

USE IT OR LOSE IT: EFFICIENCY AND REDISTRIBUTIONAL EFFECTS OF WEALTH TAXATION*

FATIH GUVENEN
GUEORGUI KAMBOUROV
BURHAN KURUSCU
SERGIO OCAMPO
DAPHNE CHEN

How does wealth taxation differ from capital income taxation? When the return on investment is equal across individuals, a well-known result is that the two tax systems are equivalent. Motivated by recent empirical evidence documenting persistent return heterogeneity, we revisit this question. With heterogeneity, the two tax systems typically have opposite implications for 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 U.S. data, replacing the capital income tax with a wealth tax in a revenue-neutral way delivers a significantly higher average welfare. Turning to optimal taxation, the optimal wealth tax (OWT) is positive and yields large welfare gains by raising efficiency and lowering inequality. In contrast, the optimal capital income tax (OKIT) is negative—a subsidy—and delivers lower welfare gains than OWT, owing to the welfare losses from higher inequality. Furthermore, when the transition path is considered, the gains from OKIT turn into significant welfare losses for existing cohorts, whereas OWT continues to deliver robust welfare gains.

* For helpful comments, we thank Mark Aguiar, Adrien Auclert, Jess Benhabib, Alberto Bisin, Richard Blundell, V. V. Chari, Mariacristina De Nardi, Andres Erosa, Priscilla Fialho, Xavier Gabaix, Mike Golosov, Joao Gomes, Pierre-Olivier Gourinchas, Luigi Guiso, Bob Hall, Hugo Hopenhayn, Nobu Kiyotaki, Henrik Kleven, Dirk Krueger, Per Krusell, Guido Menzio, Claudio Michelacci, Ben Moll, Juan Pablo Nicolini, Thomas Piketty, Victor Ríos-Rull, Martin Schneider, Rob Shimer, Stefanie Stantcheva, Aleh Tsyvinski, Maxim Troshkin, Gabriel Zucman, and seminar and conference participants at various institutions. We thank the editor and five anonymous referees for comments that substantially improved the article. For funding, Kuruscu thanks the Marie Skłodowska-Curie programme under Horizon 2020; Kambourov and Kuruscu thank the Social Sciences and Humanities Research Council; and Ocampo thanks the ERC starting grant ValuRed (no. 804104). Daphne Chen made valuable contributions in the early stages of the project.

© The Author(s) 2023. Published by Oxford University Press on behalf of the President and Fellows of Harvard College. All rights reserved. For Permissions, please email: journals.permissions@oup.com
The Quarterly Journal of Economics (2023), 1–60. <https://doi.org/10.1093/qje/qjac047>.
Advance Access publication on January 9, 2023.

These results suggest that moderate wealth taxation may be a more appealing alternative than capital income taxation, which can be significantly more distorting under return heterogeneity than under the equal-returns assumption. *JEL Codes:* E62, H21, H24.

I. INTRODUCTION

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

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

whereas under the wealth tax, it is

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

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

To see some of the implications for capital taxation, consider two entrepreneurs who start out with the same wealth level—say, \$1,000 each—but earn different returns—say, $r_1 = 0\%$ and $r_2 = 20\%$. Under capital income taxation, the unproductive (first) entrepreneur will escape taxation because he generates no income, and the tax burden will fall entirely on the more productive (second) entrepreneur. Under wealth taxation, on the other hand, both entrepreneurs will pay the same amount of tax on wealth regardless of their productivity; as a result, wealth taxation expands the tax base, shifts the tax burden toward the unproductive entrepreneur, and reduces (potential) tax distortions on the productive entrepreneur. To the extent that these differences in

productivity are persistent, a wealth tax will gradually prune the wealth of idle entrepreneurs and boost that of successful ones, leading to a more efficient allocation of aggregate capital,¹ in turn raising productivity and output. In this sense, wealth taxation has a “use-it-or-lose-it” effect absent from capital income taxation. We expand on this example in [Section II](#).

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

There are three more considerations that motivate us to study capital taxation under return heterogeneity. First, a growing number of empirical studies cast strong doubt on the assumption of homogeneous returns. Using administrative panel data sets that track millions of people over time, these studies document large and persistent differences in individual returns, even after adjusting for risk and other factors (e.g., [Bach, Calvet, and Sodini 2020](#); [Fagereng et al. 2020](#); [Smith, Zidar, and Zwick 2023](#), among others).² These new pieces of evidence make studying the tax implications of return heterogeneity more than a theoretical curiosity.

Second, the literature on power law models shows that return heterogeneity is a powerful modeling tool that can generate key features of wealth inequality that proved challenging to explain through other mechanisms.³ This is an important consideration for studying capital taxation: because of the extreme concentration of wealth in the United States and in many other countries ([Vermeulen 2018](#)), the bulk of the capital tax burden falls on a small fraction of wealthy households. In addition, a large fraction of these very wealthy households are self-made rather than inheritors, which is an important determinant of the trade-offs

1. Throughout the article we use the terms *capital* and *wealth* interchangeably.

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

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

they face and how they respond to capital taxation. Thus, for a sound quantitative analysis of capital taxation, we believe that it is important for a model to reproduce key features of the top-end wealth distribution, such as the thick Pareto tail and the very rapid wealth accumulation of the super wealthy.⁴

Third, studying wealth taxation also has a practical motivation: it is a policy tool that has long been used by governments around the world.⁵ Although its popularity was in decline until recently, the past few years have seen a revival of interest in many countries. In light of this reality, studying the implications of wealth taxation and how it differs from capital income taxation is an important step toward providing better guidance for policy makers.

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

Each individual/entrepreneur produces a differentiated intermediate good using a proprietary technology with individual-specific productivity. These intermediates are combined in a Dixit-Stiglitz aggregator by a final-goods producing firm, which pins down (together with the collateral constraint) each entrepreneur's production scale and profits. In this setup, every entrepreneur earns a monopoly profit. Individuals face idiosyncratic labor income risk, mortality risk, and various intergenerational links, although plausible variations in these details do not change the

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

5. Until about 2012, France, Germany, Spain, Italy, the Netherlands, and the Nordic countries, among others, had wealth taxation. See [Organisation for Economic Co-operation and Development \(2018\)](#) for a review of the use of wealth taxes.

substantive conclusions. The calibrated model is consistent with key features of the U.S. data, including top-end wealth inequality (including the Pareto tail), the amount of return heterogeneity, statistics on entrepreneurs, and the magnitude of borrowing by U.S. businesses, among other features.

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

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

The analogous optimal capital income taxation (OKIT) exercise delivers an optimal (linear) tax rate of -13.6% , implying a nontrivial subsidy to capital income. This finding may seem surprising in light of previous results in the literature that found a high positive tax rate (of about 35%), using Aiyagari-style models with many similarities with ours (e.g., [Conesa, Kitao, and Krueger 2009](#)). The main difference is return heterogeneity: shutting down return heterogeneity restores the high positive tax rate found in previous work. To understand why this happens, note that in Aiyagari-style models, the wealthy are workers who have been

productive in the past, but they are not any better at investing this wealth than others, so the efficiency losses from capital income taxation are not especially large. In contrast, with return heterogeneity, entrepreneurs who earn high capital income (per unit of wealth) are precisely the productive ones today, which makes taxing capital income much more distorting and the efficiency losses especially large, making a subsidy optimal. A key takeaway from these results is that beyond its implications for wealth taxation, accounting for return heterogeneity can matter greatly for studying capital income taxation.

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

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

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

readers about drawing direct conclusions regarding those tax proposals.

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

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

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

I.A. Related Literature

Although the use-it-or-lose-it feature of wealth taxes has been noted by a few authors, we are not aware of prior academic analyses of its effects.⁶ Maurice Allais was among the best-known

6. This article was first presented at the 2014 NBER Summer Institute and has been widely presented at seminars and conferences since then. It predates the

proponents of wealth taxes, and he discussed the use-it-or-lose-it rationale in [Allais \(1977\)](#). More recently, [Piketty \(2014\)](#) revived the debate on wealth taxation and proposed using a combination of capital income and wealth taxes to balance these efficiency and inequality trade-offs. Piketty focused mostly on equity considerations but also described the use-it-or-lose-it mechanism without providing a formal analysis.

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

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

recent public debate on wealth taxation that rose to prominence during the 2020 U.S. presidential election campaign.

productivity and wages. In the long run, optimal policy reverses and becomes proworker. In our framework, wealth taxation plays a similar role.

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

II. AN ILLUSTRATIVE EXAMPLE

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

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

Now consider a wealth tax imposed on beginning-of-period wealth (right panel). First, this doubles the tax base, which now includes Fredo's wealth. The tax rate on wealth is $\tau_a = \frac{5}{200} = 2.5\%$. More importantly, now half of the tax bill is paid by Fredo, and Michael's tax bill is cut by half. As a result, Fredo's after-tax return is now -2.5% , and Michael's return is 17.5%, so wealth taxation does not compress the distribution of returns or wealth across

TABLE I
SUMMARY OF THE ILLUSTRATIVE EXAMPLE

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

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

investors as much as capital income taxation does (and does not compress at all in this specific example). Notice that wealth dispersion in this example is between wealthy investors and not across the broader population. As we will see in the quantitative analysis, wealth taxation can deliver (large) distributional welfare gains (by raising wages) in addition to its efficiency benefits.

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

Two final remarks are in order. First, if this one-period example were repeated for many periods, Michael would eventually hold all the wealth, so misallocation would vanish in the long run. But this prediction would be true only if return differences were completely permanent, which is not the case in real life: the

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

fortunes of entrepreneurs vary over time and from one generation to the next, so capital misallocation will persist even in the long run. These features will be incorporated into the dynamic model we describe next. Second, we assumed that the wealth tax was imposed on the book value of assets, not their market value; the distinction between the two is critical. This is because the latter incorporates the productivity of the entrepreneur through its effect on future earnings, so taxing the market value weakens the use-it-or-lose-it mechanism by raising the tax burden of more productive entrepreneurs.

III. FULL OLG MODEL

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

III.A. Individuals

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

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

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

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

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

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

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

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

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

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

an indicator function:

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

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

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

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

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

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

$$(5) \quad e_{ih} = \rho_e e_{i,h-1} + v_{ih},$$

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

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

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

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

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

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

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

III.B. Production Technology

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

$$(8) \quad x_{ih} = z_{ih} k_{ih},$$

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

$$(9) \quad Y = Q^\alpha L^{1-\alpha},$$

where Q is the CES composite of intermediate inputs:

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

To distinguish Q from the unadjusted capital stock, $K = \int k_{i,h(i)} di$, we refer to the former as the quality-adjusted capital

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

stock, since its level depends on the allocation of capital across entrepreneurs. Total factor productivity (TFP) in the intermediate-goods sector can be written as

$$(11) \quad \text{TFP}_Q = \frac{Q}{K},$$

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

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

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

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

$$(13) \quad p(x) = \alpha Q^{\alpha-\mu} L^{1-\alpha} x^{\mu-1} \quad \bar{w} = (1 - \alpha) Q^\alpha L^{-\alpha}.$$

III.C. Markets and the Government

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

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

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

$$(14) \quad k_{ih} \leq \vartheta(\bar{z}_i) \times a_{ih},$$

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

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

2. *Government and the Tax/Transfer Systems.* In the benchmark "U.S. economy," the government imposes flat taxes at rates τ_k on capital income, τ_ℓ on labor income, τ_c on consumption, and τ_b on bequests. We alternatively refer to this case as the capital income tax economy. In the alternative wealth tax economy, the government does not tax capital income (set $\tau_k = 0$) but imposes a flat-rate tax, τ_a , on beginning-of-period wealth, a_{ih} . We consider formulations with nonlinear taxes later. We denote a tax system as $\mathcal{T} \equiv (\tau_{\text{cap}}, \tau_\ell)$, where $\text{cap} \in (k, a)$.¹³ The government runs a balanced budget and uses the tax revenues to fund Social Security

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

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

pension payments to retirees and an exogenous, fixed level of government spending, G , which does not enter individuals' utility function (see [Online Appendix A.2](#) for the equations).

The Social Security pension of a retiree ($t \geq R$) is determined according to the formula

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

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

III.D. Individuals' Decision Problem

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

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

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

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

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

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

with the associated maximized profit function

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

2. Individual's Dynamic Programming Problem. Let \mathcal{W} denote the postproduction, after-tax wealth of the individual, which includes current-period profits plus interest income from the bond market:

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

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

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

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

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

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

IV. QUANTITATIVE ANALYSIS

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

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

IV.A. Model Parameterization

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

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

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

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

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

TABLE II
PARAMETERS FOR THE BENCHMARK MODEL

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

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

the same curvature parameter as consumption:

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

as in [De Nardi and Yang \(2016\)](#). In this formulation, χ measures the importance of the bequest motive, and b_0 measures the degree to which bequests are luxury goods. When $b_0 > 0$ the marginal utility of bequests is bounded at $b = 0$ allowing some people to leave no bequests, consistent with the data. Following [De Nardi and Yang \(2016\)](#), we choose χ and b_0 to match a bequest-to-wealth

TABLE III
TARGETED AND UNTARGETED MOMENTS: MODEL VERSUS DATA

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

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

ratio of 1.18 and the 90th percentile of the bequest distribution scaled by income, which is 4.31.¹⁵

We set $\alpha = 0.4$, which implies a labor share of 0.60, and we set the depreciation rate to 5%. The curvature parameter of the CES aggregator, μ , is set to 0.9, which corresponds to a 10% markup over marginal cost. A higher μ implies less diminishing marginal returns to capital, which makes it easier to generate high

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

inequality. The reverse happens as μ goes down, although the effects are mild down to a value of $\mu = 0.75$ or so. Beyond that point, the diminishing returns in entrepreneurial production become so strong that matching the right tail of the wealth distribution becomes impossible (and is no longer Pareto). We discuss the results for $\mu = 0.8$ in [Section VII](#).

3. *Tax System.* The current U.S. tax system is modeled as a quadruplet of (flat) tax rates: $(\tau_k, \tau_\ell, \tau_c, \tau_b)$. We set $\tau_k = 25\%$, $\tau_\ell = 22.4\%$, and $\tau_c = 7.5\%$, based on [McDaniel \(2007\)](#)'s calculations for the U.S. economy between 2000 and 2003.¹⁶ In the baseline analysis, we consider flat-rate taxes; later in the article, we analyze nonlinear capital income taxes. In an earlier version, we allowed for progressive labor income taxes, but they did not change any of our substantive conclusions ([Güvönen et al. 2019](#)). Finally, we capture the U.S. estate tax system with a 40% flat-tax rate on bequests.¹⁷

4. *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 its peak at age 51. The AR(1) process has a persistence of $\rho_e = 0.9$ and a standard deviation of $\sigma_e = 0.2$, consistent with the estimates in the literature when a separate transitory shock is not modeled. The intergenerational correlation of the fixed effect is set to $\rho_\kappa = 0.5$ ([Solon 1999](#)). With these parameters fixed, we set $\sigma_{\varepsilon_\kappa} = 0.309$ to match a cross-sectional standard deviation of log labor earnings of 0.80 ([Güvönen et al. 2021](#)). Finally, \bar{y} is set to \$90,000 to map model wealth levels into data counterparts.

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

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

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

The model counterpart of a self-made billionaire is a person who is among the top 400 wealthiest in our simulated U.S. economy (whose minimum wealth is \$1.9 billion, compared with \$2 billion in the 2017 *Forbes* 400 list) and who started life with less than \$250,000 of wealth.¹⁹ The latter criterion is likely to be a generous upper bound for someone coming from an upper-middle-class family. If we use a \$100,000 cutoff instead, the fraction of self-made in the model is 50%, versus 56% under the original definition. Notice that these cutoffs imply a lower bound of 8,000-fold to 20,000-fold wealth growth over the life cycle. It is not possible to generate this incredible speed of wealth accumulation in models of wealth inequality that rely on idiosyncratic earnings shocks or heterogeneity in patience.²⁰

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

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

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

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

growth from being too persistent and therefore overstating overall inequality.

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

With these definitions, the model jointly matches the wealth share of entrepreneurs (38.6% versus 41.6% in the data) and the population share of entrepreneurs within the top 1% wealthy (68.1% versus 65% in the data) as well as their share within the top 10%, which was not targeted (40% versus 42% in the data).²¹ Although the model understates the population share of entrepreneurs we targeted (7.1% versus 11.5% in the data), other data sets and slightly different definitions yield lower empirical estimates.²² Finally, the model generates a life cycle pattern of

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

22. For example, [Cagetti and De Nardi \(2006\)](#) consider a definition that adds the condition that an active business owner also reports being self-employed (and use it as their benchmark). This group’s population share is 7.6% and wealth share is 33%, not too far from our benchmark. Using this latter definition and data from the Panel Study of Income Dynamics, [Salgado \(2020\)](#) reports a population share slightly above 7% in 2014. [Gentry and Hubbard \(2000\)](#) define an entrepreneur as an active business owner with at least \$5,000 in assets and report a population

entrepreneurship that is hump shaped (with the fraction of entrepreneurs rising until late 30s and then tapering and declining in the 50s) as has been documented in the data (see [Online Appendix Figure C.2](#)).²³

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

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

The business debt to GDP ratio can also be estimated from firm-level data. Using this approach, [Asker, Farre-Mensa, and Ljungqvist \(2011\)](#) report an average debt-to-asset ratio of 0.20 for publicly listed firms and a ratio of 0.31 for private firms in the United States. With a capital/output ratio of 3, and assuming as an upper bound that all capital stock is owned by firms, these leverage figures correspond to a debt-to-GDP ratio between 0.60 and 0.93, implying significantly tighter financial constraints than those in our benchmark. The bottom line is that our baseline model

share of 8.7% and a wealth share of 39% for entrepreneurs, figures similar to what our model generates. So overall, our calibrated model is well within the range of figures reported in the literature.

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

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

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

allows significant amounts of borrowing by U.S. firms, which are likely to be on the upper end of the empirical estimates.

7. *Alternative Parameterization: Targeting Lower Inequality:*

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

IV.B. *Performance of the Benchmark Model*

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

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

Figure I, Panel A plots the log counter-CDF of wealth against log wealth, which should be a straight line with a slope of $-\alpha$ if wealth has a Pareto distribution: $P(\omega > x) \sim x^{-\alpha}$. The blue

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

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

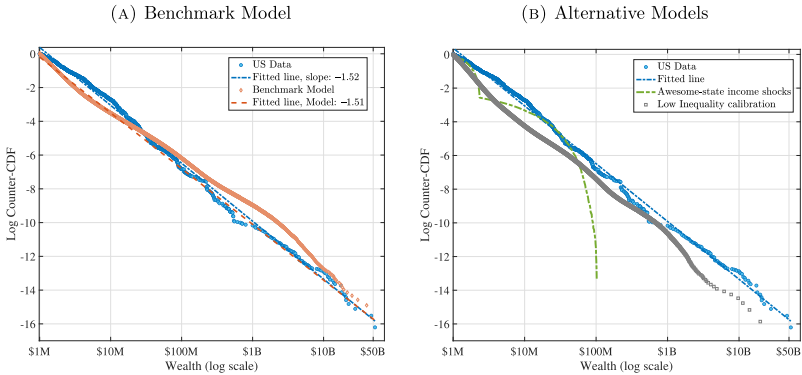


FIGURE I

Pareto Tail of the Wealth Distribution: U.S. Data versus Models

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

circles are the U.S. data (from Vermeulen 2018), which essentially form a straight line all the way up to \$50 billion, with a slope of -1.52 , confirming a thick Pareto right tail. The model counterpart (orange diamonds) aligns well with the data, especially below \$500 million and above \$10 billion, with a slight overestimation between these points. The fitted line has a slope of -1.51 , even though this moment was not targeted. That said, the top 0.1% share is somewhat higher in the model than in the data (23% versus 16%), although the robustness checks in Section VII show that this does not affect our substantive conclusions.²⁸

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

28. The 16% figure is from Smith, Zidar, and Zwick (2023, table 1).

economy has about \$100M in wealth—or 500 times less than in the data.

The baseline model is consistent with some other key facts about wealth and capital income, which we briefly discuss here and present in detail in [Online Appendix B.1](#). For example, capital income is more concentrated than wealth in the model and the U.S. data, and the magnitudes broadly align as well.²⁹ The model also generates a steeper age profile for capital income than in labor income—a 2.9-fold raise between ages 25 and 50 versus a 50% rise for labor income consistent with the U.S. data ([Piketty, Saez, and Zucman 2018](#)). Another important statistic is the inter-generational correlation of wealth, which matters for the effects of taxation and depends on many factors (parent-child correlation in abilities, bequests, estate taxes, retirement, other motives for sav-ings, etc.). [Online Appendix Figure C.3](#) shows the rank-rank plot of parent-child wealth in the model (left panel), whose shape and magnitudes match up well with the data (right) from [Fagereng et al. \(2020, figure 11\)](#). Further, the 90th–10th percentile gap in wealth of parents translates into a 16 percentile gap in wealth among their children in both the data and the model.

2. Rate of Return Heterogeneity. Rate of return heterogeneity is the key ingredient behind the results in this article, so we examine whether our calibration implies a plausible distribution for returns. As noted earlier, the availability of individual return data is very recent, and the most detailed analysis comes from Norway ([Fagereng et al. 2020](#)) with some less detailed statistics from U.S. data ([Smith, Zidar, and Zwick 2023](#)). In [Table IV](#), we report statistics about the dispersion of annual and long-run (persistent) individual returns as well as the level of returns at the top (which is critical for the right tail). Overall, the benchmark model compares well with the data for the overall population (top row) from [Fagereng et al. \(2020\)](#), with all statistics on dispersion and top returns within 1 or 2 percentage points of their empirical

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

TABLE IV
DISTRIBUTION OF RATES OF RETURN (UNTARGETED) IN THE MODEL AND THE DATA

	Annual returns			Persistent component of returns					
	Std. dev.	P90–P10	Kurtosis	Std. dev.	P90–P10	Kurtosis	P90	P99	P99.9
Data (Norway)	8.6	14.2	47.8	6.0	7.7	78.4	4.3	11.6 [†]	23.4 [†]
Data (Norway, bus. own.)	—	—	—	4.8	10.9	14.2	10.1	—	—
Data (U.S. private firms)	17.3	32.9	8.5	—	—	—	—	—	—
Benchmark model	8.4	17.1	7.6	4.1	9.2	6.1	5.8	13.9	19.7
L-INEQ calibration	6.7	13.1	9.2	3.8	9.2	4.3	5.6	11.2	15.8

Notes. Returns on wealth in percentage points. All cross-sectional returns are value weighted. †The statistics for Norway are for individual returns on wealth (net worth) taken from Fagereng et al. (2020), with the exception of P99 and P99.9, which have been kindly provided to us by the authors. The U.S. statistics are from Smith, Zidar, and Zwick (2023, table B.2) and are for S-corps' returns on investment; the authors also report statistics for partnerships, which are very similar (std. dev. of 19.0% and P90–P10 of 35.6). For each individual, the persistent component of returns is calculated as the unweighted average of annual, before-tax, returns between ages 25 and 75, after taking out the average return by age following Fagereng et al. (2020).

counterparts.³⁰ For comparison, the last row shows the L-INEQ calibration, which, as expected, displays a smaller dispersion and, more importantly, lower returns at the very top of the distribution of the persistent component of returns than those in the benchmark and the data.

The second row reports statistics that are available for business owners, which are again either in line with or higher than those in the benchmark model. The third row shows dispersion statistics of annual returns on investment for private firms from the U.S. tax data (reported by [Smith, Zidar, and Zwick 2023](#)), which are almost twice as high as in the model. This should not come as a surprise, given that these are returns on a narrower definition of investments with high returns. (Unfortunately, statistics on long-run returns are not available.) Our main takeaway from these results is that the return distribution generated by the model seems quite plausible and in line with the data, even though none of these statistics were targeted.

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

V. TAX REFORM

In this section, we analyze the effects of a tax reform that replaces a capital income tax (setting $\tau_k = 0$) with a flat-rate

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

wealth tax, τ_a , while keeping all other taxes fixed. In [Section VI](#), we conduct an optimal tax analysis in which the labor income tax is chosen jointly with τ_a or τ_k to maximize welfare. Compared with that analysis, the tax reform we study here serves two important purposes. First, it is a simpler experiment in that (i) it does not rely on the social objective function maximized, and (ii) by keeping other taxes fixed, it allows us to focus on the trade-offs between the two capital taxes in isolation of other mechanisms that would become operational when, for example, τ_ℓ were also adjusted. Second, its relative simplicity makes it more appealing from a policy perspective than an optimal policy that requires changes to several tax tools simultaneously.

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

V.A. Results

The RN reform requires a wealth tax of $\tau_a = 1.19\%$ to generate the same revenue as the benchmark U.S. economy (with $\tau_k = 25\%$).³¹ The BB reform requires a slightly higher rate, $\tau_a = 1.67\%$, mainly because of the added cost of higher pensions. A glance at the left panel of [Table V](#) shows that aggregate quantities increase across the board with the switch to a wealth tax. In the RN reform, K and Q are higher by 16.4% and 22.6%, respectively, and the $\frac{K}{Y}$ ratio rises from 3 to 3.2. The larger increase in Q relative to K reflects the improved reallocation of capital due to the wealth tax. This improvement in efficiency can be expressed as a 5.3% increase in TFP in the intermediate-goods sector, which increases aggregate TFP by 2.1%.

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

TABLE V
TAX REFORM: CHANGE IN MACRO VARIABLES FROM CURRENT U.S. BENCHMARK

	Quantities (% change)						Prices (change)			
	K	Q	TFP_Q	L	Y	C	\bar{w}	\bar{w} (net)	Δr^\dagger	Δr^\dagger (net)
RN reform	16.4	22.6	5.3	1.2	9.2	9.5	8.0	8.0	0.21	-0.36
BB reform	9.2	16.0	6.2	1.2	6.9	7.7	5.6	5.6	0.67	-0.38

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

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

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

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

V.B. Welfare Analysis

To quantify the welfare effects of the reform, we use two measures. The first one, CE_1 , is constructed at the individual level, which allows us to quantify the gains/losses experienced by different groups. For an h -year-old in state $S \equiv (a, \mathbf{S})$ in the

TABLE VI
AVERAGE WELFARE GAIN FROM TAX REFORM

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

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

U.S. benchmark, $CE_1(h, S)$ is the fixed proportional consumption transfer at all future dates and states that makes her indifferent between the stationary equilibria of the two economies. We can calculate welfare change measures for different groups by integrating $CE_1(h, S)$ over the stationary distribution of the group in the U.S. benchmark.³² We use \overline{CE}_1 to denote the aggregated measure for newborn individuals. The second one, \overline{CE}_2 , is a macro measure, à la Lucas (1987): it is the fixed proportional consumption transfer (at all dates and states) to all newborn individuals in the U.S. benchmark such that average lifetime utility is equal to that in the tax-reform economy. This measure provides a single figure that is easy to interpret and allows comparison with some previous work. The exact formulas for the calculations described here are in Online Appendix A.3.

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

How are the welfare gains distributed across the population? In Table VII, we divide the population into five age groups—age 20, 21–34, 35–49, 50–64, and retirees—and six bins for

32. Using the stationary distribution of the “reform” economy makes little quantitative difference.

TABLE VII
WELFARE GAIN/LOSS BY AGE GROUP AND ENTREPRENEURIAL ABILITY

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

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

entrepreneurial ability that become finer at the top end. Each cell reports the corresponding average welfare change. There are several takeaways. First, all newborn groups gain from the RN reform, and these gains are fairly evenly distributed across ability groups—ranging from 6.7% for the lowest 40% to 13.4% for the top 0.1% \bar{z}_i -group.³³ Second, welfare gains decline with age, which is to be expected: since wealth rises and productivity (z_{ih}) falls on average with age, the ratio of capital income to wealth falls, in turn raising the tax burden of wealth taxation relative to capital income taxation. This effect partially offsets the income gains from higher wages under wealth taxation, resulting in declining welfare gains by age. Despite this decline, the welfare change is positive for all working-age groups except those with the highest ability (the top 0.1%) whose losses from higher taxes on their large wealth outweigh their gains from higher wages.

That retirees lose from the RN reform is not surprising: by design, their pensions remain fixed at the U.S. benchmark, so they

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

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

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

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

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

VI. OPTIMAL TAXATION

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

In Sections VI.A to VI.C, we compare the stationary equilibria of OWT and OKIT with each other and with the U.S. benchmark. We conduct a full transition analysis in Section VI.D.

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

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

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

TABLE VIII
OPTIMAL TAXATION: TAX RATES AND AVERAGE WELFARE EFFECTS

	Benchmark U.S. economy	RN reform (1)	OWT (2)	OWT L-INEQ (3)	OWT-X (4)	WTE-X (5)	OKIT (6)
Tax rates							
τ_k	25.0	—	—	—	—	—	-13.6
τ_a	—	1.19	3.03	2.54	3.80 [†]	3.30	—
τ_ℓ	22.4	22.4	15.4	18.1	14.4	17.7	31.2
Δ Welfare							
\overline{CE}_1	—	6.8	9.0	6.0	9.1	8.4	4.2
\overline{CE}_2	—	7.2	8.7	5.2	8.8	8.6	5.1

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

L-INEQ calibrations will be a recurring theme in the results below (as was also the case in the tax reform).

Turning to OKIT (last column), it provides a subsidy to capital income and imposes a high labor tax: $\tau_k = -13.6\%$ and $\tau_\ell = 31.2\%$, which seems surprising at first blush. Whereas OWT shifts the tax burden from labor to capital, OKIT does the opposite—it taxes labor more heavily to subsidize capital income. This is a stark contrast between two tax systems that are equivalent without rate of return heterogeneity.

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

VI.A. Changes in Macro-Variables

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

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

TABLE IX
OPTIMAL TAXATION: CHANGES IN MACROECONOMIC OUTCOMES

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

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

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

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

In contrast to OWT, OKIT (last column) causes rather dramatic changes in the economy. The subsidy policy boosts the

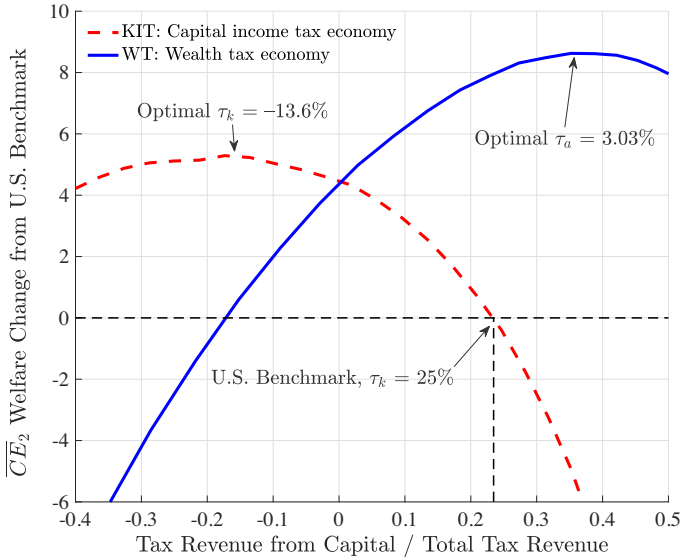


FIGURE II
Welfare Gains from Optimal Taxation

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

VI.B. Mechanisms at Play

To better understand the differences between the two tax systems, we plot how the welfare objective (CE_2 measure) varies with the share of tax revenue raised from capital in [Figure II](#). Welfare changes on the y-axis are relative to the U.S. benchmark, which is normalized to zero. Because the total tax revenue is fixed

(= $G + SSP$), as the revenue share from capital varies along the x -axis, the labor tax adjusts in the background to balance the government budget. The optimal tax rate is found where the objective value is maximized in this figure.

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

The comparison between the distortions created by each tax system can be seen more clearly in [Figure III](#). There are two major differences. First, a higher wealth tax reduces K (solid blue line) more gradually than a higher capital income tax (dashed red line). In other words, the same amount of revenue can be raised with the former with a smaller distortion to K than with the latter.³⁷ Second, and compounding the first effect, Q declines more gradually than K under wealth taxation, whereas the reverse happens under capital income taxation. Thus, wealth taxation improves the efficiency of capital allocation (Q versus K), whereas capital income taxation does the opposite. The relationship reverses for subsidies, so the capital income subsidy results in both an increase in K and a more efficient allocation.

37. A higher elasticity of aggregate capital with respect to a capital income tax than with respect to a wealth tax puts relatively more downward pressure on the optimal capital income tax, consistent with [Saez and Stantcheva \(2018\)](#), who show that a higher elasticity reduces the optimal capital tax.

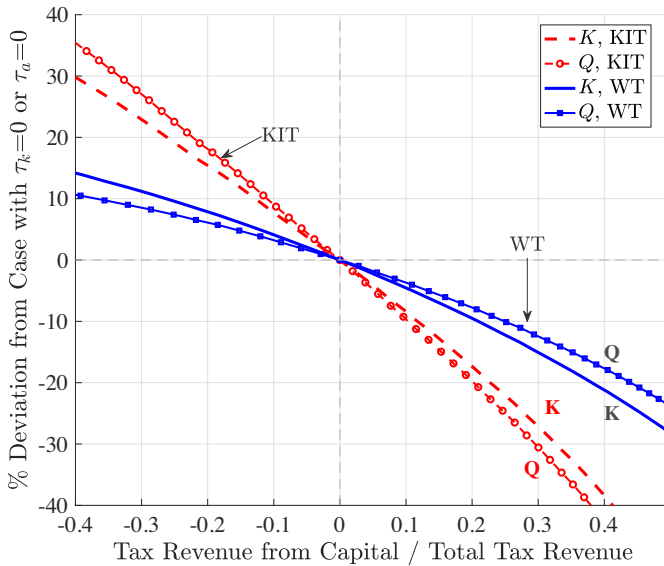


FIGURE III
How K and Q Vary with Revenue Raised from Taxing Capital

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

The optimality of a capital income subsidy stands in sharp contrast with some well-understood results in the literature. For example, [Conesa, Kitao, and Krueger \(2009\)](#) find an optimal $\tau_k = 36\%$ in a model that shares many features with ours, with

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

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

VI.C. Who Gains and Who Loses?

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

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

1. *Decomposing Welfare Changes: Levels versus Redistribution.* We decompose the average welfare gain to quantify the contribution of level versus redistributive changes stemming from each policy (reported in Table VIII).³⁹

39. The contribution of level changes is calculated by starting from the U.S. benchmark and scaling up or down the consumption and leisure allocations of all newborns by the percentage change in corresponding aggregates. The average welfare gain corresponding to these hypothetical allocations (call it $\overline{CE}_2^{\text{level}}$) gives the contribution of level changes, with the remaining portion of \overline{CE}_2 attributable

TABLE X
DECOMPOSITION OF WELFARE GAINS

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

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

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

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

to distributional changes. More specifically, the two components can be written as $1 + \overline{CE}_2 \approx (1 + \overline{CE}_2^{\text{level}})(1 + \overline{CE}_2^{\text{redistr.}})$; see Online Appendix A.3. Decomposing \overline{CE}_1 yields very similar results. We have further decomposed each component into the part coming from consumption and labor and found that most of the gains are coming from consumption, which is unsurprising given that the changes in labor supply are small.

opposite happens at older ages. In other words, the exemption level strengthens the redistribution from old to young among the most productive entrepreneurs.

Given that the optimal exemption level is fairly low, we also examine the effects of higher values. We find that values up to 100% of \bar{y} deliver only slightly lower average welfare gains than the optimum but exempts about 60% of the population from the wealth tax, delivering redistributive gains between OWT and OWT-X (WTE-X in Table X). However, further increasing $\underline{\alpha}_{ex}^*$ (not reported) leads to rapid declines in average as well as redistributive gains, suggesting that thresholds in the millions of dollars, which are often proposed in policy debates, may not improve overall welfare or achieve desirable redistributive objectives.

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

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

The second case is when the society does not have strong preferences for redistribution—or aversion to inequality—which corresponds to a lower risk aversion in our setup. In the baseline, we set $\sigma = 4$ for comparability with Conesa, Kitao, and Krueger (2009), who also studied the equity-efficiency trade-off in a related setting. We solve the model for $\sigma = 1$ and repeat the OWT

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

experiment (not reported in the table). This raises the level gains from 5.9% to 9.7% while also turning the distributional gains of 2.6% to a loss of 5.3%, for an overall welfare gain of 4%. The effect is smaller for OKIT, with a negligible decline in level gains and a moderate increase in distributional losses (from 8.3% to 10%) for a total gain that falls from 5.1% to 3.1%

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

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

VI.D. Equilibrium with Transition

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

TABLE XI
EXTENSION: POLICY ANALYSIS ACCOUNTING FOR THE TRANSITION PATH

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

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

with the equilibrium transition path, defining a new fixed-point problem.⁴¹

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

These welfare results are driven by two main factors. First, the immediate reduction in τ_ℓ causes after-tax wage income to jump, raising average consumption by 4.5% in the first year of reform. The lower τ_ℓ also raises labor supply, raising output by 1% in the first year. Moreover, the rise in consumption is larger at the lower end of the income distribution—because wage income accounts for a larger share of total disposable resources—which leads to a more even consumption distribution.

The second factor is that the OWT transition requires (almost) no capital accumulation at the aggregate level because the levels of K in pre- and postreform steady states are virtually the same, differing from each other by a few percent (Table IX). Thus,

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

there is no aggregate capital accumulation that reduces average consumption during the transition. Rather, the gains are driven by the reallocation of capital toward more productive individuals through the use-it-or-lose effect and the differential behavioral savings response across the population (with high-productivity entrepreneurs increasing their savings rate and low-productivity entrepreneurs doing the opposite), which reinforce this reallocation.⁴²

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

This stark reversal stems from two sources and can be anticipated from the discussion of the OWT case. The first one is straightforward: the transition analysis makes explicit the cost of accumulating the large capital stock of the postreform steady state (about 48% higher), which requires higher savings and lower consumption early in the transition. Second, τ_ℓ jumps from 22.4% in the benchmark to 35.9% in the first year of transition, lowering after-tax labor income both directly and indirectly (by depressing labor supply, which falls by 5% in the first year of the new policy, driving a 2.6% fall in output on impact). Both of these costs are carried over the first several decades of the transition, whereas the benefits (higher wages and consumption) are realized only gradually and are thus discounted, adding up to a large welfare loss.

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

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

VII. ROBUSTNESS AND EXTENSIONS

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

1. *Alternative Modeling of Financial Frictions.* We consider several alternative forms and calibrations of financial constraints, three of which we report in [Table XII](#).

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

ii. *Firms Going “Public” Stochastically.* We consider a second extension in which firms stochastically become “public,” by which we mean they gain access to a substantially higher credit limit: $\bar{\vartheta}_{\text{public}} \gg \bar{\vartheta}(z)$. The public/private status of a firm is inherited across generations, and public firms exit ($\mathbb{I}_{ih} = 0$) at a lower rate than private firms. The arrival rate of the “Calvo fairy” is calibrated to generate a target ratio of public to private firms of 0.5%.⁴⁴ We set $\bar{\vartheta}_{\text{public}} = 10$, which corresponds to a leverage ratio of 90%—at the very top end of values seen in the data

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

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

TABLE XII
ROBUSTNESS: OPTIMAL WEALTH TAX

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

Notes: The seven robustness experiments are as follows: (i) replacing collateral constraints with unlimited borrowing subject to a credit spread of 10%, generating a debt-to-GDP ratio of 1.1; (ii) same as (i) but with a spread of 6%; (iii) allowing firms to stochastically transition to relaxed collateral constraints; (iv) introducing a corporate sector with Cobb-Douglas production and no borrowing limits; (v) eliminating z_i heterogeneity to focus on pure monopolistic rents; (vi) a tax reform that replaces τ_k with a nonlinear capital income tax; and (vii) an optimal nonlinear capital income tax experiment (choosing $\psi, \eta, \tau_\ell, \tau_k$).

(Asker, Farre-Mensa, and Ljungqvist 2011).⁴⁵ Further details of the model and calibration are in [Online Appendix E.4](#). The calibration generates a debt/GDP ratio of 2.4, showing that the new structure indeed delivers much more borrowing than the baseline. The optimal wealth tax is slightly lower (2.76%) and generates lower but still robust welfare gains ranging from 4.8% to 5.9%.

iii. *Other Extensions Involving Financial Constraints.* Because firms become public at a gradual pace, we could ask if the results would be different if we relaxed the constraints starting at age zero. As one way to capture this, we recalibrated the baseline collateral constraint to match a debt-to-GDP ratio of 2.5, which implies a rise in ϑ from 2.8 in the baseline to 11.5 for the top ability group. The optimal wealth tax is 2.3%, and the welfare gain is 4.2% (Table E.10 in [Online Appendix E.5](#)). Finally, we solved the model with a flat collateral constraint (i.e., one that does not depend on ability: $\vartheta(z) = \bar{\vartheta}$) and recalibrated as before ([Online Appendix Table E.10](#)). The welfare gain from OWT is 11.2%, higher than the baseline, confirming our assertion in [Section III.C](#) that the baseline specification is a more conservative assumption. Finally, we have also considered a different collateral constraint of the form $k_i \leq a_i + \vartheta \cdot (z_i a_i)^\mu$, which has the same form that emerges from imperfect enforceability of collateral constraints—see the discussion in note 12. OWT yields welfare gains ranging from 4% to 4.4%. Results are available on request.

We now turn to other robustness exercises.

iv. *Adding a Corporate Sector.* We solve a version of the model with an aggregate corporate sector that produces output according to a Cobb-Douglas production function of capital and labor ($Y_c = AK_c^\alpha L_c^{1-\alpha}$) and faces no financial constraints. The final good is a Cobb-Douglas aggregate of the output of the two sectors: $Y = Y_c^\varrho Y_p^{1-\varrho}$, with subscripts denoting “corporate” and “private” sectors, respectively, and $L_c + L_p = L$. There is a common capital market for corporate firms and entrepreneurs where they can borrow at rate r . See [Online Appendix E.3](#) for further details. We set ϱ and A to match the corporate sector’s share of aggregate sales (50%) and the aggregate capital stock (60%). We choose these empirical moments to be on the high end of the estimates of the size

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

of the corporate sector so as to provide an upper bound on the potential impact of adding a corporate sector on our results.⁴⁶ The optimal wealth tax is 3.25% and delivers a welfare gain of 5.5% for newborn individuals.⁴⁷ Because private firms are a smaller fraction of the economy, misallocation is much lower than in the baseline (0.065 here versus 0.16), yet the welfare gain is almost two-thirds of the baseline figure. Most of the gains are due to reallocation within private firms, with only a small contribution from a small shift in aggregate capital from the corporate sector to private firms (a 2% shift in the capital share).⁴⁸

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

To better explain this point, we note that return heterogeneity results here from two features: heterogeneity in z_i and heterogeneity in monopoly profits due to borrowing frictions. Thus, if we eliminate heterogeneity in z_i , all firms will continue to earn monopoly rents (albeit with much less return heterogeneity). If we conduct the optimal wealth tax exercise in this setting, what do we find? [Table XII](#), fifth column shows that the optimal tax rate is $\tau_a = 1.4\%$ (and $\tau_\ell = 27\%$), but it results in a welfare loss of 1.4% to 1.7% relative to the U.S. benchmark—which features

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

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

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

capital income taxes.⁴⁹ In other words, without heterogeneity in z_i , we can replicate the result that it is better to tax excess returns with a capital income tax.⁵⁰ The optimality of a wealth tax and the large welfare gains relative to capital income taxes arise from inherent differences in entrepreneurial ability, z_i , which are powerful enough to overcome the desire to tax rents built into the model.

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

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

vii. *Adding a Wealth Tax on Top of Existing Capital Income Taxes.* The reader will notice that the tax reform and optimal tax

49. Changes in macro quantities corresponding to experiments in Table XII are reported in Online Appendix Table E.11.

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

51. This specification has a long history in public finance and has been used by Bénabou (2000) and Heathcote, Storesletten, and Violante (2014).

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

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

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

In addition, we address the effect of introducing managerial effort into our framework with an extension in [Online Appendix F](#)

(by modifying [equation \(8\)](#) to include the entrepreneur's labor supply) and discuss the conditions under which it amplifies or dampens the effects of taxes relative to the baseline model.

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

VIII. DISCUSSION AND CONCLUSIONS

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

The wealth tax we propose here differs in crucial ways from those that have been implemented by governments in the past and from those that are currently being debated by the public. Specifically, the wealth tax is levied after existing capital income taxes are repealed; it is levied on the book value of assets rather than the market value; and it is broad-based rather than targeted

at the top of the wealth distribution. As a result, its implications are also different in important ways.

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

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

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

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

leviate this problem. Moreover, eliminating the capital income tax would free up the resources of tax agencies that can then be directed toward enforcing the wealth tax.

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

UNIVERSITY OF MINNESOTA, FEDERAL RESERVE BANK OF MINNEAPOLIS, AND NATIONAL BUREAU OF ECONOMIC RESEARCH, UNITED STATES
UNIVERSITY OF TORONTO, CANADA
UNIVERSITY OF TORONTO, CANADA
WESTERN UNIVERSITY, CANADA
ECON ONE, UNITED STATES

SUPPLEMENTARY MATERIAL

An Online Appendix for this article can be found at *The Quarterly Journal of Economics* online.

DATA AVAILABILITY

The data underlying this article are available in the Harvard Dataverse, <https://doi.org/10.7910/DVN/R0YZHY> (Güvenen et al. 2022c).

REFERENCES

- Aiyagari, S. Rao, "Optimal Capital Income Taxation with Incomplete Markets, Borrowing Constraints, and Constant Discounting," *Journal of Political Economy*, 103 (1995), 1158–1175. <https://doi.org/10.1086/601445>.
- Allais, Maurice, *L'impôt Sur Le Capital et la Réforme Monétaire*. (Paris: 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. <https://doi.org/10.3386/w17394>.
- Auerbach, Alan J., Jagadeesh Gokhale, Laurence J. Kotlikoff, John Sabelhaus, and David N. Weil, "The Annuitization of Americans' Resources: A Cohort Analysis," NBER Working Paper no. 5089, 1995. <https://doi.org/10.3386/w5089>.

- Bach, Laurent, Laurent E. Calvet, and Paolo Sodini, "Rich Pickings? Risk, Return, and Skill in Household Wealth," *American Economic Review*, 110 (2020), 2703–2747. <https://doi.org/10.1257/aer.20170666>.
- Bell, Felicité 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.
- Bénabou, Roland, "Unequal Societies: Income Distribution and the Social Contract," *American Economic Review*, 90 (2000), 96–129. <https://doi.org/10.1257/aer.90.1.96>.
- Benhabib, Jess, and Alberto Bisin, "Skewed Wealth Distributions: Theory and Empirics," *Journal of Economic Literature*, 56 (2018), 1261–1291. <https://doi.org/10.1257/jel.20161390>.
- Benhabib, Jess, Alberto Bisin, and Mi Luo, "Earnings Inequality and Other Determinants of Wealth Inequality," *American Economic Review*, 107 (2017), 593–597. <https://doi.org/10.1257/aer.p20171005>.
- Benhabib, Jess, Alberto Bisin, and Shenghao Zhu, "The Distribution of Wealth and Fiscal Policy in Economies with Finitely Lived Agents," *Econometrica*, 79 (2011), 123–157. <https://doi.org/10.3982/ECTA8416>.
- Bils, Mark, Peter J. Klenow, and Ciané Ruane, "Misallocation and Mismeasurement?," Stanford University Working Paper, 2017.
- Buera, Francisco J., Joseph P. Kaboski, and Yongseok Shin, "Finance and Development: A Tale of Two Sectors," *American Economic Review*, 101 (2011), 1964–2002. <https://doi.org/10.1257/aer.101.5.1964>.
- Cagetti, Marco, and Mariacristina De Nardi, "Entrepreneurship, Frictions, and Wealth," *Journal of Political Economy*, 114 (2006), 835–870. <https://doi.org/10.1086/508032>.
- Castañeda, Ana, Javier Díaz-Giménez, and José-Víctor Ríos-Rull, "Accounting for the U.S. Earnings and Wealth Inequality," *Journal of Political Economy*, 111 (2003), 818–857. <https://doi.org/10.1086/375382>.
- Chari, V. V., and Patrick J. Kehoe, "Optimal Fiscal and Monetary Policy," In *Handbook of Macroeconomics*, vol. 1, John B. Taylor and Michael Woodford, eds. (Amsterdam: Elsevier, 1999) 1671–1745. [https://doi.org/10.1016/S1574-0048\(99\)10039-9](https://doi.org/10.1016/S1574-0048(99)10039-9).
- Conesa, Juan Carlos, Sagiri Kitao, and Dirk Krueger, "Taxing Capital? Not a Bad Idea after All!," *American Economic Review*, 99 (2009), 25–48. <https://doi.org/10.1257/aer.99.1.25>.
- De Nardi, Mariacristina, and Fang Yang, "Wealth Inequality, Family Background, and Estate Taxation," *Journal of Monetary Economics*, 77 (2016), 130–145. <https://doi.org/10.1016/j.jmoneco.2015.10.005>.
- Erosa, Andrés, and Martin Gervais, "Optimal Taxation in Life-Cycle Economies," *Journal of Economic Theory*, 105 (2002), 338–369. <https://doi.org/10.1006/jeth.2001.2877>.
- Fagereng, Andreas, Luigi Guiso, Davide Malacrino, and Luigi Pistaferri, "Heterogeneity and Persistence in Returns to Wealth," *Econometrica*, 88 (2020), 115–170. <https://doi.org/10.3982/ECTA14835>.
- Federal Reserve Statistical Release, "Z.1 Financial Accounts of the United States: Flow of Funds, Balance Sheets and Integrated Macroeconomic Accounts," Board of Governors of the Federal Reserve System Technical Report, 2015Q3.
- Gabaix, Xavier, "Power Laws in Economics and Finance," *Annual Review of Economics*, 1 (2009), 255–294. <https://doi.org/10.1146/annurev.economics.050708.142940>.
- Gabaix, Xavier, Jean-Michel Lasry, Pierre-Louis Lions, and Benjamin Moll, "The Dynamics of Inequality," *Econometrica*, 84 (2016), 2071–2111. <https://doi.org/10.3982/ECTA13569>.
- Garriga, Carlos, "Optimal Fiscal Policy in Overlapping Generations Models," Mimeo, 2003.
- Gentry, William M., and R. Glenn Hubbard, "Entrepreneurship and Household Saving," NBER Working Paper no. 7894, 2000. <https://doi.org/10.3386/w7894>.

- Golosov, Mikhail, Aleh Tsyvinski, and Ivan Werning, "New Dynamic Public Finance: A User's Guide," in NBER Macroeconomic Annual 2006, vol. 21, Daron Acemoglu, Kenneth Rogoff, and Michael Woodford, eds. (Cambridge, MA: MIT Press, 2006).
- Guvenen, Fatih, Gueorgui Kambourov, Burhanettin Kuruscu, and Sergio Ocampo, "Taxing Wealth and Capital Income when Returns are Heterogeneous," Working Paper, 2022a.
- Guvenen, Fatih, Gueorgui Kambourov, Burhanettin Kuruscu, Sergio Ocampo-Diaz, and Daphne Chen, "Use It or Lose It: Efficiency Gains from Wealth Taxation," NBER Working Paper no. 26284, 2019. <https://doi.org/10.3386/w26284>.
- Guvenen, Fatih, Gueorgui Kambourov, Burhan Kuruscu, Sergio Ocampo, and Daphne Chen, "Replication Data for: 'Use It or Lose It: Efficiency and Redistributive Effects of Wealth Taxation'," (2022c), Harvard Dataverse, <https://doi.org/10.7910/DVN/R0YZHY>.
- Guvenen, Fatih, Fatih Karahan, Serdar Ozkan, and Jae Song, "What Do Data on Millions of U.S. Workers Reveal about Lifecycle Earnings Dynamics?," *Econometrica*, 89 (2021), 2303–2339. <https://doi.org/10.3982/ECTA14603>.
- Guvenen, Fatih Jr., Raymond J. Mataloni, Dylan G. Rassier, and Kim J. Ruhl, "Offshore Profit Shifting and Aggregate Measurement: Balance of Payments, Foreign Investment, Productivity, and the Labor Share," *American Economic Review*, 112 (2022b), 1848–1884. <https://doi.org/10.1257/aer.20190285>.
- Heathcote, Jonathan, Kjetil Storesletten, and Giovanni L. Violante, "Consumption and Labor Supply with Partial Insurance: An Analytical Framework," *American Economic Review*, 104 (2014), 2075–2126. <https://doi.org/10.1257/aer.104.7.2075>.
- Hendricks, Lutz, "Bequests and Retirement Wealth in the United States," Arizona State University Working Paper, 2001.
- Hsieh, Chang-Tai, and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 124 (2009), 1403–1448. <https://doi.org/10.1162/qjec.2009.124.4.1403>.
- 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), 1–59. <https://doi.org/10.2307/2534413>.
- Imrohorglu, Selahattin, "A Quantitative Analysis of Capital Income Taxation," *International Economic Review*, 39 (1998), 307–328. <https://doi.org/10.2307/2527295>.
- Itskhoki, Oleg, and Benjamin Moll, "Optimal Development Policies with Financial Frictions," *Econometrica*, 87 (2019), 139–173. <https://doi.org/10.3982/ECTA13761>.
- Jones, Charles I., "Taxing Top Incomes in a World of Ideas," *Journal of Political Economy*, 130 (2022), 2227–2274. <https://doi.org/10.1086/720394>.
- Jones, Charles I., and Jihee Kim, "A Schumpeterian Model of Top Income Inequality," *Journal of Political Economy*, 126 (2018), 1785–1826. <https://doi.org/10.1086/699190>.
- Judd, Kenneth L., "Redistributive Taxation in a Simple Perfect Foresight Model," *Journal of Public Economics*, 28 (1985), 59–83. [https://doi.org/10.1016/0047-2727\(85\)90020-9](https://doi.org/10.1016/0047-2727(85)90020-9).
- Kelley, Donna, Slavica Singer, and Mike Herrington, "Global Entrepreneurship Monitor, 2011 Global Report," Global Entrepreneurship Research Association Report, 2011.
- Kim, Tae-Hwan, and Halbert White, "On More Robust Estimation of Skewness and Kurtosis," *Finance Research Letters*, 1 (2004), 56–73. [https://doi.org/10.1016/S1544-6123\(03\)00003-5](https://doi.org/10.1016/S1544-6123(03)00003-5).
- Kitao, Sagiri, "Labor-Dependent Capital Income Taxation," *Journal of Monetary Economics*, 57 (2010), 959–974. <https://doi.org/10.1016/j.jmoneco.2010.09.004>.
- Kochen, Federico, "Finance over the Life Cycle of Firms," New York University Working Paper, 2022.
- Kopczuk, Wojciech, "Taxation of Intergenerational Transfers and Wealth," in *Handbook of Public Economics*, vol. 5, Alan J. Auerbach, Raj Chetty, Martin

- Feldstein, and Emmanuel Saez, eds. (Amsterdam: Elsevier, 2013), 329–390. <https://doi.org/10.1016/B978-0-444-53759-1.00006-6>.
- Krusell, Per, and Anthony A. Smith Jr., “Income and Wealth Heterogeneity in the Macroeconomy,” *Journal of Political Economy*, 106 (1998), 867–896. <https://doi.org/10.1086/250034>.
- Liang, James, Hui Wang, and Edward P. Lazear, “Demographics and Entrepreneurship,” *Journal of Political Economy*, 126 (2018), S140–S196. <https://doi.org/10.1086/698750>.
- Lucas, Robert E., *Models of Business Cycles* (New York: Basil Blackwell, 1987).
- 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*, 104 (2014), 422–458. <https://doi.org/10.1257/aer.104.2.422>.
- Moll, Benjamin, “Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?,” *American Economic Review*, 104 (2014), 3186–3221. <https://doi.org/10.1257/aer.104.10.3186>.
- Organisation for Economic Co-operation and Development, “The Role and Design of Net Wealth Taxes in the OECD,” *OECD Tax Policy Studies* (2018). <https://doi.org/10.1787/9789264290303-en>.
- , “Revenue Statistics—The United States,” 2021.
- Piketty, Thomas, *Capital in the Twenty-First Century* (Cambridge, MA: Belknap of Harvard University Press, 2014).
- Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman, “Distributional National Accounts: Methods and Estimates for the United States,” *Quarterly Journal of Economics*, 133 (2018), 553–609. <https://doi.org/10.1093/qje/qjx043>.
- Quadri, Vincenzo, “Entrepreneurship, Saving, and Social Mobility,” *Review of Economic Dynamics*, 3 (2000), 1–40. <https://doi.org/10.1006/redo.1999.0077>.
- Restuccia, Diego, and Richard Rogerson, “Policy Distortions and Aggregate Productivity with Heterogeneous Establishments,” *Review of Economic Dynamics*, 11 (2008), 707–720. <https://doi.org/10.1016/j.red.2008.05.002>.
- Saez, Emmanuel, and Stefanie Stantcheva, “A Simpler Theory of Optimal Capital Taxation,” *Journal of Public Economics*, 162 (2018), 120–142. <https://doi.org/10.1016/j.jpubeco.2017.10.004>.
- Saez, Emmanuel, and Gabriel Zucman, “Wealth Inequality in the United States since 1913: Evidence from Capitalized Income Tax Data,” *Quarterly Journal of Economics*, 131 (2016), 519–578. <https://doi.org/10.1093/qje/qjw004>.
- Salgado, Sergio, “Technical Change and Entrepreneurship,” Wharton School at University of Pennsylvania Working Paper, 2020. <https://doi.org/10.2139/ssrn.3616568>.
- Scheuer, Florian, and Joel Slemrod, “Taxing Our Wealth,” *Journal of Economic Perspectives*, 35 (2021), 207–230. <https://doi.org/10.1257/jep.35.1.207>.
- Smith, Matthew, Owen Zidar, and Eric Zwick, “Top Wealth in America: New Estimates and Implications for Taxing the Rich,” Princeton University Working Paper, 2021.
- , “Top Wealth in America: New Estimates under Heterogeneous Returns,” *Quarterly Journal of Economics*, 138 (2023), 515–573. <https://doi.org/10.1093/qje/qjac033>
- Solon, Gary, “Intergenerational Mobility in the Labor Market,” in vol. 3A, Orley Ashenvelter and David Card, eds. (Amsterdam: North-Holland, 1999), 1761–1800.
- Stantcheva, Stefanie, “Dynamic Taxation,” *Annual Review of Economics*, 12 (2020), 801–831. <https://doi.org/10.1146/annurev-economics-100119-013035>
- Straub, Ludwig, and Iván Werning, “Positive Long-Run Capital Taxation: Chamley-Judd Revisited,” *American Economic Review*, 110 (2020), 86–119. <https://doi.org/10.1257/aer.20150210>
- Tørsløv, Thomas, Ludvig Wier, and Gabriel Zucman, “The Missing Profits of Nations,” *Review of Economic Studies* (forthcoming). <https://doi.org/10.1093/restud/rdac049>

- US Department of the Treasury, "A Financial System that Creates Economic Opportunities: Capital Markets," 2017.
- Vermeulen, Philip, "How Fat is the Top Tail of the Wealth Distribution?," *Review of Income and Wealth*, 64 (2018), 357–387. <https://doi.org/10.1111/roiw.12279>
- Wolff, Edward N., "Changes in Household Wealth in the 1980s and 1990s in the U.S.," in *International Perspectives on Household Wealth*, Edward N. Wolff, ed. (Cheltenham, UK: Edward Elgar, 2006), 107–150.