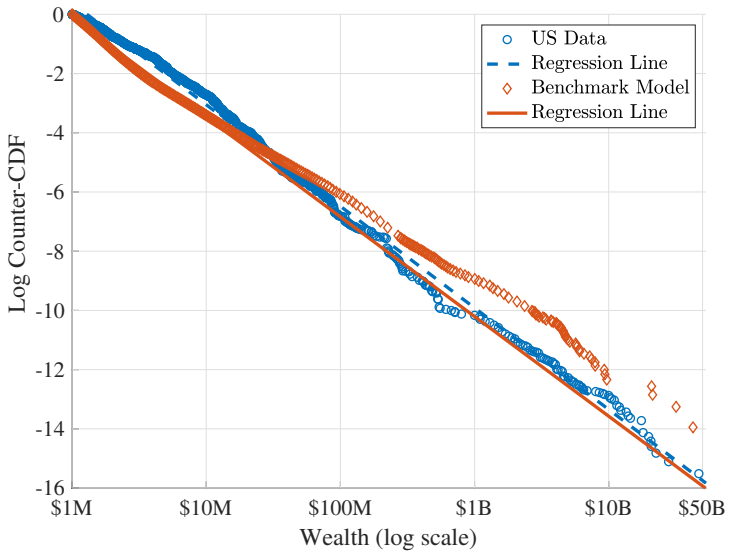
markup over marginal cost. A higher μ brings the model closer to an AK framework, which makes it easier to generate high inequality, a Pareto tail, and very high benefits from improving the tax system. The reverse happens as μ goes down, but the effects are mild down to a value of $\mu = 0.75$ or so. 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 report the results from $\mu = 0.8$ in Section 7.2. The depreciation rate of capital is set to 5%.

Financial constraints. For reasons explained before, we allow individuals with higher entrepreneurial ability to borrow more. In particular, we discretize the $\overline{z}_i$ distribution into 9 equally spaced mass points between $-3\sigma_{\overline{z}}$ and $+5\sigma_{\overline{z}}$ and assign individuals to the nearest group.[21] Then we set $\vartheta(\overline{z}_i) = 1 + 1.5(i-1)/8$ for $i = 1, ..., 9$. With this choice, the model generates the same ratio of corporate debt to GDP as in the US data. We experiment with looser constraints later.

Table II summarizes the parameters that we calibrate independently (top panel) and those that are calibrated jointly in equilibrium (bottom panel) to match the moments shown in the top panel of Table III.

[21] Adding more groups at the bottom end does not make a difference.

FIGURE 1 – Pareto Tail: Wealth above 1 Million US Dollars



4.2 Performance of the Benchmark Model

Before we can move on to the tax analyses, we need to discuss whether the parameterized model provides an acceptable description of the US economy along the dimensions relevant for the current analysis.

The top panel of Table III reports the five moments we targeted in the calibration, and the simulated model does a good job of matching them. The bottom panel reports various untargeted moments of interest. First, the share of aggregate tax revenues in US GDP is 29.5%, and the share of capital tax revenues in total tax revenues is 28%. The model counterparts are slightly lower but still reasonably close to the empirical values, both at 25%. Second, we compare the amount of bequests in the model with the US data. In the model, total bequests are about 1% of aggregate wealth, which is broadly in line with the estimates reported in the literature.[22]

Wealth inequality. To see whether the simulated model can reproduce the thick right tail of the wealth distribution, we plot the log counter-CDF of wealth against log wealth. If the right tail has a Pareto distribution—$P(\omega > x) = a \times x^{-\alpha}$—then this log-log plot will be linear with a slope of $-\alpha$. In Figure 1, the blue circles plot the US data for 2010

[22] 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 the 1989 Survey of Consumer Finances (SCF) data. An earlier estimate by Auerbach, Gokhale, Kotlikoff, Sabelhaus and Weil (1995) for the bequest-GDP ratio is 3.6%, which would translate to a 1.2% bequest-wealth ratio in our model (given $K/Y = 3$).

TABLE IV – Wealth Concentration in the Benchmark Model

	Share of aggregate wealth by top x%					*Gini*
x:	50	10	1	0.5	0.1	
US Data	0.99	0.75	0.36	0.27	0.14	0.82
Model	0.97	0.66	0.36	0.31	0.23	0.78

Note: Wealth shares are computed using data for the United States from Vermeulen (2016), who merges SCF and Forbes 400 data for 2010. The wealth Gini is based on the 2001 SCF, computed by Wolff (2006).

taken from Vermeulen (2016), who merges data from the Survey of Consumer Finances (SCF) with the Forbes 400 list (to get data on billionaires omitted from the SCF). The data align extremely well along a straight line from \$1M all the way to \$50B, confirming the well-known Pareto right tail of the US wealth distribution.

The model counterpart is shown with red diamond markers, which is also close to a straight line and overlaps very well with the data, especially up to about \$0.5 billion. The model implies a somewhat higher likelihood of billionaires than the data above that point, which is also reflected in the share of wealth held by the top 0.1% (23% in the model versus 14% in the US data; see Table IV). Given the challenges in generating a thick tail at the very top, we do not view this disparity as a concern.[23]

Assessing financial frictions and misallocation in the model. We evaluate the extent of financial frictions in the model along two dimensions. First, we ask how the aggregate business borrowing in the model compares with the US data. In 2015, the ratio of total nonfinancial business liability to US GDP was about 1.29, whereas the analogous ratio of business debt-to-output ratio is 1.26 in the model.[24] Alternative measures of the business debt imply even lower figures, suggesting that the amount of borrowing allowed in the parameterized model is probably on the high side; hence, if anything, our parameterization is conservative.[25]

[23] This overstatement is a feature of this particular calibration and can be fixed with small variations in parameter values and—it turns out—without much effect on the tax analysis. See Figure C.2a for the $\mu = 0.8$ case and Figure C.2b when wealth is measured in present value terms. See Section 7 for further discussion. We do not attempt to fine-tune the model to match it exactly since it seems inconsequential for the tax analysis, which is our main focus.

[24] The US figure is from the flow of funds accounts (Federal Reserve Statistical Release (2015Q3), Table L.102), which reports a total nonfinancial business liability of \$22.79 trillion and a nominal US GDP of \$17.65 trillion, yielding a ratio of 1.29.

[25] Federal Reserve Statistical Release (2015Q1, Table L2) a value of \$12.2 trillion for the "total credit market borrowing by Nonfinancial Sectors, implying a ratio of 0.68, or half the value in the model.

Second, we measure the extent of capital misallocation following the method proposed by Hsieh and Klenow (2009). The measure is based on the dispersion across firms of the marginal revenue product of capital (and labor), given that it would be equalized in a frictionless economy. Using data from the US Census of Manufacturers, Hsieh and Klenow estimate that the TFP gains from removing all distortions (wedges), which equalizes the "revenue productivity" (TFPR) within each industry, range between 31% and 43% from 1977 to 1997. Adapting their method to our context, the corresponding figure in our model is about 20%—or about half of what is reported by these authors (see Appendix D for details). In a recent paper, Bils et al. (2017) argued that measurement error in micro data can create an upward bias in the procedure and came up with a revised estimate around 20%. So, overall, the extent of capital misallocation in the model seems to be in line with what has been found in the US data.[26]

5 Tax Reform

In this section, we analyze the effects of a tax reform in which the government eliminates capital income taxes (setting $\tau_k = 0$) and levies a flat-rate wealth tax, while keeping τ_ℓ and τ_c unchanged. We compare both aggregate and distributional outcomes in the stationary equilibrium of this wealth tax economy with those in the benchmark. Compared with the optimal tax analysis we conduct in the next section, this tax reform analysis serves two important purposes. First, it is a simpler experiment in that it does not rely on the particular objective function maximized by the government, and it keeps policy tools other than the wealth tax fixed. Consequently, it allows for a clearer analysis of how wealth taxes work and how they differ from capital income taxes in isolation of other channels that would be present with those additional steps. Second, its relative simplicity also makes it appealing from a policy perspective, compared with an optimal policy that requires changes to various tax tools simultaneously.

To make the comparison across the two economies meaningful, we need to impose a neutrality condition. An obvious choice is to assume revenue neutrality; however notice

Asker, Farre-Mensa and Ljungqvist (2011) use firm-level data and 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. Given that the capital-to-output ratio is 3 in our model, these numbers correspond to an aggregate debt-to-output ratio between 0.60 and 0.93, which again is lower than in the model.

[26] Another comparison of interest is the distribution of lifetime returns in the model and the data. As explained earlier, the empirical measure is not available for the United States. Fagereng et al. (2016b) report select percentiles of the distribution of fixed effects for individual returns. Table B.2 in Appendix B.1 reports the model counterparts. The overall gap between the 10th and 99.9th percentile compares well with the data, whereas the dispersion in the middle (e.g., the interquartile range) is somewhat overstated.

one issue: because pension payments are anchored to the average income level ($\overline{y}$ in equation (8)), a tax reform that changes $\overline{y}$ will also change SSP, violating the budget balance if revenue is kept constant. To deal with this issue, we consider two cases. In the first, which is our main "revenue-neutral" (RN) tax reform, we keep the pension income of every individual *fixed* at its baseline value after the wealth tax reform. In the second case, the "balanced-budget" (BB) tax reform, we allow pension benefits to scale up or down with the level of $\overline{y}$ while choosing the level of wealth taxes to keep the government budget balanced. Except where we note explicitly, the results we present below pertain to the RN tax reform.

5.1 Results

The RN tax reform requires a wealth tax rate of $\tau_a = 1.13\%$ to generate the same revenue as the baseline economy (with $\tau_k = 25\%$). The BB reform requires a slightly higher rate, $\tau_a = 1.54\%$, mainly because of the added cost of higher pensions.

With respect to macro outcomes, a glance at the left panel of Table V shows that aggregate quantities increase across the board after the tax reform. Starting with the RN reform, K and Q are higher by 19.4% and 24.8%, respectively. (The K/Y ratio rises from 3 to 3.25.) The larger increase in Q reflects the reduction in capital misallocation induced by the wealth tax. This improvement in efficiency can be expressed as an increase in TFP of 4.6% in the intermediate goods sector.[27] Furthermore, aggregate labor hours and the market wage are higher by 1.3% and 8.7% respectively, clearly showing that the 10.1% rise in output is primarily accounted for by the higher Q and not L. Finally, the after-tax net interest rate falls by about 1 percentage point (right panel) since wealth taxes erode the principal and the rise in the before-tax interest rate is too small (26 basis points) to offset the principal loss.

The results for the BB reform are qualitatively the same as the RN reform (also shown in Table V). Quantities increase slightly less (by 15%–20%) than in the RN reform—owing to the somewhat higher tax rate to pay for rising pensions—with the exception of TFP, which rises slightly more than before.

Several other changes are worth noting. Wealth inequality is higher in both the RN and BB reform economies (as anticipated from the simple illustrative example in Section 2), with the top 1% share rising from 36% to 46% and the top 10% share rising from 66% to 72%. Inequality in labor earnings remains virtually unchanged, which is

[27] Equation (16) in Appendix D provides the exact formula, which in this case simplifies to the change in Q/K after the reform.

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

Tax Reforms: Change from Benchmark					
Quantities (% Change)			*Prices*		
	RN	BB		RN	BB
ΔK	19.4	12.3	$\%\Delta\overline{w}$	8.7	6.4
ΔQ	24.8	18.4	$\%\Delta\overline{w}$ (net)	8.7	6.4
ΔTFP	4.6	5.5	Δr	0.26	0.05
ΔL	1.3	1.4	Δr (net)	–0.91	–1.02
ΔY	10.1	7.9			
ΔC	10.0	8.4			

Note: 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 computed 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 $(1+r)\,(1-\tau_a)-1$, depending on the model. The TFP variable is measured in the intermediate goods market.

not surprising given the very small hours response to the reform. We will discuss the changes in consumption and leisure inequality later on. Finally, the bequest-wealth ratio barely moves (going from 1% to 1.07%).

5.2 Quantifying Three Mechanisms

The large rise in output in Table V is almost entirely driven by the rise in Q, which itself is due partly to the rise in K (level effect) and partly to the improved allocation of K across entrepreneurs (reallocation effect). Our goal here is to quantify and better understand the mechanisms that contribute to these two effects.

To this end, we begin by documenting the extent of capital reallocation that results from switching to wealth taxation. Table VI reports the percentage change in the fraction of entrepreneurs with different ability ($\overline{z}_i$) levels in various top wealth groups due to the RN tax reform. The first row shows that among the top 1% *wealth* holders, the number of entrepreneurs from the lower 90% of the $\overline{z}_i$ distribution is 10.8% to 13% lower under the wealth tax, whereas those from the top 10% of the $\overline{z}_i$ distribution are higher by 9.4% to 11.2%. The lower rows of the table show that the same compositional shift with declining magnitudes as we go down, thereby increasing the alignment of wealth with entrepreneurial skills and driving the efficiency (TFP) gains from wealth taxation.

Next, to understand the factors that drive this increase in reallocation and the rise in K, consider the three main differences between the benchmark US and RN reform

TABLE VI – Tax Reform from τ_k to τ_a: Change in Wealth Composition

	Entrepreneurial Ability Groups ($\overline{z}_i$ Percentiles)					
Top x%	0–40	40–80	80–90	90–99	99–99.9	99.9+
1	–12.0	–13.0	–10.8	10.5	11.2	9.4
5	–8.2	–3.3	1.6	8.3	8.9	7.9
10	–6.4	–1.3	2.9	6.4	6.9	6.2
50	–2.5	0.9	1.8	1.6	1.2	1.1

Note: The table shows the percentage change induced by the revenue-neutral tax reform of the share of individuals in each entrepreneurial ability group (ranked based on $\overline{z}_i$) among the top x% wealth holders. Each entry is computed as $100 \times \left(s_{ji}^{\text{RN}}/s_{ji}^{\text{US}} - 1\right)$, where s_{ji} is the share of entrepreneurs in top wealth group j with ability level in $\overline{z}_i$ group in the economy indicated by the superscript.

economies: (i) capital is taxed differently, (ii) equilibrium prices and hence returns are different, and (iii) savings behavior (decision rules) is different. To be more precise, we will refer to the effects of (i) as the use-it-or-lose-it mechanism; to (ii) as the GE price effect; and to (iii) as the behavioral savings response. The first effect, in turn, works through two channels: One, it shifts the tax burden toward less productive entrepreneurs and leave more profits in the hands of productive ones; and two, it changes the distribution of *after-tax* rates of return for *fixed* equilibrium prices. Of course, equilibrium prices do change—captured by (ii)—so the distribution of *before-tax* rates of return also change. The resulting change in the return distribution affects both the level and distribution of wealth, even when decision rules are fixed. Of course, decision rules also change—captured by (iii)—further affecting the level and distribution of capital. Although these three channels interact with each other and cannot be perfectly disentangled, we can gain some insights into their relative importance through a sequential decomposition.

Table VII reports a decomposition of these three effects on the rise in K with the wealth tax reform (left panel), and further decomposes each effect into the contribution of different entrepreneurial ability percentiles (middle panel): 0–90, 90–99, and the top 1% groups. The right panel shows the shifts in the wealth distribution across ability groups resulting from each effect.

Starting in the left panel, the top row (A) shows that the use-it-or-lose-it effect on its own (fixing prices and decision rules at the benchmark US economy equilibrium) would raise K by 14.5 log percent from the benchmark level. The middle panel shows that

TABLE VII – Tax Reform: Change in Macro Variables from Current US Benchmark

Decomposing the Three Channels							
	Δ *logK*	*Contrib. by* $\overline{z}_i$ *pctiles*			*% Change in wealth share*		
Due to:		0–90	90–99	99+	0–90	90–99	99+
A. Use-it-or-lose-it	14.5	3.8	4.4	6.3	–5.7	1.6	4.0
B. GE (price) effects	–13.1	–8.2	–3.2	–1.7	0.0	–0.7	0.6
C. Behavioral response	16.3	8.7	2.9	4.8	–1.4	–0.2	1.7
Total Effect (A+B+C)	17.7	4.3	4.0	9.4	–7.1	0.7	6.4

Note: Row A reports the effect of replacing the capital income tax in the benchmark US economy with the RN reform wealth tax, fixing the equilibrium prices and decision rules at the benchmark economy. Row B reports the incremental effect from row A when all equilibrium prices are also changed to their post-reform steady state values, while still keeping the decision rules at the benchmark. Row 3 reports the incremental effect from row B, when decision rules are also adjusted (behavioral savings response). The middle panel further decomposes each effect into the contributions of different entrepreneurial ability ($\overline{z}_i$) groups, which add up to the aggregate changes reported in the first column.

more than 40% of this rise (6.3 log pct) is accounted for by the extra wealth left in the hands of the top 1% $\overline{z}_i$ group, another 30% (4.4 log pct) from the next 9% group, and the remaining 25% (3.8 log pct) from the lower 90% of the population. Turning to the wealth distribution, the effect is large (right panel): the wealth share of the top group goes up by 4 percentage points (pp) and the share of bottom declines by 5.7 pp. These results show that the use-it-or-lose-it channel does not only increase the level of K but it also reallocates it (through the tax system) in a way that improves its efficiency.

Row B shows that if equilibrium prices are also adjusted (in addition to taxes), K would decline by 13.1 log pct relative to row A, offsetting most of the gain, so the overall change (A+B) from the benchmark US economy would be close to zero. The reason for the GE negative price effect can be explained as follows: the after-tax average return on both capital and the risk-free bond fall from a combination of diminishing returns (Q is higher in the RN economy, whereas average $\overline{z}$ is fixed) and taxing wealth as opposed to capital income. These lower returns, in turn, lead to slower wealth growth (with fixed decision rules). Although this affects all groups, it is more severe for the lower ability groups (middle panel) who heavily depend on the risk-free bond whose return falls by almost 1 percentage point (Table V). That said, the wealth concentration increases (right panel) but the change is more modest than in row A.

Finally, row C shows that the behavioral savings response—from changing the decision rules from row B to those under the RN wealth tax equilibrium—would raise K by 16.3 log pct relative to row B. Furthermore, the savings response is actually stronger at the bottom end, with the 0–90 group contributing slightly more than half of the rise in K. However, the wealth distribution further concentrates at the top.

Putting these three effects together (bottom row), we see that slightly more than half of the total rise in K comes from the top 1%, with the rest more or less equally contributed by the next top 9% and the bottom 90%. The right panel of the same row shows that overall, the wealth share of the top 1% $\overline{z}$ group grows by 6.4 pp nearly matched by the 7.1 pp decline in the bottom 90%. For the top 1% highest-ability group, the strongest positive effect comes from the use-it-or-lose followed by the behavioral response, with GE price effects having a smaller effect. As we will see later, because of the improvement in productivity resulting from this reallocation, the rise in wages improves welfare significantly even for those in lower productivity groups.

5.3 Welfare Analysis

We use two measures to quantify the welfare consequences of the tax reform. The first one, CE_1, is a micro measure—constructed at the individual level—that allows us to quantify the gains and losses experienced by different groups in the population. It can also be aggregated to obtain a welfare measure for the whole population. The second one, $\overline{CE}_2$, is a macro measure, basically an adaptation of the famous Lucas (1987) calculation to our framework. It provides a single figure that is easy to interpret and allows comparison with some previous work.

Specifically, for an h-year-old individual in state $\mathcal{S} \equiv (a, \mathbf{S})$, we compute CE_1 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. That is

$$V_h^{\text{US}}((1 + CE_1(h, \mathcal{S})) \times c^{\text{US}}(h, \mathcal{S}), \ell^{\text{US}}(h, \mathcal{S})) = V_h^{\text{RN}}(c^{\text{RN}}(h, \mathcal{S}), \ell^{\text{RN}}(h, \mathcal{S})),$$

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).[28] 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

[28] Given the utility function specification we use, $CE_1(h, \mathcal{S})$ can be computed directly from the value functions: $1 + CE_1(h, \mathcal{S}) = \left(\frac{V_h^{\text{RN}}(a, \mathcal{S})}{V_h^{\text{US}}(a, \mathcal{S})}\right)^{1/\gamma(1-\sigma)}$.

TABLE VIII – Average Welfare Gain from Tax Reform

	RN	BB
Average welfare difference:		
$\overline{CE}_1$	7.40%	5.58%
$\overline{CE}_2$	7.86%	4.71%
% with welfare gain	67.8	94.8

Note: The welfare figures report the percentage gain in consumption-equivalent terms from each tax reform relative to the current US benchmark economy.

in the benchmark economy ($\Gamma^{\text{US}}(h=1,\mathcal{S})$):[29]

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

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

$$\sum_{\mathcal{S}} \Gamma^{\text{US}}(1,\mathcal{S}) \times V_1^{\text{US}}((1+\overline{CE}_2)c^{\text{US}}(1,\mathcal{S}), \ell^{\text{US}}(1,\mathcal{S})) = \sum_{\mathcal{S}} \Gamma^{\text{RN}}(1,\mathcal{S}) \times V_1^{\text{RN}}(c^{\text{RN}}(1,\mathcal{S}), \ell^{\text{RN}}(1,\mathcal{S})),$$

and analogously for the BB reform.

Results. The welfare gains from the tax reforms are large: individuals born into the current US economy would have to be compensated on average by 7.4% ($\overline{CE}_1$) of their consumption in every date and state to be as well off as they would be if they were born into the RN tax reform economy, and by 7.86% using the $\overline{CE}_2$ measure (Table VIII). These large gains could be anticipated from the large increase in average consumption and the little change in labor hours we saw in Table V; however, the welfare changes also take into account the changes in the cross-sectional distributions, which do not seem to dampen the level gains.

How are the welfare gains distributed across the population? The individual-specific nature of the CE_1 measure allows us to answer this question. In Table IX, we divide the population into five age groups—real-life age 20, 21–34, 35–49, 50–64, and retirees—and the same six bins for entrepreneurial ability as before. Each cell reports the average

[29] We also calculated an analogous measure that integrates over individuals of all ages (and similarly for $\overline{CE}_2$ below) and found very similar results (available upon request).

TABLE IX – Welfare Gain by Age Group and Entrepreneurial Ability

	Entrepreneurial Ability Groups ($\overline{z}_i$ Percentiles)					
Age	0–40	40–80	80–90	90–99	99–99.9	99.9+
groups:	*RN Reform*					
20	7.0	7.3	7.9	8.9	10.6	11.7
21–34	6.5	6.3	6.3	6.6	7.0	6.8
35–49	5.1	4.4	3.9	3.3	1.7	0.1
50–64	2.3	1.8	1.4	0.8	–0.6	–1.8
65+	–0.2	–0.3	–0.4	–0.6	–1.2	–1.8
	BB Reform (SS Pensions Adjusted)					
20	4.9	5.3	6.0	7.2	9.3	10.5
21–34	4.7	4.6	4.8	5.4	6.1	6.2
35–49	4.2	3.7	3.4	2.8	1.4	–0.2
50–64	4.9	4.3	4.0	3.2	1.4	–0.2
65+	7.2	6.7	6.4	5.8	4.3	3.0

Note: 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.

welfare change for individuals within the $(h, \overline{z}_i)$ bin.

This table has several takeaways. First, all newborn groups gain from the RN reform, and these gains are fairly evenly distributed across ability groups—ranging from 7% for the lowest 40% by ability to 11.7% for the top 0.1% group.[30] Second, welfare gains decline with age, which is not surprising: since wealth rises and productivity (z_{ih}) falls on average with age, the ratio of capital income to wealth falls, so the tax burden of wealth taxation rises relative to capital taxation as individuals get older. This effect partially offsets the gains from higher wages experienced by all workers in the RN economy, leading to overall declining welfare gains by age. That said, the welfare change is positive for all working age groups, except those with very high ability—the top 1% of the $\overline{z}_i$ distribution—whose losses from higher taxes on their large wealth holdings outweighs their gains from higher wages.

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

The welfare losses of retirees from the RN reform follow from the fact that their pensions were kept fixed at their US benchmark level, so they do not share the wage gains experienced by workers, yet their 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 VIII) are slightly lower than the RN reform—5.58% and 4.71% for the $\overline{CE}_1$ and $\overline{CE}_2$ measures, respectively—simply because more revenue needs to be raised to pay for higher pensions. On the flip side, now all retiree groups gain significantly from the reform (lower panel of Table IX). Overall, 68% of individuals in the economy experience a welfare gain under the RN reform, and this fraction jumps to 95% under the BB reform.

Before we conclude this section, let us summarize the key conclusions. First, an economy with wealth taxes can raise the same amount of revenue as one with capital income taxes (keeping all other tax rates constant) with less distortion. The result is a reduction in the misallocation of capital, yielding higher average wages, consumption, and welfare. Second, welfare gains are relatively evenly distributed, with newborns of all entrepreneurial ability groups preferring the wealth tax economy. The gains become smaller with age and are negative for older individuals, especially those with high wealth. Third, allowing pensions to rise with average labor income (BB reform) yields somewhat lower average welfare gains but spreads the gains to the vast majority of the population. Finally, although these results show that wealth taxes outperform capital income taxes when other tax instruments are fixed, they are silent on whether either tax is desirable at all when the government can adjust the level of other taxes. In the next section, we turn to address this question of optimal taxation.

6 Optimal Taxation

In this section, we study the optimal taxation problem in which the government chooses a combination of tax instruments to maximize the ex ante lifetime utility of an individual who is born into the stationary equilibrium (implied by the chosen tax policy) subject to the constraint that it raises enough revenues to pay $G + SSP$ as before.[31] We consider two versions of this problem. In the first one, the government chooses flat-rate taxes on capital income and labor income. We refer to this as the optimal capital income tax (OCIT) system. In the second, the government chooses flat-rate taxes on wealth and labor income—the optimal wealth tax (OWT) system. Consumption taxes are held

[31] More 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 the superscript OPT refers to the relevant optimal tax economy.

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

	US Bench.	RN Tax Reform	Optimal Taxation		
			τ_k	τ_a	τ_a (thresh.)
			(1)	(2)	(3)
Tax Rates (%)					
τ_k	25.0	—	–34.4	—	—
τ_a	—	1.13	—	3.06	3.30†
τ_ℓ	22.4	22.4	36.0	14.10	14.20
Welfare gain (%)					
$\overline{CE}_1$	—	7.40	5.45	10.47	10.58
$\overline{CE}_2$	—	7.86	6.28	9.61	9.83

Notes: Percentage changes are computed with respect to the US benchmark economy calibrated in Section 4. †The optimal wealth threshold—below which $\tau_a = 0$—is found to be equal to 25% of average labor income, $\overline{y}$, in the benchmark economy.

fixed in all experiments. We will compare the two optimal tax systems with each other as well as with the current US benchmark. We also study an extension that introduces progressivity into the wealth tax system through an exemption level below which wealth is not taxed, and the government jointly chooses the exemption threshold and the wealth tax rate above that level. We consider further extensions below.

6.1 Results

We begin with an overview of the main results, summarized in Table X. In the first case, where the tax is levied on capital income, the optimal policy provides a large subsidy to capital income ($\tau_k = -34.4\%$) and a high tax rate on labor income ($\tau_\ell = 36\%$). This policy is substantially different from the current US system, and we will analyze in detail what drives this result. In the second case, where the tax is levied on wealth (without an exemption level, see column 2), the optimal policy is a wealth tax of $\tau_a = 3.06\%$ and a labor income tax of $\tau_\ell = 14.1\%$. Whereas the former, OCIT, raises higher revenues from labor income relative to the US system and uses it to subsidize capital, the wealth tax system does the opposite—raises revenues from wealth to significantly lower the tax rate on labor income.

Turning to welfare, we see that both optimal policies raise average welfare relative to the US benchmark, and the welfare gain from optimal wealth taxes (9.61%–10.47% depending on the measure) is higher than that from the optimal capital income taxes

TABLE XI – Optimal Taxation: Changes in Macroeconomic Outcomes

	Change from Benchmark			
	RN Tax Reform	Optimal Taxation		
		τ_k	τ_a	τ_a (thresh.)
$\%\Delta K$	19.4	69.0	2.8	0.4
$\%\Delta Q$	24.8	79.6	10.3	8.1
$\%\Delta L$	1.3	–1.2	3.9	3.7
$\%\Delta Y$	10.1	25.5	6.4	5.4
$\%\Delta$TFP	4.6	6.3	7.3	7.7
$\%\Delta C$	10.0	21.0	8.3	7.4
$\%\Delta w$	8.7	27.0	2.4	1.7
$\%\Delta w$ (net)	8.7	4.7	13.4	12.5
Δr	0.26	–1.51	0.68	0.78
Δr (net)	–0.91	–0.87	–1.92	–2.07

Note: Percentage changes are computed with respect to the benchmark economy without wealth taxes and capital income taxes of 25%. Changes in the interest rate are computed in percentage points. The net wage is defined as $(1-\tau_l)\,w$, and the net interest rate is defined as $(1+(1-\tau_k)\,r)\,(1-\tau_a)-1$. The TFP variable 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 $(\overline{E})$.

(5.45%–6.28%). Adding an exemption level to the OWT system (column 3) changes the results only marginally: the optimal exemption level is fairly low—about a quarter of average labor income—and the tax rates on both wealth and labor income rise only slightly. That said, given the extreme skewness of the wealth distribution, even this low threshold exempts 37% of the population from paying any wealth taxes, leading to a small improvement in average welfare compared with column 2. Now we delve into the details of each experiment.[32]

6.1.1 Changes in Macro Variables

Table XI shows the changes in aggregate variables in each optimal tax economy from their benchmark levels. The results from the RN tax reform are also reproduced here for comparison. Starting with the OCIT case, we see a substantial rise relative to the US benchmark in aggregate quantities across the board (with the exception of aggregate

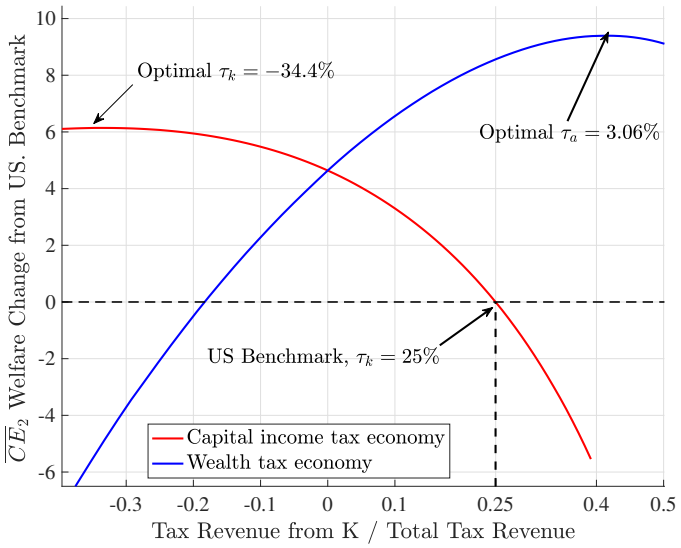
[32]Part of the welfare difference from the US benchmark arises from the differences in the stationary distributions in the two economies (e.g., $\Gamma^{\text{US}}(1,\mathcal{S})$ versus $\Gamma^{\text{OPT}}(1,\mathcal{S})$) that are used as weights in the welfare objective. To investigate how much this matters, we have also calculated average welfare in the optimal tax economies using the optimal allocations but imposing the US benchmark distribution, $\Gamma^{\text{OPT}}(1,\mathcal{S})$, and found that it makes a very small difference.

hours). In particular, K and Q are higher by almost 70% and 80%, respectively; L is lower, but only marginally (–1.2%). As a result, aggregate output and consumption are higher by 25.5% and 21%, respectively. Turning to the average wage, we see that although it also rises substantially (by 27%), most of these gains are undone by the higher optimal tax rate on labor income (36% versus 22.4% in the benchmark), resulting in an after-tax rise of only 4.7%. This latter figure is much lower than the 21% rise in consumption, which seems a bit surprising given that labor income makes up the bulk of total resources for most individuals in the model. As we will discuss later, the answer lies in the large rise in consumption inequality implied by optimal capital income taxes. (This gap was much smaller in the tax reform experiments above; see Table V.)

A very different picture emerges in the OWT experiment: K is only marginally higher (2.8%) than the US benchmark and Q is 10.3% higher, thanks to the more efficient allocation of capital. Aggregate hours rise quite a bit now—by 3.9%—spurred by the lower tax rate on labor income ($\tau_\ell = 14.1\%$). Consequently, output and consumption are higher by 6.4% and 8.3%, respectively. The average wage is only 2.4% higher than the US benchmark, but the net after-tax wage is 13.4% higher. This time, notice that consumption rises less than after-tax wages because of the higher taxes on wealth relative to the US benchmark, which leads to a smaller rise in post-production, after-tax resources (i.e., ω in (12)), moderating the rise in the consumption of the wealthy, and hence in the aggregate. Moreover, (the marginal utility of) consumption becomes more evenly distributed with wealth taxes. Finally, introducing an exemption level (column 3) makes a relatively modest impact on the aggregates, which is not surprising both because its level is low and because it affects the low-wealth individuals whose share in the aggregate economy is small. That said, the exemption level does matter for the *distribution* of welfare gains discussed later.

Comparing the two optimal tax systems reveals interesting contrasts. Broadly speaking, optimal capital income taxation results in rather dramatic changes in the economy. The subsidy to capital income both raises the income accruing to high-productivity entrepreneurs and further incentivizes them to save more, leading to a 70% higher capital level in the new economy and an increase in efficiency. Unlike wealth taxes, however, which raise revenues from capital, the subsidy on capital income requires more revenues to be raised from labor, leading to very small gains in after-tax wages relative to the growth in output and (before-tax) wages. In this sense, optimal capital taxes shift the tax burden from the wealthy to wage earners, delivering efficiency gains at the expense of large distributional losses (which we will quantify in the next subsection).

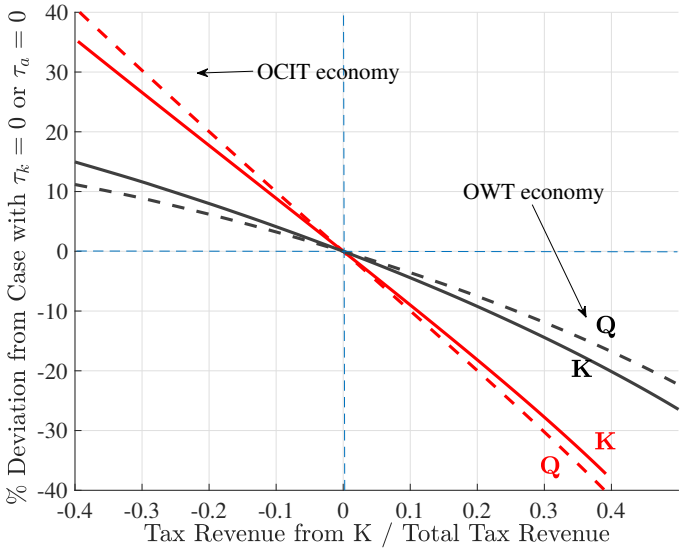
FIGURE 2 – Welfare Gains from Optimal Taxation



Contrast this with wealth taxes, which result in much smaller adjustments in the aggregate economy, with the capital stock barely moving from the US benchmark, and output and wages rising by smaller amounts. However, by enabling a lower tax rate on labor, it raises after-tax wages significantly, which in turn incentivizes work, leading to higher aggregate hours. In this sense, the OWT system shifts the overall tax burden in the opposite direction—from labor to capital—and further shifts the capital tax burden from high-productivity entrepreneurs to low-productivity ones.

Mechanisms at play. To better understand the differences between the two tax systems, in Figure 2 we plot the objective functions that are maximized in the respective optimal tax experiments. The red line shows the objective values under capital income taxes, and the blue line shows it under wealth taxes. The x-axis shows the tax revenue raised from capital as a fraction of total tax revenue. Because the latter $(G + SSP)$ is kept fixed throughout, as the revenue share of capital varies along the x-axis, the labor income tax adjusts in the background to balance the government budget. Thus, for a given tax system, the optimal tax rate is found where the objective value is maximized on this graph. The current US benchmark corresponds to the point at $x = 0.25$ on the red line, since as noted earlier, $\tau_k = 25\%$ generates a revenue share of (coincidentally) 25% from capital. To make the figure more informative, the y-axis does not report (ex ante lifetime) utilities directly but instead reports the consumption-equivalent welfare change implied by each objective value relative to the US benchmark. Thus, the current

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



US benchmark point corresponds to zero on the y-axis.

This figure offers several takeaways. The first one is the obvious contrast between the slopes of the two lines: whereas welfare declines as more tax revenues are raised from capital under capital income taxation, it rises strongly under wealth taxation. Despite this apparent contrast, both patterns are driven by the same basic principle: in the presence of persistent return heterogeneity, taxing capital has a stronger distorting effect than without heterogeneity. For example, under capital income taxation, those who pay the most 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 persistent return heterogeneity is dealt with by imposing a relatively high tax on wealth, which creates the same type of reallocation toward more productive entrepreneurs. Of course, because wealth is still taxed, the effects on savings incentives are not as strong as with capital income subsidies, which does not create a large rise in the capital stock, so the bulk of the gains come from reallocation.

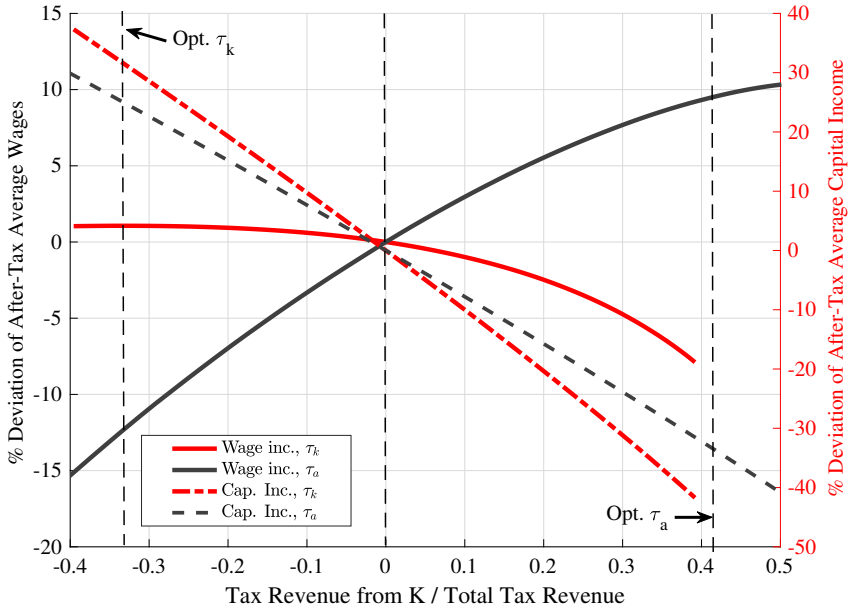
The distorting effects of the two taxes can be seen more clearly in Figure 3, which plots the average capital level—both raw and quality adjusted—under each tax system as the revenue share is varied. As expected, both K and Q decline as capital is taxed under both systems and rise as it is subsidized. But, the two taxes have two major

differences. First, under wealth taxation, K (solid black line) declines much more slowly with higher taxation than under capital income taxation (solid red line). In other words, the same amount of revenue can be raised with the former, with a smaller distortion to the capital level than with the latter. Second, and to compound the first effect, Q declines more slowly than K under wealth taxation whereas the reverse happens—it declines more steeply under capital income taxation. Thus, whereas wealth taxation improves the efficiency in how capital is allocated (Q versus K), capital income taxation does the opposite. The relationship reverses for subsidies, so the capital income subsidy results in both an increase both K and a more efficient allocation of it.

To better illustrate the trade-offs between capital and labor income implied by each tax system, Figure 4 plots the average after-tax labor income and capital income under both tax systems. Although the welfare objective plotted in the previous figure also takes into account preferences and the resulting distributional effects beyond average consumption, this figure is still quite informative about the key trade-offs. With this in mind, notice first that average after-tax capital income is declining with the capital tax (dashed lines) under both tax systems, which is obvious. What is less obvious is that it declines more slowly under wealth taxation, a point we expand on in a moment. Second, the main contrast is in average after-tax labor income (the solid lines), which declines with the tax rate on capital under capital income taxation and rises strongly under wealth taxation, largely because of the labor income tax rate rising with capital tax revenues in the former and falling in the latter. This is the main driver of the contrast in how welfare varies with capital taxation we saw above.

Third, notice also that, under capital income taxation, the gains in after-tax labor income flatten out considerably once the tax crosses into the subsidy territory. For example, after-tax labor income is 0.60% higher when the revenue share is –0.1 and is 1.05% higher at the optimum when the revenue share is –0.34. Hence, the main reason the optimal capital income subsidy is large is the nearly linear gains in (after-subsidy) capital income, which accrues to the top end. Of course, higher capital subsidies and labor income taxes also raise inequality, which keeps the optimal subsidy from being higher than it is. Notice also that the flatness of the average labor income line, which is also seen as a relatively flat spot in the average welfare function in Figure 2, suggests that the large magnitude of the capital income subsidy would be sensitive to modest changes in model details. Our robustness analysis in the next section confirms this, although most of the tax rates we obtain range from –20% to –40%, so the sign of the tax seems quite robust.

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



These findings—of optimal capital subsidies—are in sharp contrast with some well-understood results in the literature, which find that taxing capital income is optimal. For example, Conesa et al. (2009) find an optimal $\tau_k = 36\%$ in an overlapping-generations model that shares many similarities with ours. One key difference is return heterogeneity—which is abstracted from in their model. Without this heterogeneity, the wealthy are wealthy because they have been lucky in the labor market and saved their high earnings; however, they are not any better at investing their wealth than others. It is clear from the mechanisms we have discussed so far that this disconnect—between being wealthy and being a good investor/entrepreneur in such models—makes capital income taxation significantly less distorting than shown here. As a result, the distributional *benefits* of a *positive* tax on capital income can outweigh the costs of distortion and make a positive tax rate optimal. In the next section, we examine this and related issues more closely and reconcile our findings with these earlier studies.

Next, we turn to wealth taxes. As seen again in Figure 4, the rise in after-tax wage income with the level of wealth taxes is very steep throughout. These gains are balanced against the nearly linear decline in capital income due to higher taxation, and the optimal level trades off these two considerations against each other. Because labor income constitutes the bulk of total lifetime resources for the substantial majority of individuals, the utilitarian social welfare function places a large implicit weight on it, so overall the optimal tax rate leans heavily toward a rate that maximizes average after-tax

TABLE XII – Welfare Gain by Age Group and Entrepreneurial Ability

	Entrepreneurial Ability Groups ($\overline{z}_i$ Percentiles)					
Age	0–40	40–80	80–90	90–99	99–99.9	99.9+
groups:	*Optimal Capital Income Taxes*					
20	4.0	5.6	7.2	9.5	13.0	14.9
21–34	3.7	5.0	6.2	7.9	10.4	11.4
35–49	2.7	3.3	3.8	4.0	3.5	2.5
50–64	1.1	1.4	1.6	1.5	0.6	–0.4
65+	–0.1	0.1	0.2	0.2	–0.2	–0.8
	Optimal Wealth Taxes					
20	10.0	9.7	10.1	11.1	13.1	14.4
21–34	9.2	7.9	7.3	7.1	6.6	5.6
35–49	6.8	4.9	3.7	2.1	–1.3	–4.4
50–64	2.7	1.4	0.6	–0.8	–3.7	–6.2
65+	–0.6	-0.9	–1.2	–1.8	–3.2	–4.5

Note: Each entry reports the average welfare gain (CE_1) from the corresponding optimal tax experiment for individuals in a given age and entrepreneurial productivity group (ranked based on the permanent component of entrepreneurial productivity $\overline{z}$). The average is computed with respect to the benchmark distribution.

labor income. In the sensitivity analyses that we summarize in the next section, we found optimal tax rates ranging from about 2% to 3% under reasonable calibrations. For example, when we introduced progressive labor income taxes, we found the OWT rate to be 2.4%, with slightly higher welfare gains than the baseline analysis conducted here.

6.2 Who Gains and Who Loses?

So far, we have focused on aggregate outcomes, but as we have seen earlier in the tax reform analyses, there are important distributional consequences. In Table XII, we report the distribution of welfare gains and losses (CE_1) across the population, in an analogous format to Table IX above for the tax reforms.[33] The top panel shows the results for the OCIT (relative to the US benchmark) and the bottom panel does the same for the OWT system. Broadly speaking, we see some of the same patterns as

[33]The results based on an optimal (linear) wealth tax with a threshold limit are similar to those in the OWT system, and we refer to them when appropriate.

we did with the tax reform earlier: the young gain more than the old; the gains for the newborns (age 20) increase with entrepreneurial ability level, $\overline{z}$, especially at the very top; but the relationship between gains and $\overline{z}$ is reversed at older ages, so older individuals with very high $\overline{z}$ often experience welfare losses.

Beyond these general patterns, however, note some important differences. First, starting with the newborns (whose expected lifetime utility is maximized in the optimal tax system), the welfare gains from the OWT system are distributed quite evenly across ability groups: they rise from a 10% gain for those in the bottom 40% of the $\overline{z}$ distribution, to 11.1% for those in the 90th–99th percentiles group, and to a 14.4% gain for the very top—the 0.1% group. This is not the case under OCIT, where the gains are only 4% for the bottom group and rise all the way to 14.9% at the very top. Notice that the key difference is not at the top but at the bottom: individuals with low $\overline{z}$ do not gain too much because the rise in average after-tax labor income is very modest. At the top, the gains are similar (14.4% vs. 14.9%) under both systems.

A second notable difference is that, under wealth taxation, welfare gains decline more sharply with age at the top end—turning into larger losses—for the top 1% or so. The reason is clear: these individuals have a lot of accumulated wealth but face declining productivity (and hence a declining profit rate), so looking ahead, the burden of wealth taxes is heavier than that of the capital income taxes imposed in the US benchmark (not to mention the OCIT system, which involves large subsidies). Finally, retirees have lower welfare under wealth taxes for the same reasons explained earlier for the tax reform.[34] This is one place where introducing the exemption threshold into the OWT system makes a big difference and mitigates these welfare losses substantially. In particular, of the retirees in the bottom 90% of the entrepreneurial ability distribution, the percentage of those who have higher welfare under wealth taxation rises from 2% in the baseline case without the exemption level to about 70% with the optimal exemption level (see Table B.5 in Appendix B.2).[35]

[34] To be more specific, the after-tax interest rate (r(net) in Table XI) is 1.92% lower under OWT relative to the benchmark but only 0.87% lower under OCIT. So, retirees' after-tax capital income falls more under wealth taxation than under the latter. (They also have much higher wealth in the latter stationary equilibrium, but that brings back the transition discussion above.)

[35] Another straightforward change would be to allow retirement pensions to rise with the average labor income, as we studied in the BB tax reform analysis earlier, but we have not attempted to solve this extension.

TABLE XIII – Decomposition of Welfare Gains

	RN Reform	OCIT	OWT
$\overline{CE}_2$ (%)	7.86	6.28	9.61
	Consumption		
Total	8.27	5.90	11.02
Level	10.01	21.04	8.28
Distribution	–1.58	–12.51	2.53
	Leisure		
Total	–0.38	0.36	–1.27
Level	–0.66	0.73	–2.21
Distribution	0.27	–0.38	0.76

The table reports the decomposition of the average welfare gains of newborns as measured by $\overline{CE}_2$. See the text for details.

Decomposing the welfare gains

To understand the sources of welfare gains and losses, we implement a useful decomposition following Conesa et al. (2009). Specifically, the aggregate welfare gain, $\overline{CE}_2$ (which corresponds to the objective function maximized by the government), is first decomposed into effects from changes in consumption and leisure. In turn, each component is further decomposed into level effects and distributional effects (e.g., the change in average consumption versus in its distribution, and the same for leisure).[36,37]

[36] A similar decomposition was earlier used by Flodén (2001), where total welfare changes are expressed in terms of changes in levels, changes in uncertainty, and changes in inequality.

[37] More concretely, let CE be the aggregate welfare gain, and CE_C and CE_L be the components of CE due to changes in consumption and leisure allocations, respectively, such that $1+CE = (1+CE_C)(1+CE_L)$. Then, CE_C is given (for example, for the OWT economy) by

$$V_0((1+CE_C(h,\mathcal{S}))\times c^{\text{US}}(h,\mathcal{S}),\ell^{\text{US}}(h,\mathcal{S})) = \widetilde{\mathbb{V}}_0(c^{\text{OWT}}(h,\mathcal{S}),\ell^{\text{US}}(h,\mathcal{S})),$$

and CE_L is

$$V_0((1+CE_L(h,\mathcal{S}))c^{\text{US}}(h,\mathcal{S}),\ell^{\text{US}}(h,\mathcal{S})) = \widetilde{\mathbb{V}}_0(c^{\text{US}}(h,\mathcal{S}),\ell^{\text{OWT}}(h,\mathcal{S})).$$

Further, CE_C can be decomposed into a level and a distribution component, $CE_{\overline{C}}$ and CE_{σ_C}, respectively:

$$V_0((1+CE_{\overline{C}}(h,\mathcal{S}))\times c^{\text{US}}(h,\mathcal{S}),\ell^{\text{US}}(h,\mathcal{S})) = \widehat{\mathbb{V}}_0(\widehat{c}(h,\mathcal{S}),\ell^{\text{US}}(h,\mathcal{S})),$$

where $\widehat{c}^{\text{US}}(h,\mathcal{S}) = c^{\text{US}}(h,\mathcal{S})\frac{\overline{C}}{\overline{C}^{\text{US}}}$ and $\widehat{\mathbb{V}}_0\left((1+CE_{\sigma_C})\,\widehat{c}^{\text{US}}(h,\mathcal{S}),\ell^{\text{US}}(h,\mathcal{S})\right) = \widetilde{\mathbb{V}}_0(c^{\text{OWT}}(h,\mathcal{S}),\ell^{\text{US}}(h,\mathcal{S}))$ where one can show that $1+CE_C = (1+CE_{\overline{C}})(1+CE_{\sigma_C})$. A similar decomposition holds for leisure.

Table XIII reports the decomposition results. First, starting with the OWT system (the last column), the 9.61% average welfare gain under OWT arises from an 11.02% welfare gain from changes in consumption and a 1.27% welfare loss from changes in leisure. The small change coming from leisure is not very surprising given the small change in aggregate hours we saw earlier. So, focusing on consumption, of the 11.02% gain from consumption, 8.28% is due to the positive level effect (higher average consumption), and 2.53% is due to the positive distributional effect (improved distribution of the marginal utility of consumption). Therefore, wealth taxation improves welfare by both growing the economy *and* making it more equitable. This result is very different from the usual intuition about the trade-offs involved in capital income taxation in incomplete markets models. As mentioned above, in that case, capital income taxation results in a level loss due to distortions of the positive tax rate, which is weighed against the less dispersed consumption distribution it implies (thereby reducing inequality or serving as insurance against idiosyncratic uncertainty). Here gains arise from both components, with the bulk of the gains coming from the improved efficiency and growth component.

Turning to capital income taxation, we see that the usual mechanism just described operates here too (middle column of Table XIII), but is manifested in the opposite direction: because the optimal policy is a subsidy, there are large level gains combined with large distributional losses: of the 5.90% welfare gain from consumption, a substantial 21.04% is from a positive level effect, and –12.51% is from a negative distributional effect.[38] So, the same efficiency versus equity trade-off is present in our model, except that the sign is reversed because the optimal policy is a capital income subsidy rather than a tax. Finally, despite the small contribution of leisure to welfare changes, it is worth noting that the level and distributional effects under the two systems again go in opposite directions: whereas wealth taxes generate a 2.21% welfare loss from a level effect (individuals work more), it also produces a small gain of 0.76% from an improved distribution. The signs are reversed for capital income taxes, with a small level gain and smaller distributional loss.

In robustness analysis, we have examined alternative versions of this decomposition (e.g., decomposing the average welfare of the population rather than of newborns, using the $\overline{CE}_1$ measure, among others). The substantive conclusions remained consistent with this baseline analysis, despite some changes in magnitudes. For example, when the CE_1 measure is used, wealth taxes no longer produce a welfare gain from distributional effects; it turns into a small loss of 0.62%. However, the distributional gains from leisure are

[38] Note that the decomposition is multiplicative, so $(1 + 0.059) \approx (1 + 0.2104)(1 - 0.1251)$.

larger than before, so the overall contribution of leisure is no longer negative. Perhaps more importantly, the bulk of the welfare gains come from the level effect of consumption, as in the baseline case. Furthermore, the distributional losses coming from consumption are always far greater under capital income taxes as before. These results are available upon request.

7 Robustness

We begin, in Section 7.1, with an extension that introduces a transition path into the OCIT and OWT analyses of the previous section. Then, in Section 7.2, we present eight robustness experiments that include: (i) changing the timing of wealth taxes from post-production to beginning-of-period wealth, a; (ii) modeling progressive labor income taxation; (iii) eliminating the stochastic variation in entrepreneurial productivity ($z_{ih} \equiv \overline{z}_i$ for all i, h); (iv) removing the borrowing constraint faced by entrepreneurs (i.e., $\vartheta = \infty$); (v) reducing the CES curvature to $\mu = 0.8$; (vi) eliminating return heterogeneity ($z_{ih} = 1$ for all i, h, *and* $\mu = 1$) which allows a comparison with Conesa et al. (2009); (vii) introducing estate taxation; and (viii) measuring wealth in present value rather than in book value. In each case, we follow the same calibration procedure as in our benchmark economy, with two exceptions: in case (iii) we do not target the fraction of self-made billionaires, and in case (v) the model is unable to match the wealth concentration in the data. We repeat the RN tax reform experiment and the two optimal tax experiments and report the results in Tables XV and XVI, respectively.

7.1 Accounting for the Transition Path

We extend the optimal policy analysis of Section 6 by modeling the transition path after the switch to a new optimal policy regime to understand how the individuals who are alive at the time of the switch fare. To be precise, the goal here is not to do an optimal tax analysis with transition but rather to solve for an equilibrium that holds throughout the transition and in the new steady state while minimally deviating from the OCIT and OWT tax rates found above. To this end, we fix one of the two policy instruments (e.g., τ_ℓ) at its non-transition optimum found earlier (Table X), allow the government to run a budget deficit (or surplus) during the transition, and choose the other policy instrument (e.g., τ_a) 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. We consider different versions of this setup.

TABLE XIV – Extension: Policy Analysis Accounting for the Transition Path

Policy instrument fixed:	OCIT		OWT	
	τ_k	τ_ℓ	τ_a	τ_ℓ
τ_k	–34.38*	–26.80	0.00	0.00
τ_a	0.00	0.00	3.06*	3.40
τ_ℓ	37.41	36.00*	15.40	14.10*
$\overline{CE}_2$ (newborn)	–5.30	–4.85	7.71	8.43
$\overline{CE}_2$ (all)	–3.86	–3.55	4.65	5.07

Note: Both the tax rates and welfare figures are reported in percents. *indicates that the tax rate is fixed at its non-transition optimum shown in Table X.

Starting with the OCIT policy, first, we fix τ_k at the non-transition optimum (–34.38% from Table X) and choose τ_ℓ as just described. The new equilibrium value of τ_ℓ is 37.41% (Table XIV, first column) compared with 36.0% without transition, and K is 67.2% higher (vs. 69%) in the new steady state. Most of the transition is completed between 45 to 50 years: 95% of the gap (between the first year of transition and the new steady state) is closed in 41 to 43 years for output, consumption, and employment; and 54 years for K versus 42 years for Q, where the faster convergence of the latter reflects the contribution of improving capital allocation with the new optimal policy. Full convergence (99.9%) takes about 80 years. During the first 36 years of the transition, the government runs a deficit, accumulating debt that reaches 116% of output in that year and 130% by year 80, after which the debt is rolled over.[39] The higher labor income tax rate, now applied to a larger aggregate wage income in the new steady state, pays for both the new capital income subsidy and the interest on this debt.

The results for welfare changes are especially interesting. We start by comparing the welfare of newborns—those who enter the economy the first year of the reform—to the alternative without the new policy. The average welfare of this cohort *falls* by 5.30% in consumption-equivalent terms, in contrast to the 6.28% *gain* we found without transition in Section 6. Broadening the comparison to include the entire population alive at the

[39]The reason for this result is clear: the new policy involves an immediate loss of capital income tax revenue and need for funds for the capital income subsidy. While the higher labor tax rate increases revenues from labor income, this rise is tempered by an immediate 9.6% fall in labor supply in response to the higher tax rate. The resulting revenue shortfall causes the early budget deficit, which eventually closes as the growing economy raises wages, reversing the decline in employment, and generating more revenue from labor income.

time of the reform shows an average loss of 3.86% compared with a 4.87% gain without transition (not reported earlier).

This dramatic reversal stems from two sources. The first one is straightforward: the transition analysis makes explicit the cost of accumulating the large capital stock (about 67% higher) found in the new steady state, which requires higher savings and lower consumption early in the transition. Second, τ_ℓ jumps from 22.4% in the benchmark to 37.41% in the first year of transition, lowering the after-tax wage labor income both directly and indirectly—by depressing labor supply, which falls by 9.2% in the first year of the new policy, driving the 4.7% fall in output upon impact. Both of these costs are borne in the earlier 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.

The other polar case—fixing $\tau_\ell = 36\%$ and finding τ_k—yields a similar picture, which we discuss more briefly in the interest of space. The capital income subsidy is now smaller, $\tau_k = -26.8\%$, compared with -34.38% without transition. More importantly, both newborns and the overall population experience welfare losses, albeit slightly smaller than before: by 4.85% and 3.55%, respectively. The conclusion from these two experiments is clear: accounting for the transition path is critical for the OCIT analysis and turns the apparent gains into nontrivial losses.

Turning to optimal wealth taxes (OWT) with transition, first, when τ_a is fixed (at 3.06%), the new equilibrium value of τ_ℓ is 15.40% (versus 14.10% without transition). As for welfare, newborns continue to experience large gains: average welfare rises by 7.71% compared with the 9.62% gain found without transition. Looking at the overall population, the transition barely changes the results, with welfare rising by 4.65% versus 4.79% without transition. The other polar case ($\tau_\ell = 14.10\%$ fixed) delivers an optimal wealth tax of $\tau_a = 3.40\%$ and yields a gain of 8.43% for the newborn cohort. Perhaps more surprisingly, the gain for the overall population is 5.07%, which is slightly *higher* than the 4.79% figure without transition, so there is no presumption that accounting for transition automatically implies smaller welfare gains.[40]

These welfare results—as well as why they are so different from the OCIT case—are driven by two main factors. The first one works through the decline in labor taxes and

[40]In this case, the welfare losses are confined to older individuals at the top of the $\overline{z}$ distribution, whereas the young all gain regardless of their $\overline{z}$ group. This is the exact opposite of the OCIT transition: those in the bottom 80% of the $\overline{z}$ distribution experience welfare losses with the young below the median bearing the brunt of the losses, while the older individuals at the top experience larger gains (see Table B.7).

has both a level and a distributional effect. Basically, the immediate fall in τ_ℓ causes after-tax wage income to jump, raising average consumption by 4.7% in the first year of reform.[41] Moreover, the rise in consumption is proportionally 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. Another way to state the same result is that the OWT system reduces the share of tax revenue raised from labor income by shifting it toward wealth, yielding a more even distribution of consumption. The second key difference of the OWT transition is that it 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 less than 2%. Thus, unlike in the OCIT case, there is no aggregate capital accumulation that can reduce consumption during the transition. Rather, the gains are driven by the reallocation of capital toward more productive individuals.[42]

Overall, these extensions show that considering the transition path has little effect on the implications of optimal wealth taxes found in the baseline analysis, whereas it upends the welfare gains from large capital subsidies that emerged without considering the transition. In this sense, this analysis strengthens the case for wealth taxes and weakens it for capital income subsidies.

7.2 Extensions and Sensitivity Analysis

I. Alternative timing: taxing beginning-of-period wealth. As discussed in Section 3.3, in the baseline model, the wealth tax was assumed to be imposed on post-production wealth (i.e., $\tau_a \times (a + (\pi(a, z) + ra))$ to align the timing with that of capital income taxes. Consequently, the tax applied not only to assets a but also to profits earned during the period. To see whether this timing assumption is critical, we solve the model again but this time, taxing beginning-of-period wealth: $\tau_a \times a$. As seen in Tables XV and XVI, this change makes only a modest difference, and when it does, it actually raises the welfare gains relative to the baseline (by about 0.4% and 1.2% for the tax reform and OWT cases, respectively). The reason for these improvements should be intuitive: in the original timing, the wealth tax is also partly on profits (or capital

[41] The reduction in τ_ℓ also raises labor supply, further raising output, although this effect is not large early in the transition.

[42] That said, K is not constant during the transition but follows a nonmonotonic path, falling for the first 13 years and then taking another 35 years to rise back to its pre-reform level. This is because switching to wealth taxes reduces the after-tax return of many older and wealthier individuals who now find it optimal to increase consumption and spend down their wealth. Although the opposite happens for young and productive individuals, their wealth is a smaller fraction of the aggregate. However, note that Q rises monotonically—thanks to reallocation—so output rises throughout the transition.

income), which, as explained before, is highly distorting in this framework. The alternative formulation relieves the latter from wealth taxation, making it a pure wealth tax which is more effective.

II. Progressive labor income tax. To introduce progressivity into the labor income taxes, we follow Heathcote, Storesletten and Violante (2014) and write after-tax labor income as $(1-\tau_\ell)(wy_h n)^\psi$ with $\psi = 0.815$. We set τ_ℓ to obtain an average labor income tax rate of 22.4%—the same rate as in the US benchmark. As before, the labor income tax is kept fixed in the tax reform experiment, whereas in the optimal tax analyses, τ_ℓ and ψ are jointly chosen with τ_k or τ_a. As seen in both Tables XV and XVI, the results remain very similar to the baseline case, with a slightly smaller welfare gain in the tax reform case and a slightly higher one in the optimal tax experiments. One notable difference is that the optimal labor income tax is more progressive than the US benchmark, $\psi = 0.72$, which allows a lower wealth tax rate of $\tau_a = 2.4\%$.

III. Constant z over the life cycle. As discussed in Section 3, without the stochastic fluctuations in z_{ih} as modeled in equation (2), even the best calibration cannot simultaneously match the steady state wealth concentration and the fraction of self-made billionaires (the latter is 18.5% here versus 50% in the baseline). The average welfare gains from both the tax reform and OWT are lower than in the baseline but still large at about 5%–6%. One notable difference is in τ_k^{OPT}, which is no longer a large subsidy but is instead close to zero (–2.33%). A second finding is that switching to wealth taxes causes a larger efficiency gain—equivalent to a TFP rise of 11.7%, which is up to twice the gain in the baseline calibration (ranging from 4.6% to 7.7% in Tables V and XI).[43]

IV. No financial constraint: $\vartheta = \infty$. In this extreme case scenario, marginal returns are equalized across entrepreneurs, so capital is no longer misallocated. Yet, switching to a wealth tax still increases welfare. The gain is about 1.5% across the tax reform and optimal tax case. Although this gain is significantly smaller than in the baseline case, that there is any gain at all may seem surprising. First note that both K and Q rise by the same amount (by 6.28% in Table XV), so there is no TFP improvement because

[43]Both results can be understood as follows. With constant z_{ih}, returns are much more persistent over the life cycle (declining gradually because $\mu < 1$), which creates two important effects. First, there is less capital misallocation in the model, both because one source of misallocation—stochastic z_{ih}—is eliminated and because the new calibration implies less heterogeneity in $\bar{z}_i$, thereby implying effectively looser constraints. Second, the efficiency gains from a given reallocation are larger because now the correlation between current wealth and future productivity is higher, so there are potentially larger efficiency gains from wealth taxation. These explain both the lower gain in welfare and the fact that the larger part of it coming from TFP gains. Of course, the model cannot match the dynamics of wealth, which is one of the reasons why this is not our preferred baseline.

TABLE XV – Robustness: Tax Reform Experiments

	Baseline	Tax beginning of period wealth	Progressive labor tax	Constant z	No constraints $\vartheta = \infty$	Lower curvature $\mu = 0.8$	Homogeneous rates of return	Estate taxes	Present value
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
τ_a	1.13%	1.18%	0.90%	1.23%	1.65%	1.24%	1.92%	0.95%	1.26%
	Welfare Change								
$\overline{CE}_1$	7.40	7.75	6.48	5.46	1.86	7.54	0	8.22	6.07
$\overline{CE}_2$	7.86	8.25	7.06	5.36	1.43	7.85	0	8.80	6.48
	Change in Macro Variables (%)								
K	19.37	20.41	21.27	9.56	6.28	16.43	0	21.05	15.60
Q	24.79	26.10	25.61	22.37	6.28	21.25	0	27.90	19.87
w	8.70	9.10	9.25	7.66	2.10	7.77	0	9.75	7.08
Y	10.10	10.60	10.01	9.54	3.02	8.38	0	11.25	8.18
L	1.28	1.33	0.69	1.75	0.91	0.57	0	1.37	1.04
C	10.01	10.50	10.01	11.25	2.93	8.33	0	11.31	8.17

Note: The eight robustness experiments are as follows: (i) taxing beginning-of-period wealth; (ii) modeling progressive labor income taxation; (iii) eliminating the stochastic variation in entrepreneurial productivity ($z_{ih} \equiv \overline{z}_i$ for all i, h); (iv) removing borrowing constraints, $\vartheta = \infty$; (v) eliminating return heterogeneity ($z_{ih} = 1$ for all i, h and $\mu = 1$); (vi) reducing the CES curvature to $\mu = 0.8$; (vii) introducing estate taxation; and (viii) measuring wealth in present value rather than in book value.

there is no misallocation to begin with. Thus, the gain is entirely from the effect of wealth taxes on savings incentives and the resulting rise in the capital stock. Turning to the optimal tax experiments, we see that τ_k^{OPT} is now positive, 13.6%, but still smaller than the benchmark level of 25%, whereas τ_a^{OPT} is 1.57%, half the rate in the baseline case. As before, the OWT system delivers a higher welfare gain than the OCIT system.

V. Reducing the CES curvature: $\mu = 0.80$. As noted earlier, μ is an important parameter as it determines the extent of diminishing returns in entrepreneurial production. Holding other parameters fixed, a lower μ implies lower efficiency gains from wealth taxation, since diminishing returns lowers the correlation between past wealth and future productivity. However, the same mechanism implies that for the same parameter values, wealth inequality will be lower in the model, so a recalibration requires a higher dispersion in $\bar{z}_i$ ($\sigma_{\varepsilon_{\bar{z}}}$) to match the top (1%) wealth concentration, which undoes the first effect. As a result, the welfare results remain broadly similar to the baseline model: gains are almost unchanged for the tax reform, slightly lower (8.32%) in the OWT system (with a lower tax rate $\tau_a^{\text{OPT}} = 2.12\%$), and slightly higher (7.38%) in the OCIT system.[44]

VI. Eliminating return heterogeneity. As emphasized before, return heterogeneity is the key new ingredient in our model that distinguishes our model from previous work, so this experiment allows us to investigate whether eliminating it reconciles our findings with earlier ones. We check several things. First, we noted earlier that with $z_{ih} = 1$ and $\mu = 1$, capital income and wealth taxes become equivalent. This is confirmed in column (5) of Table XV for the tax reform case, leading to no changes in any variable, as well as in Table XVI where both optimal tax systems yield the same exact outcomes.

Second, and more importantly, the optimal capital income tax is now positive and large: $\tau_k^{\text{OPT}} = 25.4\%$ compared with -34.4% in the baseline with return heterogeneity. This new tax rate is also much closer to Conesa et al. (2009) but is still below their 36% figure. Another small difference between the two models is in the treatment of accidental bequests. They are passed on to one's offspring in our model but are distributed equally to all newborns in their model. When we change this assumption to match theirs, the new optimal rate we find is $\tau_k^{\text{OPT}} = 42.4\%$, slightly exceeding their figure. Therefore, this experiment confirms their findings of a large positive capital income tax rate and also

[44] Another effect of the stronger diminishing returns implied by the lower μ is the extent to which the model can generate a Pareto tail. As shown in Appendix C, Figure C.2a, the model with $\mu = 0.8$ continues to generate an almost linear Pareto tail—which overlaps better with the data than the baseline calibration. In additional analyses, we found that the Pareto tail starts to curve down around $\mu = 0.75$ and disappears for $\mu = 0.7$ and lower.

TABLE XVI – Robustness: Optimal *Tax* Experiments

	τ_k	τ_ℓ	ψ	τ_a	$\overline{CE}_2$ (%)
			Baseline Model		
US Bench.	25%	22.4%		—	—
OCIT	–34.4%	36.0%		—	6.28
OWT	—	14.1%		3.06%	9.61
		I. *Timing: Tax on Beginning-of-Period Wealth*			
OWT		12.0%		3.62%	10.87
			II. *Progressive Labor Income Tax*		
US Bench.	25%	15.0%	0.815	—	—
OCIT	–38.8%	29.3%	0.720	—	9.31
OWT	—	12.7%	0.720	2.40%	10.71
			III. *Constant z over the Life Cycle*		
OCIT	–2.33%	29.0%		—	3.27
OWT	—	18.5%		2.21%	5.80
			IV. *No Financial Constraint:* $\vartheta = \infty$		
OCIT	13.6%	26.0%		—	0.41
OWT	—	22.7%		1.57%	1.43
			V. *Lower Curvature in CES:* $\mu = 0.80$		
OCIT	–38.6%	37.7%		—	7.38
OWT	—	18.6%		2.12%	8.32
			VI. *Homogeneous Rates of Return*		
OCIT	25.4%	22.33%		—	0.005
OWT	—	22.33%		1.93%	0.005
			VII. *Estate Taxes*		
OCIT	–32.2%	33.7%		—	9.26
OWT	—	13.0%		3.12%	11.02
			VIII. *Present Value*		
OCIT	–18.3%	33.56%		—	4.16
OWT	—	16.45%		2.64%	7.38

shows the importance of return heterogeneity for studying capital taxation, regardless of whether this is done through taxing capital income or wealth.

VII. Estate taxes. We now introduce estate taxes that mimic the US system, with a 40% tax rate above an exemption level of $5 million. The rest of the model is recalibrated with this new feature. The results are similar, with wealth taxes resulting in a slightly higher welfare increase than in the baseline case (by 1%–1.5%) in both the reform and optimal tax cases. The reason is that estate taxes provide another source of wealth taxation, so as wealth concentration rises because of the reform or optimal taxation, more revenue collection occurs through estate taxation, which improves the overall efficiency of the tax system. Notice that welfare gains are larger even under OCIT because part of the system now involves a wealth tax (on estates).

VIII. Present value. In the baseline analysis, the notion of wealth we used was an individual's assets, a. Although this approach is fairly standard, a sensible alternative is to use the "market value" of the firm that each entrepreneur owns, the value of which depends not only on a, the "book value," but also on her z_{ih}.[45] To see whether the difference matters, we replaced a with the expected present value of future earnings, discounted by the average rate of return in the economy. The wealth tax is still based on the book value. For given parameter values, market value wealth is more dispersed than the book value (because it also depends on z_{ih}), so the recalibration to the same inequality targets requires a lower $\sigma_{z_{ih}}$ than before, leading to slightly smaller efficiency gains. As a result, welfare gains from wealth taxes are about 1%–2% smaller than in the baseline but are still substantial (ranging between 6%–7.5% in the tax reform and optimal tax case).

7.2.1 Further Extensions

We have conducted some other experiments that we do not discuss to save space. We have considered alternative assumptions for borrowing constraints: (i) $\vartheta(\overline{z}_i) = \overline{\vartheta}$ for all i, (ii) relaxing the constraint so that the debt-to-GDP ratio is 3 (rather than 1.29), and (iii) allowing for unlimited borrowing at a higher interest rate than the rate for lending, and recalibrated the model. The upshot was that case (i) strengthened the results for wealth taxes, case (ii) yielded lower welfare and efficiency gains, but closer to the baseline than the $\vartheta = \infty$ case presented above, and case (iii) had a small quantitative effect and no substantive ones. We have also explored the effects of alternative values for some key parameters, such as ρ_z (0.5 instead of 0.1 in the baseline) and the risk aversion (2 instead of 4).

[45] For a given a, profits and, consequently, their present discounted value are rising in z_{ih}, driving a wedge between market value and book value.

8 Discussions and Conclusions

Wealth taxation has been a widely used policy tool for governments around the world, although its popularity has declined significantly in recent decades. One major reason for this decline has been the lack of a good rationale for its use, which itself reflects a dearth of academic research on its effects. In this paper, we presented a case for wealth taxes based on efficiency and distributional benefits and quantitatively evaluated its effects.

Under capital income taxation, 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, which expands the tax base and shifts the tax burden toward unproductive entrepreneurs. Furthermore, wealth taxes reduce the after-tax returns of high-productivity entrepreneurs less than low-productivity ones, which creates a behavioral savings response, which further shifts the wealth distribution toward the productive ones. Finally, the general equilibrium response of prices to wealth taxes can dampen the aggregate savings incentives, but its effect on reallocation is still in the same direction as the first two effects. The resulting reallocation increases aggregate productivity and output.

An optimal tax system with wealth taxation involves a positive tax on wealth (that ranges from about 2% to 3% a year across different specifications) and improves average welfare for every cohort starting with the tax reform. The higher revenue generated by the wealth tax allows the government to substantially reduce the tax rate on labor, raising after-tax wages relative to after-tax capital and interest income, making the consumption distribution more even. These findings lead us to conclude that wealth taxation can increase efficiency, grow the economy, and reduce inequality all at once.

An optimal tax system with capital income taxation involves a large subsidy to capital income and results in several drastic changes in the economy, including a much higher capital stock, output and consumption, coupled with much higher consumption inequality. Although average welfare in the OCIT steady state is higher compared to the benchmark US economy, both the population alive at the time of the reform and the cohorts that enter the economy soon after the reform experience large welfare losses.

Overall, our analysis lends strong 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, reduce consumption inequality, and improve welfare for large parts of the US population. In building our quantitative model, we tried hard to incorporate features that we believed were key for the first-order trade-offs

between wealth taxation and capital income taxation. As is the nature of quantitative analysis, we inevitably left out some other features of the real world that may be relevant for the overall effects of wealth and capital income taxation. We hope that these results would provide impetus for exploring the issues that we have not addressed in this paper. One such broad topic concerns some practical considerations that come up in implementing these policies. These questions are on our current and future research agenda.

References

Aiyagari, S. Rao, "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting," *Journal of Political Economy*, 1995, *103* (6), 1158–1175.

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

Asker, John, Joan Farre-Mensa, and Alexander Ljungqvist, "Comparing the Investment Behavior of Public and Private Firms," *NBER Working paper, No. 17394*, 2011.

Auerbach, Alan J, Jagadeesh Gokhale, Laurence J Kotlikoff, John Sabelhaus, and David N Weil, "The Annuitization of Americans' Resources: A Cohort Analysis," Working Paper Series 5089, National Bureau of Economic Research 1995.

Bach, Laurent, Laurent E. Calvet, and Paolo Sodini, "From Saving Comes Having? Disentangling the Impact of Saving on Wealth Inequality," Working Paper 2018.

Bell, Felicite C. and Michael L. Miller, "Life Tables for the United States Social Security Area: 1900-2100"," Actuarial Study 116, Office of the Actuary, Social Security Administration 2002.

Benhabib, Jess, Alberto Bisin, and Mi Luo, "Earnings Inequality and Other Determinants of Wealth Inequality," *American Economic Review*, May 2017, *107* (5), 593–97.

_ , _ , and Shenghao Zhu, "The Distribution of Wealth and Fiscal Policy in Economies With Finitely Lived Agents," *Econometrica*, 2011, *79* (1), 123–157.

_ , _ , and _ , "The Wealth Distribution in Bewley Models with Investment Risk," *Working Paper*, 2013.

_ , _ , and _ , "The Distribution of Wealth in the Blanchard-Yaari Model," *Macroeconomic Dynamics, Forthcoming*, 2014.

_ and _ , "Skewed Wealth Distributions: Theory and Empirics," *Journal of Economic Literature*, December 2018, *56* (4), 1261–91.

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

Boar, Corina and Matthew Knowles, "Entrepreneurship, Agency Frictions and Redistributive Capital Taxation," Working paper, New York University 2018.

Buera, F., J. Kaboski, and Yongseok Shin, "Finance and Development: A Tale of Two Sectors," *American Economic Review*, August 2011, pp. 1964–2002.

Cagetti, Marco and Mariacristina De Nardi, "Estate Taxation, Entrepreneurship, and Wealth," *American Economic Review*, 2009, *99* (1), 85–111.

Carroll, Christopher, Jiri Slacalek, Kiichi Tokuoka, and Matthew N. White, "The Distribution of Wealth and the Marginal Propensity to Consume," *Quantitative Economics*, 2017, *8*, 977–1020.

Chari, V. V. and Patrick J. Kehoe, "Optimal Fiscal and Monetary Policy," in John B.

Taylor and Michael Woodford, eds., *Handbook of Macroeconomics*, Vol. 1, Elsevier, 1999, chapter 26, pp. 1671–1745.

Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger, "Taxing Capital? Not a Bad Idea After All!," *American Economic Review*, 2009, *99* (1), 25–48.

De Nardi, Mariacristina and Fang Yang, "Wealth inequality, family background, and estate taxation," *Journal of Monetary Economics*, 2016, *77*, 130 – 145.

_ , Giulio Fella, and Gonzalo Paz Pardo, "The Implications of Richer Earnings Dynamics for Consumption, Wealth, and Welfare," Working paper 21917, National Bureau of Economic Research 2016.

Erosa, Andres and Martin Gervais, "Optimal Taxation in Life-Cycle Economies," *Journal of Economic Theory*, 2002, *105* (2), 338–369.

Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri, "Heterogeneity and Persistence in Returns to Wealth," Working paper, Stanford University 2016.

_ , _ , _ , and _ , "Heterogeneity in Returns to Wealth and the Measurement of Wealth Inequality," *American Economic Review*, 2016, *106* (5), 651–655.

_ , _ , _ , and _ , "Heterogeneity and Persistence in Returns to Wealth," Working paper, Stanford University 2018.

Federal Reserve Statistical Release, "Z.1 Financial Accounts of the United States: Flow of funds, balance sheets and integrated macroeconomic accounts," Technical Report, Board of Governors of the Federal Reserve System 2015Q1.

_ , "Z.1 Financial Accounts of the United States: Flow of funds, balance sheets and integrated macroeconomic accounts," Technical Report, Board of Governors of the Federal Reserve System 2015Q3.

Flodén, Martin, "The effectiveness of government debt and transfers as insurance," *Journal of Monetary Economics*, 2001, *48* (1), 81–108.

Gabaix, Xavier, "Zipf's Law for Cities: An Explanation," *The Quarterly Journal of Economics*, 1999, *114* (3), 739–767.

_ , "Power Laws in Economics and Finance," *Annual Review of Economics*, 2009, *1*, 255–93.

_ , "The Granular Origins of Aggregate Fluctuations," *Econometrica*, May 2011, *79-3*, 733–772.

_ , Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "The Dynamics of Inequality," *Econometrica*, 2016.

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

Golosov, Mikhail, Aleh Tsyvinski, and Ivan Werning, "New Dynamic Public Finance: A User's Guide," in "NBER Macroeconomic Annual 2006," MIT Press, 2006.

Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song, "What Do Data on Millions of U.S. Workers Say About Labor Income Risk?," Working Paper 20913, National Bureau of Economic Research 2015.

_ , _ , _ , and _ , "What Do Data on Millions of U.S. Workers Say About Labor Income

Dynamics?," Working Paper, University of Minnesota 2016.

Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L Violante, "Consumption and Labor Supply with Partial Insurance: An Analytical Framework," *American Economic Review*, 2014.

Hendricks, Lutz, "Bequests and Retirement Wealth in the United States," Working Paper, Arizona State University 2001.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, November 2009, *124* (4), 1403–1448.

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

Hubmer, Joachim, Per Krusell, and Anthony A. Smith, "The Historical Evolution of the Wealth Distribution: A Quantitative-Theoretic Investigation," Working paper, Yale University 2017.

Imrohoroglu, Selahattin, "A Quantitative Analysis of Capital Income Taxation," *International Economic Review*, 1998, *39*, 307–328.

Itskhoki, Oleg and Benjamin Moll, "Optimal Development Policies with Financial Frictions," *Econometrica*, 2019.

Jones, Charles I. and Jihee Kim, "A Schumpeterian Model of Top Income Inequality," *Journal of Political Economy*, 2018, *126* (5), 1785–1826.

Judd, Kenneth L., "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, 1985, *28* (1), 59–83.

Kitao, Sagiri, "Labor-Dependent Capital Income Taxation," *Journal of Monetary Economics*, 2010, *57* (8), 959–974.

Li, Huiyu, "Leverage and Productivity," 2016.

Lucas, Robert E., "On the Size Distribution of Business Firms," *The Bell Journal of Economics*, 1978, *9* (2), 508–523.

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

Luttmer, Erzo G.J., "Selection, Growth, and the Size Distribution of Firms," *Quarterly Journal of Economics*, 2007, *122* (3), 1103–1144.

_ , "On the Mechanics of Firm Growth," *Quarterly Journal of Economics*, 2011, *78* (3), 1942–1068.

McDaniel, Cara, "Average Tax Rates on Consumption, Investment, Labor and Capital in the OECD: 1950-2003," Arizona State University mimeo 2007.

Midrigan, Virgiliu and Daniel Yi Xu, "Finance and Misallocation: Evidence from Plant-Level Data," *American Economic Review*, February 2014, *104* (2), 422–458.

Moll, Benjamin, "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?," *American Economic Review*, October 2014, *104* (10), 3186–3221.

OECD, "The Role and Design of Net Wealth Taxes in the OECD," *OECD Tax Policy Studies*, 2018, (26).

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

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

Restuccia, Diego and Richard Rogerson, "Policy distortions and aggregate productivity with heterogeneous establishments," *Review of Economic Dynamics*, 2008, *11* (4), 707 – 720.

Shourideh, Ali, "Optimal Taxation of Wealthy Individuals," Technical Report, Wharton School at University of Pennsylvania 2013.

Smith, Matthew, Danny Yagan, Owen Zidar, and Eric Zwick, "Capitalists in the Twenty-First Century," Working Paper, UC Berkeley 2017.

Solon, Gary, "Intergenerational Mobility in the Labor Market," in Orley Ashenfelter and David Card, eds., *Handbook of Labor Economics, Vol 3A*, Vol. 3A, North-Holland, 1999, pp. 1761–1800.

Stachurski, John and Alexis Akira Toda, "An Impossibility Theorem for Wealth in Heterogeneous-agent Models with Limited Heterogeneity," *Journal of Economic Theory*, 2019, *182*, 1–24.

Straub, Ludwig and Iván Werning, "Positive Long-Run Capital Taxation: Chamley-Judd Revisited," Working paper, MIT 2018. Test.

Vermeulen, Philip, "How Fat is the Top tail of the Wealth Distribution?," *Reivew of Income and Wealth*, 2016.

Wolff, Edward N., "Changes in Household Wealth in the 1980s and 1990s in the U.S.," in Edward N. Wolff, ed., *International Perspectives on Household Wealth*, Eduaward Elgar Publishing, 2006, chapter 4, pp. 107,150.

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:

$$y^R(\kappa, e) = \Phi(\kappa, e)\,\overline{y},$$

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

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

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 $e_R = e$:

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

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 such 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:

$$SSP \coloneqq \int_{h\geq R,a,\mathbf{S}} y^R(\kappa, e)\, d\Gamma(h, a, \mathbf{S}). \tag{14}$$

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 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* $L = \int_{h,a,\mathbf{S}} (y_h\,(\kappa,e)\,\ell_h(a,\mathbf{S}))\, d\Gamma\,(h,a,\mathbf{S})$*, where* $\log y_h = \kappa + g(h) + e$*.*

4. *The bond market clears:*

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

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

$$\begin{aligned} G + SSP = {} & \tau_k \int_{h,a,\mathbf{S}} (ra + \pi\,(a,z))\, d\,\Gamma\,(h,a,\mathbf{S}) \\ & + \tau_a \int_{h,a,\mathbf{S}} ((1+r)a + \pi\,(a,z))\, d\,\Gamma\,(h,a,\mathbf{S}) \\ & + \tau_\ell \int_{h<R,a,\mathbf{S}} (\overline{w} w_h\,(\kappa,e)\,\ell_h(a,\mathbf{S}))\, d\,\Gamma\,(h,a,\mathbf{S}) \\ & + \tau_c \int_{h,a,\mathbf{S}} c_h(a,\mathbf{S}) d\,\Gamma\,(h,a,\mathbf{S})\,, \end{aligned} \tag{15}$$

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 eq. (14).*

B Additional Tables

B.1 Dispersion of Individual Rates of Return

The heterogeneity in the rates of return is an important mechanism in the model for generating a wealth distribution that is consistent with the data in numerous dimensions. Therefore, it is of interest to compare the dispersion in the rates of return in the model and the data. Even though the empirical evidence is scarce, Fagereng et al. (2016a) report the rates of return in the Norwegian data. Rather encouragingly, the dispersion observed in the model matches well with the facts reported in Fagereng et al. (2016a). Let r^p_{ih} denote the annual return on the investment portfolio of individual i at age h:

$$r^p_{ih} = \frac{ra_{ih} + \pi\,(a_{ih}, z_{ih})}{a_{ih}},$$

TABLE B.1 – Forbes Self-made Index

	Description	Fraction 2015
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 and 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**
	Our definition of "Self-made": groups 8 to 10	**54.25**

TABLE B.2 – Distribution of *Lifetime* Returns around the Median

	p10	p25	P50	p75	p90	p99	p99.9
Norwegian Data	–2.4%	–1.3%	0%	2.1%	4.1%	9.7%	19.9%
Working life	–3.7%	–2.9%	0%	4.5%	7.8%	13.3%	19.4%
Ages 25-65	–3.4%	–2.9%	0%	4.7%	8.0%	13.6%	19.9%
Ages 20-24	–10.0%	–5.6%	0%	4.9%	13.3%	32.7%	55.4%

Note: Lifetime returns are weighted by the individual's wealth at each age. All numbers are before tax. All numbers are presented as differences from the median. The Norwegian data are taken from Fagereng et al. (2016a), Table 4, which reports percentiles of fixed effects of individual returns to wealth. The data for P99.9 were kindly provided to us by the authors.

where π is the profit as defined in equation (9). The lifetime rate of return, denoted r_i^p, is computed as a weighted average of annual returns:

$$r_i^p = \sum_{h=1}^{R} \varphi_{ih} r_{ih}^p,$$

where $\varphi_{ih} = a_{ih}/\Sigma_{h=1}^{R} a_{ih}$ is the relative wealth share at age h.

Table B.2 reports various percentiles of r_i^p in the data and the model, relative to the median return in the data and the model, respectively. The lifetime rate of return at the 99.9th percentile, relative to the median return, is around 20% in both the model and the data. The lifetime returns at other percentiles above the median, however, are slightly higher than the returns observed in the data; for example, the lifetime return at the 99th percentile is around 10% in the data and around 13–14% in the model. As expected, the rates of return are substantially higher at high percentiles when individuals

TABLE B.3 – Cross-Sectional Dispersion in Annual Rates of Return: Benchmark vs. RN Reform Economy.

	P10	P50	P90	P95	P99
	Before-tax				
Benchmark	2.00	2.00	17.28	22.35	42.36
Wealth Tax	1.74	1.74	14.62	19.04	36.91
	After-tax				
Benchmark	1.50	1.50	12.96	16.76	31.77
Wealth Tax	0.59	0.59	13.32	17.69	35.35

Note: Each cell reports the annual rate of return in percentages.

TABLE B.4 – Welfare Change: Optimal Wealth Taxes with Exemption Threshold

	Productivity group (Percentile)					
Age	0–40	40–80	80–90	90–99	99–99.9	99.9+
20	9.9	9.8	10.3	11.4	13.4	14.6
21–34	9.1	8.0	7.4	7.2	6.6	5.6
35–49	6.7	4.9	3.6	1.9	–1.6	–4.9
50–64	2.7	1.5	0.6	-0.8	–3.9	–6.5
65+	–0.4	–0.7	–1.0	–1.6	–3.2	–4.6

Note: This table reports the welfare change distribution for the optimal wealth tax with exemption to complement Table XII.

are young. As productive individuals experience significant growth early in the life cycle, between the ages of 20 and 24, they experience rates of return as high as 55% at the 99th percentile. Overall, the distribution of lifetime rates of return in the model is consistent with the distribution observed in the Norwegian data.[46]

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

[46]The overall message remains unchanged if we instead compute the lifetime rates of return percentiles in the model relative to the median return in the Norwegian data.

TABLE B.5 – Fraction with Positive Welfare Gain by Age Group and Entrepreneurial Ability

(A) Optimal Capital Income Taxes

	Productivity group (Percentile)					
Age	0–40	40–80	80–90	90–99	99–99.9	99.9+
20	95.4	98.6	99.3	99.6	99.8	99.9
21–34	96.3	97.7	97.7	97.3	96.0	94.6
35–49	91.7	92.8	91.1	87.8	80.3	73.4
50–64	74.2	76.2	73.8	69.4	60.3	52.8
65+	13.8	18.6	18.7	18.2	16.6	15.0

(B) Optimal Wealth Taxes

	Productivity group (Percentile)					
Age	0-40	40-80	80-90	90-99	99-99.9	99.9+
20	94.5	93.1	93.3	94.6	95.8	96.1
21–34	95.7	92.6	90.5	88.8	84.2	78.2
35–49	91.3	82.8	76.5	68.2	53.6	43.6
50–64	72.6	62.9	56.1	49.4	39.8	33.8
65+	2.1	2.3	1.8	1.4	0.9	0.7

(C) Optimal Wealth Taxes with Exemption Threshold

	Productivity group (Percentile)					
Age	0-40	40-80	80-90	90-99	99-99.9	99.9+
20	94.5	93.1	93.3	94.6	95.8	95.9
21–34	95.6	92.4	90.4	88.5	83.8	77.6
35–49	91.1	82.4	76.0	67.8	53.2	43.3
50–64	76.4	66.7	59.6	52.5	42.3	35.8
65+	75.9	68.6	63.7	57.9	48.7	42.1

Note: Each entry reports the share of individuals in a given age and entrepreneurial productivity group (ranked based on the permanent component of entrepreneurial productivity $\overline{z}$) that would experience a positive welfare gain (CE_1) from the corresponding optimal tax experiment. The shares are computed with respect to the benchmark distribution.

TABLE B.6 – Total Tax Revenue and Tax Revenue by Source

	Benchmark	Tax Reform		OWT	OCIT
		RN	BB		
Total tax rev./GDP	0.25	0.23	0.24	0.24	0.20
Rev. capital taxes/Total tax rev.	0.25	0.17	0.22	0.41	–0.34
Rev. from labor/Total tax rev.	0.54	0.59	0.56	0.36	1.08

TABLE B.7 – Welfare Change: Optimal Tax Reforms with Transition

(A) Optimal Capital Income Taxes

	Fixing τ_k						Fixing τ_l					
	Productivity group (Percentile)						*Productivity group (Percentile)*					
	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	-6.1	-2.7	0.5	5.4	13.0	17.5	-5.6	-2.6	0.3	4.5	11.4	15.4
21–34	-5.6	-0.8	3.6	10.0	20.6	27.4	-5.1	-0.8	3.0	8.6	18.0	24.1
35–49	-5.0	-0.4	2.9	7.0	12.7	16.0	-4.6	-0.4	2.5	6.1	11.1	14.2
50–64	-3.2	0.0	2.0	4.2	7.2	9.0	-2.9	0.0	1.7	3.7	6.4	8.0
65+	0.7	1.9	2.6	3.6	5.1	6.1	0.6	1.7	2.3	3.2	4.5	5.4

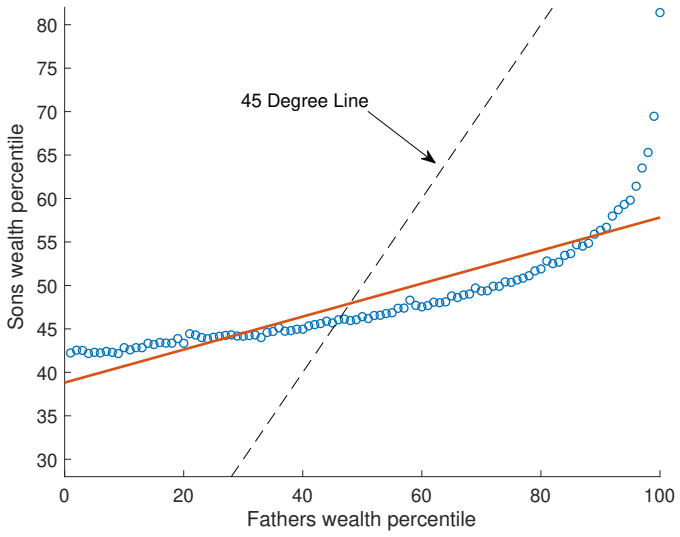
(B) Optimal Wealth Taxes

	Fixing τ_a						Fixing τ_l					
	Productivity group (Percentile)						*Productivity group (Percentile)*					
	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	6.4	6.6	7.4	9.0	11.8	13.5	7.1	7.1	7.9	9.4	12.1	13.8
21–34	5.8	5.2	5.3	5.9	6.8	6.7	6.4	5.6	5.5	5.9	6.4	6.1
35–49	3.9	2.8	2.1	1.3	-1.0	-3.3	4.3	2.9	2.1	1.0	-1.7	-4.4
50–64	1.2	0.4	-0.1	-1.1	-3.3	-5.3	1.3	0.4	-0.3	-1.4	-4.0	-6.2
65+	-0.5	-0.7	-0.9	-1.4	-2.6	-3.8	-0.6	-0.9	-1.1	-1.7	-3.1	-4.4

C Additional Figures

FIGURE C.1 – Intergenerational Rank-Rank Correlation in Wealth

(A) Baseline



(B) Fagereng et al. (2016a)

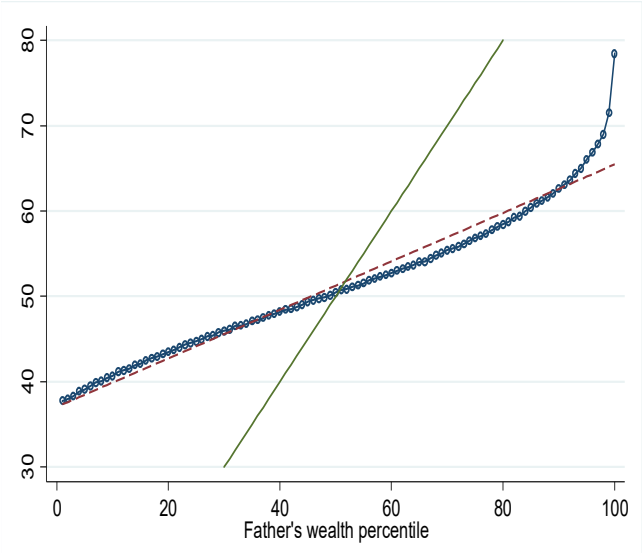
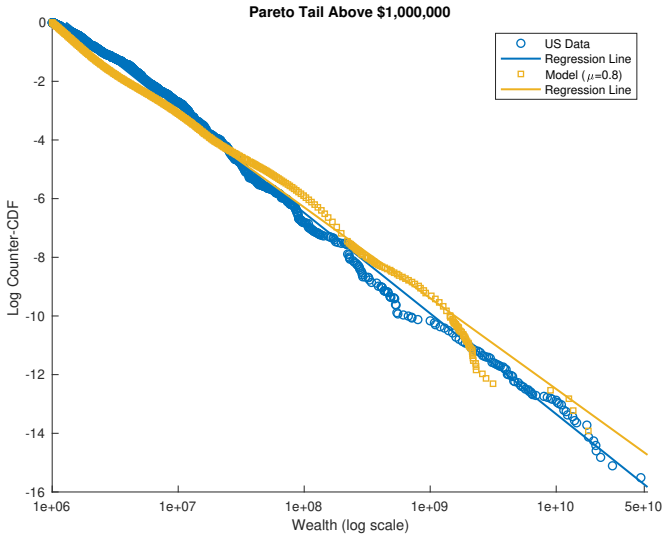
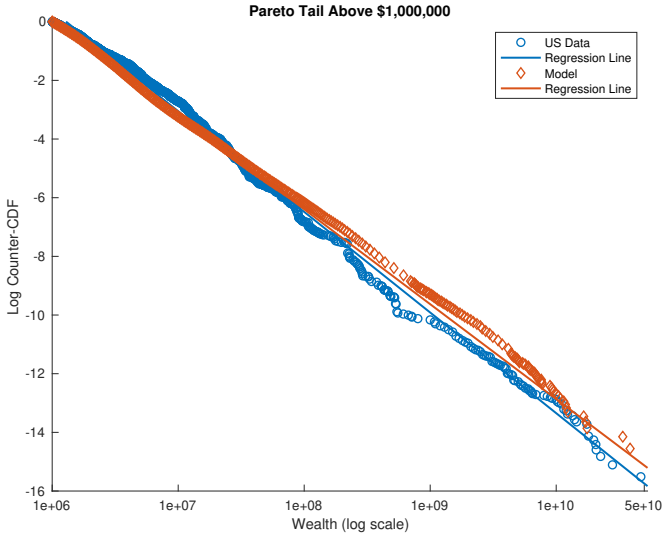


FIGURE C.2

(A) Stronger Diminishing Returns in Entrepreneurial Production, $\mu = 0.8$



(B) Measuring Wealth in Present Values



D Misallocation in the Benchmark Economy

Our benchmark economy is distorted because of the existence of financial frictions in the form of borrowing constraints, and we can measure the effects of these distortions on aggregate TFP and output and compare them with those obtained in other studies. A large and growing literature frames the discussion on misallocation in terms of various wedges, such as capital, labor, and output wedges. The analysis in Hsieh and Klenow (2009) is particularly useful since, in a similar model environment, they study the degree of misallocation and its effect on TFP in manufacturing in China, India, and the United States. Hsieh and Klenow use detailed firm-level data from the US Census of Manufacturers (1977, 1082, 1987, 1992, and 1997) and find that the TFP gains from removing all distortions (wedges), which equalizes the "Revenue Productivity" (TFPR) within each industry are 36% in 1977, 31% in 1987, and 43% in 1997.

We will follow the approach in Hsieh and Klenow (2009) and compute the same measures of misallocation for the united States as in their analysis. It is useful to briefly describe their approach as it applies to our framework. The final goods producer behaves competitively and uses an aggregated good, Q, and labor, L, in the production of the final good,

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

where Q aggregates the intermediate goods x_i in the following way:

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

Each intermediate goods producer i produces a differentiated intermediate good using the production function $x_i = z_i k_i$, where z_i is individual i's entrepreneurial ability and k_i is the amount of capital.

Instead of modeling and capturing the effect of a particular distortion, or distortions, the approach of Hsieh and Klenow and the related misallocation literature is to 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 in the economy (or in a particular industry). This is based on the insight that absent any distortions, the marginal revenue products of capital and labor have to be equalized across all firms.[47]

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

TFP in the Q sector. We will first focus on the Q sector, the sector that produces the composite intermediate input Q by aggregating all the intermediate goods x_i. Under this alternative capital-wedge approach, the problem of each intermediate goods producer is

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

where τ_i^k is a firm-specific capital wedge. The only input in the production function of the intermediate goods producer is capital, and as a result only one wedge can be identified in the analysis. We choose to specify that wedge to be the capital wedge, but in principle it should be understood as capturing the effect of an output wedge.

The revenue TFP in sector Q for each firm i is

$$TFPR_{Q,i} \equiv \frac{p\left(x_i\right) x_i}{k_i} = \frac{1}{\mu}\left(1+\tau_i^k\right)\left(R+\delta\right).$$

The aggregate TFP in sector Q can be expressed as

$$TFP_Q = \left(\int_i \left(z_i \frac{\overline{TFPR_Q}}{TFPR_{Q,i}}\right)^{\frac{\mu}{1-\mu}} di\right)^{\frac{1-\mu}{\mu}}, \tag{16}$$

where the average $TFPR_Q$ is given by

$$\overline{TFPR_Q} = \left(\int \frac{1}{TFPR_{Q,i}} \frac{p\left(x_i\right) x_i}{p_q Q} di\right)^{-1}.$$

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

$$TFP_Q^* = \left(\int_i \left(z_i\right)^{\frac{\mu}{1-\mu}} di\right)^{\frac{1-\mu}{\mu}} \equiv \overline{z}.$$

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 borrowing constraints:

$$\Omega_Q = \frac{TFP_Q^*}{TFP_Q} = \left(\int_i \left(\frac{\overline{z}}{z_i} \frac{TFPR_{Q,i}}{\overline{TFPR_Q}}\right)^{\frac{\mu}{1-\mu}} di\right)^{\frac{1-\mu}{\mu}}.$$

Table D.8 reports Ω_Q for various economies: the TFP in the Q sector in the non-distorted economy is 58% higher than in the benchmark economy, 51% higher than in the economy with a wealth tax, 54% higher than in the economy with a consumption tax, 49% higher

TABLE D.8 – Hsieh and Klenow (2009) Efficiency Measure - Benchmark Model

	Benchmark	Tax Reform (τ_a)	Opt. Taxes (τ_k)	Opt. Taxes (τ_a)
TFP_Q	1.001	1.047	1.064	1.074
$\frac{TFP_Q^*}{TFP_Q}$	1.582	1.514	1.489	1.475
Mean TFPR	0.145	0.131	0.106	0.145
StD TFPR	0.054	0.048	0.039	0.053
p99.9	0.68	0.61	0.5	0.66
p99	0.35	0.32	0.27	0.35
p90	0.19	0.17	0.14	0.19
p50	0.14	0.12	0.1	0.14
p10	0.1	0.09	0.07	0.1

than in the economy with an OCIT, and 47% higher than in the economy with an OWT.

Wealth taxes give the higher TFP gains, allowing for better allocation of capital across firms, even without eliminating the borrowing constraints. The tax reform experiment implies a TFP gain of 4.6%, and optimal wealth taxes give a TFP gain of 7.3% with respect to our benchmark economy.

This can also be seen in the dispersion of TFPR in the different models. Recall that absent any constraints on the firms, TFPR would be equated across all of them, so there is higher misallocation in the economy the higher the dispersion of TFPR across firms. Table D.8 reports the standard deviation of TFPR and some of its percentiles.

Comparison with the Hsieh and Klenow (2009) results for the United States

In order to compare these results with the results reported in Hsieh and Klenow (2009) for the United States, we need to note that the improvement in aggregate output, Ω_Y, as a result of eliminating the capital wedges in the economy can be expressed as

$$\Omega_Y = \frac{Y^*}{Y} = \left(\frac{TFP_Q^*}{TFP_Q}\right)^{\alpha} \left(\frac{K^*}{K}\right)^{\alpha} \left(\frac{L^*}{L}\right)^{1-\alpha}.$$

Since the model with capital wedges is static, the effect of the removal of the capital wedges on aggregate capital, K, and labor supply, L, cannot be taken into account. The analysis in Hsieh and Klenow (2009) measures the improvement in total output as a result of an improvement in TFP in all industries. In our model, this corresponds to the improvement in TFP in the Q sector. Therefore, removing the capital wedges would

increase total output, through its effect on TFP in the Q sector, by 20%.[48]

Two things are important to point out. First, the magnitude of the misallocation in our benchmark economy is substantial, although a bit lower than the one measured in Hsieh and Klenow (2009) using micro data from manufacturing firms: 36% in 1977, 31% in 1987, and 43% in 1997. However, it is in line with the level reported in ongoing research by Bils et al. (2017), who take into account measurement error in micro data and find gains from removing distortions for the United States in the range of 20%. In any case, it is worth noting several differences between our framework and that of Hsieh and Klenow (2009). Our benchmark economy is parameterized based on moments from the entire economy, not just the manufacturing sector. Second, our benchmark model is a dynamic model, and any changes in the financial frictions will affect aggregate capital accumulation and aggregate labor supply. The misallocation calculations above do not take those changes into account. It is clear, however, that eliminating the financial friction would increase the aggregate capital stock K and lead to larger increases in total output than measured above. The effect on aggregate labor supply is less obvious.

E 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, here we show 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 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.

E.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 (4), so the entrepreneurs' problem (9) becomes:

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

[48] Note that $\tilde{\Omega}_Y = \Omega_Q^{\alpha} = \Omega_Q^{0.40} = 1.20$.

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

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

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, i.e., leisure is low, the marginal benefit (MB) of extra work is lower. Switching to a wealth tax increases MB since 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 $\ell \leq 0.43$ for the capital-constrained entrepreneur and $\ell \leq 0.88$ 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:

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

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

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

E.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:

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

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:

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

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

1. $-\left(r+\delta\right)\vartheta\left(z\right)a+ra<0$,
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 $\left(\tau_k-\tau_a\right)R\left(z\vartheta\left(z\right)a\ell\right)^{\mu}>\Delta C$. 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

$$\begin{aligned}\frac{1-\ell}{\ell} &\geq 1.3\\ \frac{1}{\ell} &\geq 2.3\\ \ell &\leq 0.43\end{aligned}$$

Of course, this is a sufficient condition. So, if the entrepreneur were not working too much initially (i.e. $\ell\leq 0.43$), 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

$$\ell \leq 0.28.$$

B. Capital-unconstrained entrepreneur

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

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

and 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 increases 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.[49] Consider what happens to the left-hand and the right-hand sides for a given ℓ if we switch from a capital income tax to a wealth tax:

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

and

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

Note that if the $\Delta RHS < 0$, the switch to a wealth tax definitely increases entrepreneurial hours. So, we will focus on the case, in which $\Delta RHS > 0$. 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 $\Delta LHS > \Delta RHS$ is that

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

[49] 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 gives

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

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

which implies

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

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