Wealth Taxation: The Key to Unlocking Capital Gains*

Guttorm Schjelderup and Floris T. Zoutman

December 15, 2023

Abstract

We study how a wealth tax and a realization-based capital gains tax affect capital market efficiency. We develop a two-period model with investors that are heterogeneous in both the value of an initial investment, and the future return on the initial investment. We show that the realization-based capital gains tax reduces the required rate of return on existing investment below the required rate of return on new investments, resulting in lock-in. A comprehensive wealth tax can eliminate this lock-in effect. We then develop an optimal-tax model that trades of equity gains from the capital-gains and wealth tax to efficiency losses related to intertemporal choice, and lock-in. We derive a criterion for the desirability of a wealth tax based on elasticities that can be estimated empirically. In addition, we find an upper bound on the optimal wealth tax. Finally, we consider an extension in which (part of) long-run capital gains escape taxation, which provides a strong rational for much higher wealth tax rates.

JEL classification: H24, D14, G51, H21,M21
Keywords: Wealth Tax, Capital-gains Tax, Dividend Tax, Lock-in Effect, Capital-market Efficiency

*For helpful discussions and comments we would like to thank Lukasz Mayr, Jarle Møen and Joel Slemrod. The authors declare that they have no competing interests.

1 Introduction

Capital gains are usually taxed on a realization, rather than an accrual basis. Similarly most countries levy a dividend tax which applies only when shareholders elect to pay out dividends. The presence of such realization-based taxes implies that investors exert substantial control on the timing of tax payments. Piketty et al. (2018) compute inequality statistics for the US by creating distributional national accounts. Retained earnings are part of national income and are included in their study on an accrual basis. One of their most striking findings is that a significant share of the rise in top incomes is explained by capital gains from equity and bonds since the late 1990s.

There exists robust evidence that higher realization-based tax rates result in a postponement of realization. Using US state-level data, Agersnap and Zidar (2021) find a negative relationship between the capital-gains tax rate and realized capital gains. For the 2013 tax hike, Saez (2017) shows that the reduction in realized capital gains is primarily driven by a timing effect rather than a permanent reduction. Using a panel of French firms and a dividend tax hike, Bach et al. (2021) find that following the reform, closely-held firms reduce their dividend payments, and reinvest in financial assets substituting for portfolio savings at the private level.

Investor control over the timing of tax payments has a profound effect on the progressiveness of the tax system. Using US tax returns and data from the Fortune 400, Yagan (2023) finds that the effective income tax rate on income earned by the 400 wealthiest American households equals only 9.6 percent. Similarly, research from France, (Bozio et al., forthcoming), and the Netherlands, (Bruil et al., 2022), uncovers tax regressivity towards the top of the income distribution. The sharp difference between statutory and effective tax rates is mostly driven by the fact that control of retained earnings and unrealized capital gains is highly concentrated towards the top of the income distribution.

Realization-based taxes also negatively affect the efficiency of capital markets. The reason is that investors may forego profitable investment opportunities, if new investments require the realization of capital gains or retained earnings. Consequently, realization-based taxes reduce the required rate of return on existing investments, relative to new investments.

In this paper we show that in the presence of realization-based taxes, a wealth tax can simultaneously reduce inequality, and enhance capital market efficiency. Unrealized capital gains and retained earnings are a part of net wealth, and can hence be taxed through a comprehensive wealth tax reducing inequality. Further, and this is the primary contribution of our paper, the wealth tax provides an incentive to realize capital gains when the rate of return on those assets lies below the market-rate of return.

To understand why the wealth tax incentivizes realization, consider a capital-gains tax that applies at realization. Realizing capital gains results in a capital-gains tax payment which mechanically reduces net wealth, and therefore reduces future wealth tax payments. Hence, whereas the realization-based capital gains tax provides an incentive on investors to postpone realization, the wealth tax encourages immediate realization. We derive a simple formula for the wealth tax that fully restores capital-market efficiency. The formula only depends on the capital-gains tax rate, and the market interest rate. For an interest rate of 2.5 percent, and a capital gains tax rate of 35 percent, a comprehensive wealth tax of 1.6 percent ensures that the required rate of return on existing and new investment is equalized, restoring capital-market efficiency.

Our model is set up as follows. We consider a two-period model with a population of investors. Each investor is born with an existing investment whose value consists of a principal investment, and an initial capital gain. Investors are heterogeneous along two dimensions: i.) their initial capital gain, and ii.) the rate of return they will (with certainty) obtain in period 1 if they keep the initial investment. Investors consume in both periods, and can borrow or lend at the market interest rate. They face a portfolio choice, since they can either keep the initial investment, or sell the initial investment and lend the proceeds at the market interest rate. We assume interest income/expenditure is taxed/deducted at the same rate as realized capital gains which henceforth we refer to as the (investment)-income tax.[1]

Capital-market efficiency implies that investors sell their initial investment whenever the rate of return on the initial investment lies below the market-interest rate, and retain the investment otherwise. In our model, this outcome arises in the absence of taxation. However, the income tax

[1]We relax the assumption that all capital-income is taxed at the same rate in an extension.

distorts this decision, because postponing the sale reduces the net-present value of tax payments on the initial capital gain. Hence, in the presence of the income tax, the required rate of return on the initial investment lies below the market-interest. This implies that a portion of investors keeps their initial investment, even though the its return lies below the market-interest rate, reducing the surplus in the economy.

A wealth tax can alleviate, or even eliminate the inefficiency, because realizing capital gains in period 1 mechanically reduces wealth tax payments in period 2. Hence, in the presence of the income taxes, increasing the wealth tax reduces the capital market distortion.[2] When the wealth tax, τ equals $\hat{\tau} \equiv \frac{(1-t)r}{1+r}$ where t is the income tax rate and r the market-interest rate, the required rate of return equals the market-interest rate for all investors. Hence, setting the wealth tax to this level fully restores capital-market efficiency.

In the second part of our paper, we consider optimal-tax policy. Introducing an income tax in conjunction with a wealth tax equal to $\hat{\tau}$ restores capital-market efficiency. However, the downside is that both taxes contribute to a significant intertemporal wedge. Optimal tax policy thus balances obtaining tax revenue with efficiency losses related to lock-in and intertemporal distortions.

We use our model to derive sufficient-statistics formulas for the optimal (linear) tax on income and on wealth. These formulas equate the mechanical welfare gain resulting from an increase in the capital-gains and wealth tax to the marginal excess burden of each instrument. The excess burden of the taxation depends on two sets of elasticities. The first set relates taxable income and wealth to the net-of-income -and -wealth tax rate through the intertemporal channel. Under standard intertemporal preferences, these elasticities are positive indicating that both taxes distort taxable income and wealth downwards. A second set of elasticities between taxable income/wealth and the net-of-income/wealth tax rate arises through portfolio choice. Higher income taxes reduce the fraction of investors who realize their capital gain in period 1, which in turn reduces taxable income. Contrary, higher wealth taxes increase the fraction of investors who realize their capital gain in period 1, increasing taxable income. This implies that the

[2]In the absence of realization-based taxes the wealth tax does not affect capital-market efficiency.

income tax is more distortive than it would be in the absence of portfolio choice, whereas the wealth tax is less distortive.

We employ the model to generate three sets of results. First, we examine the optimal wealth tax for a given (exogenous) income tax rate. Here our main result is a sufficient condition for assessing the desirability of the wealth tax. If the total cross-elasticity of taxable income with respect to the net-of-wealth tax is negative, the government always optimally sets a positive wealth tax. Intuitively, in this case, the portfolio effect of the wealth tax outweighs the intertemporal effect. Therefore, the wealth-tax is efficiency-enhancing, implying that the government should optimally levy a wealth tax, even if it has no social preference for reducing wealth inequality.

Next, we consider joint optimal taxation of capital gains and wealth in a setting where social welfare weights are constant across the population, such that the government does not exhibit preferences for redistribution. In this case, the government's objective reduces to minimizing the deadweight loss associated with financing exogenous expenditure. The social optimum is thus obtained when the marginal excess burden per unit of tax revenue of the income and the wealth tax are equalized.

Our first result in this setting is that in the absence of portfolio-responses, the government will finance all expenditure through either the income tax, or the wealth tax. Whether the government uses the income tax or the wealth tax depends on the amount of excess returns in the economy, and on the government's discount rate. The advantage of the income tax is that it can tax excess returns at a smaller intertemporal distortion than the wealth tax. The reason is that the income tax only taxes the return, whereas the wealth tax also taxes the principal investment. On the other hand, the wealth tax taxes all investors in both periods, whereas the income tax only applies upon realization. Hence, the wealth tax is better suited when the government strongly discounts future tax payments relative to the current tax payers. Hence, if excess returns are large relative to the time-valuation of tax revenue, the income tax generates a lower intertemporal distortion per unit of tax revenue than the wealth tax. Empirically, we believe that this condition is likely to be met, since excess returns represent a large, and increasing share of the economy (e.g. Barkai (2020)), whereas the interest rate at which the government can borrow is typically low.

Adding back in portfolio responses, we show that the government should optimally apply both the capital-gains tax and the wealth tax when i.) the income tax generates a smaller intertemporal distortion per unit of tax revenue than the wealth tax, and ii.) evaluated at a zero wealth tax, the overall distortion of the income tax exceeds the overall distortion of the income tax. This latter condition, is likely met when portfolio-responses are large, since portfolio responses increase the distortion of the income tax and reduce the distortion of the wealth tax.

We derive an upper bound on the optimal wealth tax when the income tax exhibits a lower intertemporal distortion per unit of revenue than the wealth tax. In this case, we show that the optimal wealth-tax lies strictly below the level required for full capital-market efficiency (i.e. below $\hat{\tau}$). Quantitatively, this upper limit is in broad agreement with wealth taxes observed around the world.

Next, we relax the assumption that welfare weights are constant, and allow them to correlate negatively with taxable wealth and income, such that the government values redistribution from rich to poor. It is generally more difficult to arrive at strong policy prescriptions in this setting. However, we show that relative to the case with constant welfare weights, the government levies higher income (wealth) tax rates when welfare weights correlate stronger to income (wealth) than to wealth (income).

One setting with redistributive preference where we are able to arrive at strong conclusions is the setting with a Rawlsian planner. Similar to the case with constant welfare weights, the Rawlsian planner also equates the marginal distortion per unit of tax revenue for the two tax instruments. Hence, the upper limit on the wealth tax attained under constant welfare weights also applies for Rawlsian preferences.

We consider an extension in which the government taxes long-run (period-2) capital-gains taxes at a lower rate than short-run capital gains, and interest income/expenditure. This extension corresponds to the US tax system, which taxes short-run capital gains and interest in a comprehensive income-tax system with a top rate of 37 percent. The capital gains tax rate on assets held for more than a year is (at most) 20 percent.

The reduced tax rate on long-run capital provides a strong incentive for postponing the realization of capital gains, and hence, exacerbates capital-market inefficiency. Moreover, lock-in can now occur even among investors who do not have an initial capital gain, since postponing the realization reduces the effective tax rate on the return obtained in period 1.

The wealth tax also improves capital-market efficiency in this setting, because postponement of capital gains results in a higher wealth level in period 2. However, there is no longer a single wealth-tax rate that restores capital-market efficiency for all investors. Instead, we focus on the wealth-tax rate that restores efficiency for investors with a large initial capital gain. For the US, the wealth tax rate that attains capital-market efficiency among investors with large capital gains is more than an order of magnitude larger than in our base model. Hence, if the capital-gains tax is unable to fully tax long-run gains, far larger wealth tax rates are warranted.

Literature The wealth tax is a controversial topic that has been debated by economists, policymakers, and the public for decades. A core topic in the academic literature is whether wealth taxes stimulate or impede economic activity. The effect of the wealth tax has, for example, been studied in relationship with entrepreneurial risk taking (Hall and Woodward (2010), migration (Young et al. (2016); Advani and Tarrant (2021), tax evasion (Guyton et al. (2020)), and tax avoidance (Alvaredo and Saez (2009); Hemel (2019), There is also a large literature on the effect of taxation on portfolio choice starting with the seminal paper by Sandmo (1977), and a nascent literature studying the effects of wealth taxes on taxable wealth (Brülhart et al., 2022; Seim, 2017; Jakobsen et al., 2018; Londoño-Vélez and Ávila-Mahecha, 2021; Ring, 2020; Garbinti et al., 2023).

To our knowledge the only other paper that highlights an efficiency motivation for levying a wealth tax is Guvenen et al. (2023). Guvenen et al. (2023) develop a paper in which agents have heterogeneous returns and the government can tax both capital-income and wealth.[3] They show that in the presence of a capital-market imperfection in which entrepreneurs are credit constrained, the wealth tax can be used to redistribute funds from entrepreneurs with low productivity to entrepreneurs with high productivity. Our findings are largely complimentary to their findings. Unlike, Guvenen et al. (2023) we assume that capital markets are perfect in the absence of taxation. We instead assume that the tax system is imperfect by (realistically) assuming that the capital-gains tax applies upon realization rather

[3] In addition to Guvenen et al. (2023) a number of other papers consider capital income taxation in the context of heterogeneous returns (e.g. Gerritsen et al., 2020,Boadway and Spiritus, 2021,Ferey et al., 2021). However, these papers do not consider wealth taxation.

than accrual. We show that in this setting the wealth tax can also enhance efficiency. Hence, jointly the papers show that the wealth tax can enhance efficiency if either capital markets, or the tax system contain imperfections.

Saez and Zucman (2019) and Piketty et al. (2023) both consider wealth taxation in the presence of capital gains. They discuss how a (progressive) wealth tax can aide in restoring tax progressivity by taxing capital gains prior to realization. This effect of wealth taxation on inequality is also present in our paper. However, our primary contribution lies in showing that the wealth tax can also encourage investors to realize capital gains earlier, which improves capital-market efficiency.

The literature discusses a number of other mechanisms to reduce inefficient lock-in. Most notable examples are the retrospective taxation of capital gains, Auerbach (1988), and a withholding tax on capital gains, Saez et al. (2021). Relative to these proposals, the wealth tax has two advantages. First, it has been implemented in several developed and developing countries. Second, the wealth tax can be administered without keeping records over time, whereas both a retrospective tax on capital gains, and a withholding tax on capital gain require record-keeping for the duration in which the asset is held.

2 Set-up

We set up a model with 2 periods where time is indexed by $k \in \{1, 2\}$. We consider a unit-mass of investors indexed by i. Each investor initially holds an asset worth W_1^i, and no other assets/debt such that W_1^i also represents initial wealth. The value consists of an exogenously given capital gain R^i and a price for which the asset is purchased, which we normalize to 1, formally $W_1^i \equiv (1 + R^i)$. We assume investors initially own no other assets or debts such that wealth W_1^i represents the initial wealth of an investor. If the investor keeps the investment, it will (with certainty) yield a net rate of return equal to ρ^i.

Investors are heterogeneous across two dimensions. The initial locked-in capital gain R^i is distributed according to marginal cumulative density function $F^R(R^i)$ defined within the support $[R_0, R_1]$, with $R_1 > R_0 \geq 0$. Note that heterogeneity in the initial locked-in capital gain also generates inequality in the distribution of initial wealth W_1^i. The rate of return ρ^i is distributed according to the conditional cumulative density function $F^{\rho}(\rho^i|R^i)$ with support $[\rho_0, \rho_1]$. Note that by conditioning the density of ρ^i on R^i we are effectively allowing for any arbitrary correlation between the rate of return and the initial locked in capital gain. We do assume that both $F^R(\cdot)$ and $F^{\rho}(\cdot|R^i)$ are continuous and strictly increasing within their respective support. These assumptions respectively rule out mass points and holes in the joint distribution of (R^i, ρ^i). In addition, we assume the conditional distribution of ρ^i is differentiable with corresponding probability density function $f^{\rho}(\cdot|R^i)$.

In period 1 investors can lend or borrow at rate $r > 0$. We model this through investment in a risk-free asset B^i, where $B^i > 0$ corresponds to lending, and $B^i < 0$ corresponds to borrowing. This setup allows us to capture the classical approach to modelling lock-in in which an investor foregoes investing in a more profitable investment, since this requires the realization of capital gains (see for instance the model in Auerbach, 1988). In addition, it allows us to model a setting in which the investor finances consumption by borrowing funds rather than realizing capital gains, which appears to coincide with the tax strategy of some of the wealthiest Americans (see for instance the leaked tax returns that are discussed in Eisinger et al., 2021).

The objective of the investor is to maximize the utility function $u(C_1^i, C_2^i)$, where C_k^i denotes consumption of investor i in period k. We make standard restrictions on preferences by assuming that the utility function is twice differentiable, increasing and concave.

The investor faces three tax instruments. The first is a tax on investment income at rate t. In the continuation we refer to this as the income tax. The income tax is levied on realized capital gains, as well as on interest income. Interest expenditure is deductible at the same rate. The second, is a wealth tax which taxes net wealth (i.e. assets minus debt) at a rate τ. We assume the wealth tax applies to beginning-of-period wealth.[4] The third is a lump-sum transfer M, which can take positive or negative values (signifying a lump-sum tax), and which the investor receives/pays in period 1.

To finance consumption investors require liquid funds. In the first period,

[4]An alternative interpretation is that the wealth tax is calculated on the basis of end-of-period wealth, but payable in the next period. The second interpretation is consistent with the way wealth taxes are typically administered (see e.g. Thoresen et al., 2022).

they can attain these by selling a share ϕ^i of their initial investment. We assume $1 \geq \phi^i \geq 0$, such that investors cannot expand or short-sell their initial holdings. Taxable income in period 1 is thus given by $I_1^i \equiv \phi^i R^i$. In addition, they receive funds through the lump-sum transfer. Finally, investors can attain liquid funds through borrowing. The intratemporal budget constraint in period 1 is thus given by:

$$\begin{aligned}
\underbrace{\phi^i W_1^i + M - B^i}_{\text{Liquid Funds}} &= C_1^i + \underbrace{\tau W_1^i}_{\text{Wealth Tax Liability}} + \underbrace{t I_1^i}_{\text{Income Tax Liability}} \leftrightarrow, \\
\phi^i W_1^i + M - B^i &= C_1^i + \tau W_1^i + tR^i - (1-\phi^i)tR^i, \qquad (1)
\end{aligned}$$

where the final term on the right hand side in (1) represents the period-1 tax benefit the investor obtains from (partially) postponing his capital gains.

Investors accrue a capital gain on the remainder of their initial investment equal to the remaining investment multiplied by the rate of return: $\rho^i(1-\phi^i)W_1^i$. In addition, they earn interest income equal to rB^i. In period 2 investors sell off their remaining assets/debt, and use the proceeds to finance consumption and tax expenditure.

To understand tax expenditure in more detail, note first that the wealth tax burden depends on wealth at the beginning of period 2 given by the sum of the left-over initial investment, and the investment in bonds:

$$W_2^i \equiv (1-\phi^i)(1+\rho^i)W_1^i + (1+r)B^i. \qquad (2)$$

Taxable income in period 2 consists of the capital-gain on the remaining investment, $(1-\phi^i)((1+\rho^i)W_1^i - 1)$, and interest income. Hence, taxable income in period 2 is given by:

$$I_2^i \equiv (1-\phi^i)((1+\rho^i)W_1^i - 1) + rB. \qquad (3)$$

Using these variables, we can write the period 2 intratemporal budget constraint as follows:

$$\begin{aligned}
W_2^i &= C_2^i + \underbrace{tI_2^i + \tau W_2^i}_{\text{Tax Expenditure}} \leftrightarrow, \\
(1-\phi^i)(1+\tilde{\rho}^i)W_1^i + (1+\tilde{r})B^i &= C_2^i + t(1-\phi^i)R^i \qquad (4)
\end{aligned}$$

where $\tilde{r} \equiv (1-t-\tau)r - \tau$, denotes the return on the risk-free asset after wealth and income taxes, and $\tilde{\rho}^i \equiv (1-t-\tau)\rho^i - \tau$ denotes the after-tax return on W_1^i in the counterfactual scenario in which capital gains are taxed on accrual, rather than on realization. The first term on the left-hand side of (4) represents the value of the remaining investment in W_1^i after taxes, whereas the second term represents the value of the investment in B^i. The second term on the right-hand side represents the income tax payment investors have to make on the share of the initial investment they did not sell.

To arrive at the intertemporal budget constraint we solve equation (4) for B and substitute the resulting expression into (1):

$$\underbrace{C_1^i + \frac{C_2^i}{1+\tilde{r}}}_{\text{NPV Consumption}} + \underbrace{tR^i\left(\frac{1+\phi^i\tilde{r}}{1+\tilde{r}}\right)}_{\text{NPV tax on } R^i} + \underbrace{\tau W_1^i}_{\text{Initial Wealth tax}} = \underbrace{W_1^i\left(\frac{1+\tilde{\rho}^i}{1+\tilde{r}} + \phi^i\frac{\tilde{r}-\tilde{\rho}^i}{1+\tilde{r}}\right)}_{\text{NPV Asset}} + M, \quad (5)$$

which states that the net present value of consumption plus tax expenditure equals the net present value of the original investment. The most interesting part of this equation is the net present value of taxes on the initial capital gain R^i. This term depends positively on the fraction of the investment sold in period 1, ϕ^i, (assuming the after-tax interest rate $\tilde{r} > 0$). A larger value of ϕ^i implies a larger part of the original capital gain is taxed immediately, hence raising the net-present value of tax expenditure. This term therefore gives rise to an incentive to postpone the realization of capital gains, as we discuss in more detail in the next section.

Before turning to the equilibrium, it is useful to discuss a number of reinterpertations of the model. First, we have described W^i as a (financial) asset which can be sold to realize a capital gain. An alternative interpretation is that W^i is an investment in a closely-held corporation that is fully controlled by the investor. In that case, ρ^i denotes the return on assets within the corporation and ϕ^i denotes the share of liquid funds attained through the corporation by either selling off (part of) the company, or paying out dividends. This alternative interpretation is entirely in line with our model provided that dividends and capital gains are taxed at the same rate t.

Additionally, we have assumed that the initial capital gain R^i is exogenous. Alternatively, we can microfound this by adding a period 0, in which all investors are endowed with $W_0 = 1$ which is invested at rate of return R^i. Again, this reinterpretation is fully in line with our model, as long as we assume the investor does not consume in period 0.[5][6]

3 Equilibrium

In this section, we derive the equilibrium, and derive our main result showing that wealth taxation enhances capital market efficiency. We first consider the first-order conditions of the agents under general income and wealth taxes, and rewrite them in a familiar consumption-Euler condition, and a condition on the optimal realization of capital gains in period 1. We use these to derive Proposition 1 showing that in the absence of wealth taxes, a positive income tax will result in inefficient lock-in for a positive fraction of investors. Proposition 2 is our main result showing that capital-market efficiency can be restored with the wealth tax.

Assigning Lagrange-multiplier λ to the budget constraint (5), we arrive at the following first-order conditions for the investors:

$$\frac{U_{C_1}(C_1(\cdot), C_2(\cdot))}{U_{C_2}(C_1(\cdot), C_2(\cdot))} = 1 + \tilde{r}, \tag{6}$$

$$\lambda \frac{\tilde{r}(W_1^i - tR^i) - \tilde{\rho}^i W_1^i}{1 + \tilde{r}} \lesseqgtr 0, \tag{7}$$

where henceforth for any equilibrium quantity X, $X(\rho^i, R^i)$ denotes its equilibrium value for an investor with rate of return ρ^i and initial gain R^i. The equilibrium quantities also depend on the tax variables t, τ, M but for brevity these arguments are omitted. In addition, functional dependence is usually shortened using $X(\rho^i, R^i) = X(\cdot)$ notation. For future reference $V(\rho^i, R^i)$ denotes indirect utility as a function the rate of return and the initial capital gain.

Equation (6) is the investor's Euler equation for consumption. Note that

[5] Adding consumption in period 0 would significantly increase the technical complexity of the model, without yielding additional intuition.

[6] To be fully consistent with our model, we would also need to assume no wealth tax applies in period 0. However, since all individuals are endowed with the same principal investment, wealth taxes in period 0 are equivalent to the lump-sum instrument M.

the right-hand side of (6) is the same for all investors even though investors are heterogeneous in their rate of return on the initial asset. The reason is that investors have access to a common asset B, which in equilibrium serves as the marginal source of liquid funds. Hence, at the margin all investor's face the same relative price of period 2 vs period 1 consumption.

Equation (7) described optimal portfolio choice. Here, the sign $\lesseqgtr$ should be understood as follows. When the left-hand side of (7) is positive, resources in the budget constraint (5) strictly increase in ϕ^i, and hence it is optimal for the investor to sell off all of his/her initial asset, $\phi^i = 1$. Contrary, when the left-hand side is negative, resources decrease in ϕ^i, and it is optimal for the investor to keep the initial asset, $\phi^i = 0$. We ignore the knife-edge case in which the left-hand side equals zero since our assumption that the distribution of (R^i, ρ^i) does not contain mass-points, implies that the set of investors for which this knife-edge condition holds, constitutes a zero measure.

The information in first-order condition (7) can be summarized into the following formula that defines the required rate of return:

$$\rho(R^i) \equiv \frac{1 + (1-t)R^i}{1 + R^i} r + \frac{t\tau R^i}{(1 - t - \tau)(1 + R^i)} \tag{8}$$

Investors whose rate of return exceeds the required rate $\rho(R^i)$ will keep the initial investment, $\phi^i = 0$. Investors, whose rate of return lies below $\rho(R^i)$ will sell, $\phi^i = 1$.

3.1 Taxation and Capital-Market Efficiency

Next, we show the relationship between taxation and capital-market efficiency using equation (8). We show that i.) the income-tax can push the required rate of return below the market-interest rate resulting in a capital-market inefficiency, ii.) the wealth tax does not, on its own, create a similar inefficiency, and iii.) the wealth tax can be used to mitigate the capital-market inefficiency when the income tax is positive.

In this subsection we ignore intertemporal distortions, and focus solely on the efficiency of capital markets. We will revisit the intertemporal distortion in the next section when we discuss optimal-tax policies. Here, we will instead work with a more narrow concept of efficiency that ignores intertemporal efficiency, and is defined below:

Definition 1 Inefficient lock-in*: There is inefficient lock-in when a positive fraction of investors has a rate of return on their initial investment $\rho^i < r$, but nevertheless optimally chooses $\phi^i < 1$. Formally, the fraction of locked-in investors is defined as:*

$$F^{locked-in} \equiv \mathbb{E}\big[L(\phi(\cdot) < 1, \rho^i < r\big], \tag{9}$$

where $L(\cdot)$ is an indicator function that returns value 1 if the argument(s) inside are true and zero otherwise, and $\mathbb{E}[\cdot]$ is the (conditional) expectation operator, which takes the expectation with respect to the distribution of (R^i, ρ^i).

To see that this definition of inefficient lock-in captures an efficiency loss, note that keeping part of the initial investment, when it yields a return (ρ^i) below the market interest rate (r) constitutes a reduction in the total surplus of the economy. We consider capital markets to function efficiently when the fraction of inefficiently locked-in investors equals zero. Our first Proposition shows that inefficient lock-in is the result of the realization-based income tax, whereas a wealth tax on its own does not generate inefficient lock-in.

Proposition 1 *Assume $t + \tau < 1$ and $t, \tau \geq 0$. The equilibrium then satisfies the following properties:*

1. *The fraction of locked-in investors can be written as:*

$$F^{locked-in} = \mathbb{E}\big[L\left(\rho(R^i) < \rho^i < r\right)\big]. \tag{10}$$

2. *If the income tax is zero, $t = 0$, there is no inefficient lock-in i.e. $F^{locked-in} = 0$ for all τ.*

3. *A positive income tax $t > 0$ in combination with a zero wealth tax, $\tau = 0$ implies that some investors are inefficiently locked in, $F^{locked-in} > 0$.*

Proof. Note that the required rate of return $\rho(R^i)$ is undefined for $t+\tau = 1$, which is why we need to assume this upper bound on the tax rates. For part 1, substitute (8) into (9) for $\phi(\cdot)$ noting that investors will keep the initial investment if ρ^i is above the required rate of return. For Part 2 substitute $t = 0$ into (8) to arrive at $\rho(R^i) = r$, implying that the set $\{\rho^i : \rho(R^i) < \rho^i < r\}$ has zero measure. For part 3, substitute $\tau = 0$ into

(8) to arrive at:

$$\rho(R^i) = \frac{1 + (1-t)R^i}{1 + R^i} r < r, \tag{11}$$

for all $t, R_1 > 0$. Since we assumed that the distribution of ρ^i, R^i is continuous without holes and that r is contained in the support of the distribution, this implies there exist investors in the economy whose rate of return lies above $\rho(\cdot)$ and below r. ■

Proposition 1 shows that in the absence of a wealth tax, positive income tax rates result in inefficient lock-in for some investors. The intuition is that postponing the capital gain reduces the net-present value of income tax payments. This can make it attractive to keep the initial investment even if the rate of return lies below the market interest rate. The Proposition also shows that the wealth tax does not cause a similar distortion. The reason is that wealth taxes apply on an accrual, rather than a realization basis, such that postponing realization does not result in a reduction in the net-present value of wealth tax payments.

The Proposition is graphically depicted in Figure 1, where the downward-sloping curve represents the required rate of return as a function of the capital gain R^i for a given (positive) income tax rate t. The curve starts at $(0, r)$, since there is no lock-in for investors that have no initial capital-gain. The required rate of return converges to $1 - t$ as R approaches infinity. The area between the curves represents the region where investors are inefficiently locked in. Increasing the income tax increases the curvature of the required rate of return $\rho(\cdot)$, and rotates it clock-wise around the point $(0, r)$. This results in a larger fraction of inefficiently locked-in investors.

The next Proposition derives our main result by showing that when $t > 0$, the wealth tax can be used to restore capital-market efficiency.

Proposition 2 *Assume the income-tax rate is positive, $t \in (0, 1)$. The fraction of locked-in investors, decreases in the wealth tax for $\tau \in [0, \hat{\tau})$, and equals zero when the wealth tax equals $\hat{\tau}$, where $\hat{\tau}$ is defined as:*

$$\hat{\tau} \equiv \frac{r(1-t)}{1+r}. \tag{12}$$

Proof. From (10) it follows that inefficient lock-in only occurs when the required rate of return lies below the market-interest rate, $\rho(R^i) < r$ for at least one value of R^i. Additionally, the fraction of locked-in investors

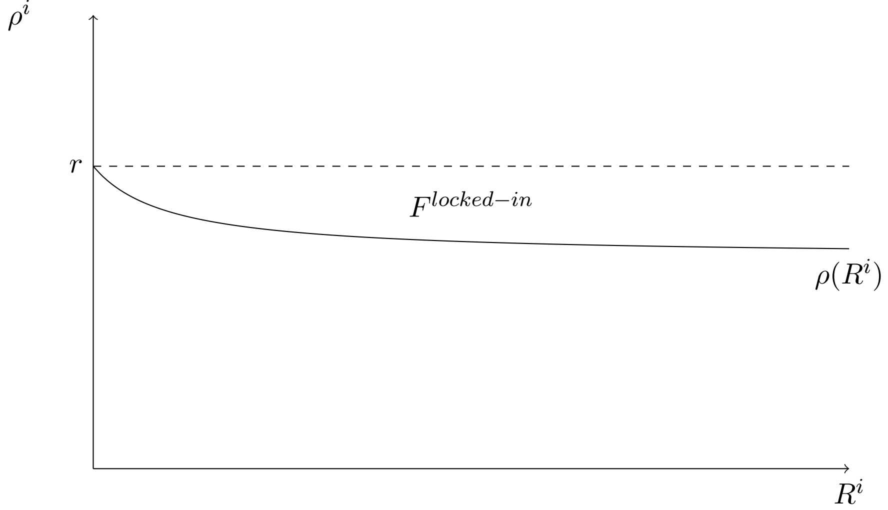

Figure 1: Effect of taxation on inefficient lock-in

Note: The Figure shows the required rate of return $\rho(R^i)$ for a given positive income-tax rate t, alongside the market-interest rate r. The area between the two curves contains inefficiently locked-in investors, since their rate of return exceeds the required rate of return, but lies below the interest rate. Increasing the income-tax results in an increase in the curvature of $\rho(R^i)$ alongside a rotation around $(0, r)$. Increasing the wealth tax has the opposite effect.

decreases in the required rate of return as long as $\rho(R^i) < r$. Hence, we can prove the Proposition, by showing that i.) the required rate of return increases in the wealth tax and that ii.) $\rho(R^i) = r$ when $\tau = \hat{\tau}$, such that $\hat{\tau}$ implies there is no inefficient lock-in.

Note first that $\tau \leq \hat{\tau}$ implies that $t + \tau < 1$ such that the denominator of the second term in (8) is positive. Further τ only appears in the second term of (8) and this term increases in τ when $t > 0$ as we have assumed. Hence, the fraction of locked-in investors decreases in the wealth tax for $\tau \in [0, \hat{\tau})$. Second, substituting $\hat{\tau}$ into equation (8) yields:

$$\rho(R^i) = \frac{1 + (1-t)R^i}{1 + R^i} r + \frac{t\hat{\tau}R^i}{(1 - t - \hat{\tau})(1 + R^i)} = r.$$

■

Proposition 2 derives a formula for the wealth tax (12) that restores capital-market efficiency. A surprising feature is that $\hat{\tau}$ does not depend on the (distribution of) locked-in capital gains and rates of return (ρ^i, R^i). To understand this, note that keeping the initial investment postpones the payment of income taxes. However, when the wealth tax rate equals $\hat{\tau}$

the after-tax interest rate $\tilde{r} = 0$.[7] This implies that under this wealth-tax rate investors do not discount future tax payments relative to current tax payments, and hence, the tax incentive to postpone the realization of capital gains is eliminated.

Proposition 2 can be represented in Figure 1 as follows. When $t > 0$ increasing the wealth tax flattens the required rate of return $\rho(\cdot)$ and rotates the curve counter-clockwise around $(0, r)$. At $\tau = \hat{\tau}$ the required rate of return-curve is fully flattened at $\rho(\cdot) = r$.[8]

Since $\hat{\tau}$ only depends on the income-tax rate and the interest rate, it is easily quantifiable. If we assume the interest rate is 2.5 percent, consistent with the average (real) return on bonds in the US over a long time horizon (see Jordà et al. (2019)), and we assume an income-tax rate of 37 percent consistent with the top income-tax rate in the US then we arrive at a wealth tax rate of 1.6 percent.

It should be noted that setting the wealth tax equal to $\hat{\tau}$ constitutes a large intertemporal distortion, since it eliminates private returns to saving. Therefore, when the income tax is positive an optimal wealth tax rate must trade off the intertemporal distortion to the capital market distortion. In the next section, we study this trade-off in greater detail through the lens of an optimal-tax model, in which the government chooses welfare-maximizing tax rates. The model provide us with a better understanding of the circumstances under which inefficient lock-in can be part of optimal tax policy, and to what extend policy makers should use wealth taxes to enhance capital-market efficiency.

4 Optimal Tax Rates

In this section we describe optimal tax policy taking into account both intertemporal distortions and capital-market efficiency. We first describe the government's objective and budget constraint, before turning to a general

[7] To see this simply substitute $\hat{\tau}$ into the definition of $\tilde{r}$.

[8] Setting the wealth tax above the efficient rate $\tau > \hat{\tau}$ results in inefficient lock-out, a phenomenon in which some investors sell their initial investment, although the rate of return lies above the market-interest rate. The reason is that $\tau > \hat{\tau}$ implies a negative after-tax interest rate $\tilde{r} < 0$ and investors discount current cashflows relative to future cashflows providing an incentive to pay taxes in the first, rather than the second period. We ignore this phenomenon here, since as we discuss in the next section, a wealth tax $\tau > \hat{\tau}$ is rather unlikely to be part of optimal tax policy.

description of optimal-tax policy phrased in terms of sufficient statistics. We then consider several cases of interest. First, we derive a condition under which the government should optimally tax wealth for exogenously given income taxes. Second, we consider optimal wealth and income taxes in the case in which welfare weights are constant, but the government is restricted from using lump-sum taxes as a method for financing all expenditure. Finally, we discuss the extend to which insights attained with restricted welfare weights can be generalized.

4.1 Government

The government's objective is to maximize a weighted sum of individual indirect utility:

$$\mathbb{E}\big[\alpha_i V(\rho^i, R^i)\big], \tag{13}$$

where α_i is the Pareto weight attached to each individual. We assume the government collects tax revenue in order to finance the lump sum transfer M and exogenous expenditure whose net present value equals E. The government budget constraint can thus be written as:

$$E + M \leq \mathbb{E}[(tI(\cdot) + \tau W(\cdot))], \tag{14}$$

in which $I(\rho^i, R^i) \equiv I_1(\rho^i, R^i) + \frac{I_2(\rho^i, R^i)}{1+r}$ denotes (the net-present value of) taxable income, and $W(\rho^i, R^i) \equiv W_1^i + \frac{W_2(\rho^i, R^i)}{1+r}$ denotes taxable wealth.

We assume that the government is restricted in his choice of optimal-tax policy through lower-bound constraints on the tax instruments. We assume that the wealth-tax rate and the income-tax rate have to be non-negative. In addition, we introduce a lower-bound M_0 on the lump-sum transfer. We do not make a priori restrictions on the sign of M_0 such that the government might still be able to set lump-sum taxes (i.e. negative values of M). However, in our analysis we will sometimes consider the case where the lower-bound restriction on M binds.

The full optimal-tax problem of the government is hence to maximize objective (13) subject to the budget constraint (14) and lower-bound constraints on the tax and transfer instruments. In the next subsection, we use this formulation to derive optimal-tax formulas.

4.2 Optimal Tax Formula

Using the objective (13) and the budget constraint (14) we can write the government's Lagrangian as:

$$
\begin{aligned}
\mathcal{L} &= \int_{\underline{R}_1}^{\overline{R}_1} \int_{\underline{\rho}}^{\rho(R^i)} \left(\frac{\alpha_i V(\cdot)}{\eta} + tI(\cdot) + \tau W(\cdot) \right) dF(\rho|R) dF(R) \\
&+ \int_{\underline{R}_1}^{\overline{R}_1} \int_{\rho(R^i)}^{\overline{\rho}} \left(\frac{\alpha_i V(\cdot)}{\eta} + tI(\cdot) + \tau W(\cdot) \right) dF(\rho|R) dF(R) \\
&-M - E, \qquad (15)
\end{aligned}
$$

where η denotes the Lagrange-multiplier on the government's budget constraint. In addition, we have switched from the expectation operator to integral-notation which allows us to split between investors who sell their initial investment $\rho^i < \rho(R^i)$, and investors who keep it, $\rho^i > \rho(R^i)$. We make this split to account for the fact that given our assumption on the utility function, we can guarantee that the equilibrium quantities $X(\rho^i, R^i)$ are differentiable everywhere, except potentially at the required rate of return $\rho^i = \rho(R^i)$. The reason is that at this point investors switch between selling and keeping the initial investment. This discontinuity subsequently causes a discontinuity in reported income and wealth, as we show below. Splitting the integral into two differentiable parts allows us to apply Leibniz's rule for taking the derivative underneath the integral sign to each part, and then recombining the integrals afterwards.

We maximize the Lagrangian (15) with respect to the tax rates $t, \tau \geq 0$ and $M \geq M_0$. First-order conditions are given by:

$$
\begin{aligned}
\frac{\partial \mathcal{L}}{\partial t} \leq 0: \quad 0 &\geq \mathbb{E}\left[\frac{\alpha_i}{\eta} \frac{\partial V}{\partial t} + t \frac{\partial I}{\partial t} + \tau \frac{\partial W}{\partial t} + I \right] + \qquad (16) \\
&\quad \mathbb{E}\left[\frac{\partial \rho}{\partial t} \left(t \Delta I(\cdot) + \tau \Delta W(\cdot) \right) f^{\rho}(\rho^i|R_i) \,\middle|\, \rho^i = \rho(\cdot) \right], \\
\frac{\partial \mathcal{L}}{\partial \tau} \leq 0: \quad 0 &\geq \mathbb{E}\left[\frac{\alpha_i}{\eta} \frac{\partial V}{\partial \tau} + t \frac{\partial I}{\partial \tau} + \tau \frac{\partial W}{\partial \tau} + W \right] + \qquad (17) \\
&\quad \mathbb{E}\left[\frac{\partial \rho}{\partial \tau} \left(t \Delta I(\cdot) + \tau \Delta W(\cdot) \right) f^{\rho}(\rho^i|R_i) \,\middle|\, \rho^i = \rho(\cdot) \right], \\
\frac{\partial \mathcal{L}}{\partial M} \leq 0: \quad 0 &\geq \mathbb{E}\left[\frac{\alpha_i}{\eta} \frac{\partial V}{\partial M} + t \frac{\partial I}{\partial M} + \tau \frac{\partial W}{\partial M} - 1 \right]. \qquad (18)
\end{aligned}
$$

Note that the first-order conditions are written as inequalities, rather than equalities. The reason is that we have imposed lower-bound constraints on the tax-instruments. For values t, τ, M above their lower-bound the first-order equations (16)-(18) must hold with equality, but at the lower bound inequality is sufficient.

Additionally $\Delta I(\cdot)$ and $\Delta W(\cdot)$ are defined through the following limits.

$$\Delta I(R^i) \equiv \lim_{\rho^i \xrightarrow{-} \rho(R^i)} I(\rho^i, R^i) - \lim_{\rho^i \xrightarrow{+} \rho(R^i)} I(\rho^i, R^i), \tag{19}$$

denotes the difference in reported income between the investor who marginally sells his initial investment, and the investor who marginally keeps his investment, and

$$\Delta W(R^i) \equiv \lim_{\rho^i \xrightarrow{-} \rho(R^i)} W(\rho^i, R^i) - \lim_{\rho^i \xrightarrow{+} \rho(R^i)} W(\rho^i, R^i), \tag{20}$$

denotes the corresponding difference in reported wealth. Typically, taxable income and wealth exhibit a discontinuity around $\rho^i = \rho(\cdot)$, (i.e. $\Delta I(\cdot), \Delta W(\cdot) \neq 0$) because portfolio choice has a direct and mechanical impact on taxable wealth and income in the presence of realization-based taxes. Lemma 1 proves the existence of the limits in the definitions of $\Delta I(\cdot)$ and $\Delta W(\cdot)$, and expresses their value as a function of the initial locked-in gains R^i.

Lemma 1 *Assume the tax rates satisfy* $t, \tau \geq 0$ *and* $t + \tau < 1$*. In that case, the limits in* (19) *and* (20) *exist and are given by:*

$$\Delta I(R^i) = \frac{rR^i(1-t-\tau) - t\tau R^i}{(1-t-\tau)(1+r)} \tag{21}$$

$$\Delta W(R^i) = -\frac{tR^i(1-t)}{(1-t-\tau)(1+r)} \tag{22}$$

Proof. *The proof can be found in the Appendix.* ■

Lemma 1 summarizes the difference in reported income and wealth between the investor that marginally sells his initial investment, and the investor that marginally keeps his investment. From equation (21) we can see that, absent wealth taxation ($\tau = 0$), the taxable income of the marginal investor selling his investment is higher than that of the marginal investor keeping the initial investment. Intuitively, keeping the investment postpones the realization of income, and hence reduces the net present value of taxable income.

On the other hand, the difference in reported wealth $\Delta W(\cdot)$ is non-positive and strictly negative when $t > 0$. Intuitively, the marginal investor who sells his investment pays income taxes in period 1 which mechanically reduces taxable wealth in period 2.

Using these terms we rephrase the government's first-order conditions in sufficient statistics as follows. Marginally increasing tax rates has the following welfare implications. First, increasing the income (wealth) tax mechanically increases tax revenue by taking funds from investors proportional to their income (wealth). At the margin, these funds are used to increase the lump-sum transfer. To evaluate the welfare-impact of this mechanical transfer we rely on the Diamond definition of the social marginal value of one unit of private income. The Diamond social welfare weight (henceforth social welfare weight) includes the direct welfare impact of transferring a unit of income to investor i. In addition, it includes the tax-revenue effect that follows from the income effect generated by the transfer (see Diamond, 1975; Jacobs, 2018). Formally:

$$g(\rho^i, R^i) \equiv \frac{\alpha^i U_{C_1}}{\eta} + t\frac{\partial I}{\partial M} + \tau\frac{\partial W}{\partial M}. \tag{23}$$

The mechanical welfare impact of increasing the income (wealth) tax is positive, when social welfare weights decrease in income (wealth). The reason is that the government in that case values redistribution from investors with high income (wealth) to low income (wealth).

The second welfare implication follows from the fact that increasing the income and wealth tax decreases the after-tax interest rate. This results in a substitution response that decreases the accumulation of wealth and income on the intertemporal margin. This effect is welfare-relevant because it creates a fiscal externality. A reduction in the accumulation of wealth reduces both taxable income and taxable wealth. The strength of this response margin can be characterized by the compensated elasticity of tax-base y with respect to net-of-tax rate $1 - \sigma$ through the intertemporal channel:

$$\varepsilon_{y,1-\sigma}(\rho^i, R^i) \equiv -\frac{\partial y}{\partial \sigma}|_{V=V_0}\frac{1-\sigma}{y} \quad \text{for} \quad y \in \{I, W\}, \sigma \in \{t, \tau\}. \tag{24}$$

Third, changing the income and wealth tax affects portfolio choice by affecting the required rate of return $\rho(R_i)$ for investor's to keep/sell their initial investment. This also creates a fiscal externality, because the reported income and wealth depends on whether investors keep or sell their initial investment. The effect of a change in net-of-tax rate $1-\sigma$ on tax base y at a particular capital-gain R_i via this portfolio channel consists of the product between: i.) the change in the required rate of return as a result of increasing $1-\sigma$, $-\frac{\partial \rho(R^i)}{\partial \sigma}$, ii.) the density of investors that are indifferent between buying and selling their initial investment conditional on capital gain R_i, $f^\rho(\rho(R^i)|R_i)$, and iii.) the difference in tax payments between the marginal investor who sells his investment, and the marginally locked in investor Δy. To convert this portfolio response into a portfolio elasticity we multiply by the net-of-tax rate $1-\sigma$ and divide by the average reported income/wealth at R_i:

$$\xi_{y,1-\sigma}(R^i) \equiv -\frac{\partial \rho(R^i)}{\partial \sigma} \frac{(1-\sigma) f^\rho(\rho(R_i)|R_i)\Delta y(R_i)}{\mathbb{E}[y(\rho, R) \mid R = R^i]} \quad \text{for} \quad y \in \{I, W\}, \sigma \in \{t, \tau\}. \tag{25}$$

Using these definitions, we can write the government's first-order conditions as follows:

Proposition 3 *Equations* (16)-(18) *can be rewritten in sufficient statistics as:*

$$1-\bar{g}^I \leq \frac{t}{1-t}\left(\underbrace{\bar{\varepsilon}^I_{I,1-t}}_{>0}+\underbrace{\bar{\xi}^I_{I,1-t}}_{>0}\right)+\frac{\tau}{1-t}\frac{\bar{W}}{\bar{I}}\left(\underbrace{\bar{\varepsilon}^W_{W,1-t}}_{>0}+\underbrace{\bar{\xi}^W_{W,1-t}}_{\leq 0}\right), \tag{26}$$

$$1-\bar{g}^W \leq \frac{\tau}{1-\tau}\left(\underbrace{\bar{\varepsilon}^W_{W,1-\tau}}_{>0}+\underbrace{\bar{\xi}^W_{W,1-\tau}}_{\geq 0}\right)+\frac{t}{1-\tau}\frac{\bar{I}}{\bar{W}}\left(\underbrace{\bar{\varepsilon}^I_{I,1-\tau}}_{>0}+\underbrace{\bar{\xi}^I_{I,1-\tau}}_{\leq 0}\right), \tag{27}$$

$$\bar{g}-1 \leq 0, \tag{28}$$

In which for any sufficient statistic $y(\cdot)$, $\bar{y} \equiv \mathbb{E}[y(\cdot)]$ denotes the average value of $y(\cdot)$, and $\bar{y}^x \equiv \frac{\mathbb{E}[y(\cdot)x(\cdot)]}{\mathbb{E}[x(\cdot)]}$ the average value of $y(\cdot)$ weighted by $x(\cdot)$. For the sign-restrictions on $\bar{\xi}^I_{I,1-\sigma}$ we have assumed that the wealth tax rate is weakly below the level that eliminates lock-in ($\tau \leq \hat{\tau}$).

Proof. *The proof can be found in the Appendix.* ∎

The left-hand side of equations (26)-(28) denotes the mechanical welfare gain associated with increasing, respectively, the income tax, the wealth tax and the lump-sum transfer by one unit of income. For the income (wealth) tax this mechanical effect consists of increasing tax revenue by 1 unit of income, and reducing private utility proportional to income (wealth) resulting in a welfare loss equal to the income-weighted average welfare welfare weight $\bar{g}^I$ (wealth-weighted average welfare welfare weight $\bar{g}^W$). Increasing the lump-sum transfer is costly in terms of tax revenue, but transfers a unit of income to all investors resulting in a welfare gain of $\bar{g} - 1$.

The right-hand side of equations (26)-(28) represents the marginal excess burden per unit of tax revenue associated with each instrument. The lump-sum transfer is non-distortive and hence its excess burden equals zero (equation (28)).

For both the income tax (equation (26)) and the wealth tax (equation (27)) the excess burden can be decomposed in an own-base response (i.e. the effect of the income (wealth) tax on taxable income (wealth)) represented by the first-term on the right-hand side and a cross-base response represented by the second term (i.e. the effect of the income (wealth) tax on taxable wealth (income)). Cross-base responses are weighted by the ratio between the size of the cross-base and the size of the own-base (for the income tax $\bar{W}/\bar{I}$. The reason is that the cross-base response becomes more important when the cross-base is larger relative to the own-base.

The responses can further be decomposed in an intertemporal effect (represented by the compensated intertemporal elasticity $\varepsilon_{y,1-\sigma}$) and a portfolio-effect ($\xi_{y,1-\sigma}$). The intertemporal elasticities $\varepsilon_{y,1-t}$ are all positive, since a (compensated) change in income and wealth taxes reduces the after-tax interest rate, discouraging the accumulation of wealth. The portfolio own-base elasticities, ($\xi_{I,1-t}$,$\xi_{W,1-\tau}$) are also positive, but the portfolio cross-base elasticities ($\xi_{I,1-\tau}$,$\xi_{W,1-t}$) are negative. Intuitively, increasing the income-tax rate reduces the fraction of investors who sell their initial investment. This in turn reduces taxable income, but increases taxable wealth. Conversely, increasing the wealth-tax rate increases the fraction of investors who sell their initial investment, increasing taxable income, but reducing taxable wealth.

4.3 Exogenous income taxes

The optimal wealth-tax expression (27) can be used to evaluate the desirability of a wealth tax in the context of an exogenous positive income tax. When t is exogenous, a positive wealth tax is desirable if, evaluated at $\tau = 0$, the mechanical benefits of the wealth tax exceed the margininal excess burden. Mathematically:

$$1-\bar{g}^W > \frac{t\bar{I}}{\bar{W}}\left(\bar{\varepsilon}^I_{I,1-\tau}+\bar{\xi}^I_{I,1-\tau}\right). \tag{29}$$

Note that because the formula is evaluated at $\tau = 0$, the marginal excess burden of the wealth tax depends only the cross-base response, that is, the effect of the wealth tax on taxable income. For a sufficiently small wealth tax the own-base response is second order.

From the inequality we can see that a wealth tax is more likely to be desirable if the negative correlation between reported wealth and welfare weights becomes stronger (i.e. more negative). The intuition is that in that case, redistribution through the wealth tax is more desirable, since the government's objective increases more in wealth redistribution. Additionally, the wealth tax is more desirable if the compensated effect of the wealth tax on income through the intertemporal channel becomes smaller, measured by a smaller value of $\bar{\varepsilon}^I_{I,1-\tau}$, and when the wealth tax is more effective in restoring capital market efficiency (i.e. a more negative value of the portfolio response $\bar{\xi}^I_{I,1-\tau}$).

Using equation (29) we formulate a sufficient condition for the government to optimally set a positive wealth tax $\tau > 0$ which we describe in the Proposition below.

Proposition 4 *Assume an exogenously given income tax rate $t > 0$. In addition, assume that the government weakly values wealth distribution, $\bar{g}^W \leq 1$. In that case, a sufficient condition for a positive wealth tax is that*

$$\bar{\varepsilon}^I_{I,1-\tau}+\bar{\xi}^I_{I,1-\tau} < 0, \tag{30}$$

Proof. *Substitute the restriction $\bar{g}^W \leq 1$ into inequality* (30) *to arrive at :*

$$1-\bar{g}^W \geq 0 > \frac{t\bar{I}}{\bar{W}}\left(\bar{\varepsilon}^I_{I,1-\tau}+\bar{\xi}^I_{I,1-\tau}\right).$$

Cancelling out $\frac{t\bar{I}}{\bar{W}}$ (which are all positive) from the right-hand side results in (30) ■

Interestingly, Proposition (4) describes a sufficient condition for a non-zero optimal wealth tax that does not depend on the government's desire to reduce wealth inequality. Most previous research that calls for a positive wealth tax (e.g. Piketty, 2013; Saez and Zucman, 2019) as well as optimal wealth-tax models (e.g. Scheuer and Slemrod, 2021; Piketty et al., 2023), justify a positive wealth tax from the perspective of reducing inequality in the (initial) wealth distribution. Contrary, Proposition 4 shows that a positive wealth tax may be desirable even if the government is indifferent with respect to wealth inequality, $\bar{g}^W = 1$. The reason is that the wealth-tax alleviates the lock-in distortion associated with realization-based income taxes, and therefore enhances overall efficiency.

Equation (30) also provides an empirical test for the desirability of the wealth-tax. If the total cross-elasticity between taxable investment income, and the net-of-wealth tax is negative the government should optimally set $\tau > 0$. The intuition is that in that case, the wealth tax reduces the portfolio-elasticity associated with the income tax, and this effect outweights the intertemporal distortion associated with the wealth tax.

Berzins et al. (2019) study the effect of the Norwegian wealth tax on dividends paid out to owners of privately-held companies. They study variation in the wealth tax that results from changes in the valuation method of real estate valuation. They find that an increase in the wealth tax rate results in an increase in dividend payments. This finding is consistent with inequality (30) being satisfied. Their result would hence imply that a wealth tax is desirable in the Norwegian context.

4.4 Constant welfare weights

Next, we consider optimal-tax policy in which the government optimally sets both the income tax and the wealth tax. However, we simplify the model by assuming that the government assigns the same welfare weight to all investors, $g(\cdot) = g_0$. In practice, since Diamond welfare weights depend on both government's preferences for redistribution, and income effects, this assumption restricts both. In addition, we assume that the lower-bound constraint on lump-sum taxes binds such that the government is unable to finance all of its expenditure through lump-sum taxes. In the absence of this restriction, constant welfare weights would imply that the government would not use distortive tax instruments. We formalize this assumption below:

Assumption 1 *Suppose that the lower-bound constraint on the lump-sum* *transfer M binds such that lump-sum taxes are insufficient to finance exogenous expenditure, $-M_0 < E$, and by equation (28) $\bar{g} < 1$. In addition, assume (equilibrium) welfare weights are constant across the population such that $g(\cdot) = g_0$ for some constant $1 > g_0 > 0$.*

Note that this setup closely resembles the canonical Ramsey-tax model. In Ramsey-tax models the government faces a lower-bound restriction on lump-sum transfers (typically lump-sum transfers are not allowed to be negative). The government then maximizes the utility of a representative agent subject to its budget constraint. By virtue of the representative-agent setting, the government does not value redistribution, and hence, the government's objective is equivalent to minimizing the excess burden of taxation. In our setting, we do allow heterogeneous investors but we rule out equity motives for taxation by assuming welfare weights are constant across the population. Hence, in our model the objective of the government also condenses to minimizing the distortionary cost of taxation.

Assumption 1 simplifies the optimal-tax equations significantly, since the left-hand side of both the optimal-income tax expression (26), and the wealth-tax expression (27) reduces to $1 - g_0$. Hence, a necessary condition for an internal optimum, $t, \tau > 0$ is that the marginal excess burden of each tax instrument is equal:

$$\frac{t}{1-t}\left(\bar{\varepsilon}^{I}_{I,1-t} + \bar{\xi}^{I}_{I,1-t}\right) + \frac{\tau}{1-t}\frac{\bar{W}}{\bar{I}}\left(\bar{\varepsilon}^{W}_{W,1-t} + \bar{\xi}^{W}_{W,1-t}\right) = \frac{\tau}{1-\tau}\left(\bar{\varepsilon}^{W}_{W,1-\tau} + \bar{\xi}^{W}_{W,1-\tau}\right) + \frac{t}{1-\tau}\frac{\bar{I}}{\bar{W}}\left(\bar{\varepsilon}^{I}_{I,1-\tau} + \bar{\xi}^{I}_{I,1-\tau}\right). \tag{31}$$

A violation of (31) in which the marginal excess burden of the income tax (the left-hand side), is larger than the marginal burden of the wealth-tax (the right-hand side) implies that there exists a welfare-improving and revenue neutral-tax reform, which reduces the income tax and increases the wealth tax, and vice versa if the marginal excess burden of the wealth tax is larger than that of the income tax.

Below we use equation (31) to derive three results. First, we show that in the absence of portfolio-responses the government will collect revenue through either the income, or the wealth tax, and derive a condition under which the government would only use the income tax in that setting. Second, adding back in portfolio responses, it derives sufficient conditions for an internal optimum $t, \tau > 0$. Finally, we derive an upper bound on the optimal wealth tax when the optimal income tax is positive.

Proposition 5 *Under assumption 1 the following holds:*

1. *Suppose the portfolio-choice elasticities $\xi_{y,1-\sigma}(R^i) = 0$ for all R^i and all $y \in \{W, I\}, \sigma \in \{t, \tau\}$ and that the required rate of return equals the market interest rate $\rho(R^i) = r$. In this case, the government will always select a corner solution $t > 0, \tau = 0$ when the distribution of (R^i, ρ^i) satisfies:*

$$\int_{R^0}^{R^1} \left(\int_{\rho_0}^{r} (R - r) + \int_{r}^{\rho^1} \frac{R - r}{1 + r} + \frac{(\rho - r)(1 + R)}{1 + r} \right) dF^{\rho}(\rho|R) dF^R(R) >$$
$$\int_{R^0}^{R^1} \int_{r}^{\rho^1} \frac{r^2(1 + R)}{1 + r} dF^{\rho}(\rho|R) dF^R(R). \tag{32}$$

   Vice versa it will always choose a corner solution $t = 0, \tau > 0$ when the right-hand side of equation (32) exceeds the left-hand side.

2. *Now allow for portfolio-choice according to our model, $\xi_{y,1-\sigma}(R^i) \neq 0$, and assume the required rate of return $\rho(R^i)$ satisfies (8). In that case, the government will select an internal equilibrium $t, \tau > 0$ when the inequality (32) is satisfied and additionally, the following inequality evaluated at $(t, \tau) = (t^*, 0)$ holds:*

$$\frac{\bar{\varepsilon}^I_{I,1-t} + \bar{\xi}^I_{I,1-t}}{(1 - t^*)\bar{I}} > \frac{\bar{\varepsilon}^I_{I,1-\tau} + \bar{\xi}^I_{I,1-\tau}}{\bar{W}}, \tag{33}$$

   where t^* denotes the optimal income-tax rate when τ is restricted to be zero.

3. *When (32) is satisfied the optimal wealth-tax lies strictly below the wealth tax that restores capital-market efficiency, $\hat{\tau}$.*

Proof. The proof can be found in the Appendix. ■

Part 1 of Proposition 1 shows that in the absence of an effect of taxation on portfolio choice, the government will either set the wealth tax rate or the income tax rate equal to zero. To see why, note that both the income and the wealth tax affect intertemporal decision making through their effect on the after-tax rate of return $\tilde{r}$. Therefore, from a tax-revenue perspective the question which instrument the government applies depends on whether for a given reduction in $\tilde{r}$ the wealth tax or the income tax yields more tax revenue. The answer to this question depends on the presence of excess returns. The income tax taxes the rate of return but does not tax the principal investment. On the other hand, the wealth tax taxes both the rate of return and the principal and thus yields more tax revenue, but also a higher intertemporal distortion. The income tax is more desirable when returns are large relative to the principal (i.e. when the economy contains positive excess returns). The wealth tax is more desirable when returns are small relative to excess returns (i.e. when the economy contains negative excess returns). An additional benefit of the wealth tax is that it yields revenue in both period 1 and 2, whereas investors that keep their initial investment only pay income taxes in period 2.

This intuition is formalized in equation (32). The left-hand side represents the amount of excess returns in the economy. The first term represents the excess return as a result of the difference between the initial capital gain R^i and the market interest rate r for those that sell the initial asset. The second term represents the same difference for investors that keep their initial asset. Note that this term is discounted by $1 + r$, since investors that keep their initial investment only pay taxes on their initial capital gain in period 2. The final term describes the excess return on the initial investment attained in period 1. The right-hand side represents the revenue loss associated with the fact that the income tax only attains revenue in period 2 whereas the wealth tax collects revenue in both periods.

We consider it likely that equation (32) is satified in the real world. There is empirical evidence from the US which suggests that the economy yields ample excess returns. For instance, Barkai (2020) find that pure profits represent around a 10 percent share of GDP or close to 50 percent of the normal return to capital. On the other hand, the right-hand side of (32) is proportional to r^2, which is small when the market-interest rate is small. Hence, in the absence of portfolio responses the government would optimally not use the wealth tax in our model.

However, part 2 of the Proposition shows that the government should nevertheless optimally set a positive wealth tax when we consider portfolio-response margins and inequality (33) holds. The left-hand side of (33) represents the ratio of the total distortion of the income tax on the income-tax base relative to the remaining income-tax base after taxing it at rate t^*. Here the total distortion of the income-tax is the sum of the intertemporal elasticity of the income-tax base with respect to income taxes and the portfolio elasticity of the income-tax base with respect to income taxes. The right-hand side represents the ratio of the total distortion of the wealth tax on the income-tax base and the wealth-tax base. Hence, the inequality states that the government should optimally apply the wealth tax when evaluated at $\tau = 0$, the wealth tax yields a smaller distortion relative to its tax base than the income tax.

In the absence of portfolio-responses ($\xi_{y,1-\sigma} = 0$) inequality (32) and (33) are mutually exclusive. However, portfolio-responses make it more likely that (33) is satisfied since they make the income tax more distortive ($\bar{\xi}^I_{I,1-t} > 0$) and the wealth tax less distortive ($\bar{\xi}^I_{I,1-t} < 0$). Therefore, if excess returns are not too large and/or portfolio-elasticities are sufficiently large both inequalities can be satisfied simultaneously.

Finally to understand part 3 note that by assuming that inequality (32) holds we have implicitly assumed that there are sufficient excess returns, such that the income tax generates revenue at a lower intertemporal distortion than the wealth tax. Hence, the only role for the wealth tax is to restore capital-market efficiency. As we show in Proposition 2 setting $\tau = \hat{\tau}$ eliminates all distortions on the portfolio-choice margin, implying that at this point the distortion of the wealth tax exceeds the distortion of the income tax per unit of tax revenue.

Proposition 4 provides clear policy guidance on the optimal level of the wealth tax. When (32) holds, the optimal wealth tax rate is in the interval $[0, \hat{\tau})$. This interval is independent of the behavioral responses by investors. When additionally equation (33) holds the optimal wealth tax is in the interval $(0, \hat{\tau})$.

Because the optimal wealth tax is strictly smaller than $\hat{\tau}$ some capital-market inefficiency will remain. Intuitively, eliminating all inefficient lock-in is inefficient, because it yields a too large intertemporal distortion.

We have derived clear bounds on the wealth tax in a setting where we restricted government's preferences for redistribution through Assumption 1. In the next subsection we consider to which extend these results generalize to the a setting with other preferences for redistribution.

4.5 The full problem

We now consider optimal-tax expressions (26) and (27) without imposing assumptions on the welfare weights. We are particularly interested in understanding which parts of Proposition (5) continue to apply when the government values redistribution from rich to poor. For general welfare weights and internal equilibria the marginal excess burden of the income and wealth tax must satisfy the following relationship:

$$\frac{t}{1-t}\left(\bar{\varepsilon}^{I}_{I,1-t}+\bar{\xi}^{I}_{I,1-t}\right)+\frac{\tau}{1-t}\frac{\bar{W}}{\bar{I}}\left(\bar{\varepsilon}^{W}_{W,1-t}+\bar{\xi}^{W}_{W,1-t}\right)=$$
$$\frac{\tau}{1-\tau}\left(\bar{\varepsilon}^{W}_{W,1-\tau}+\bar{\xi}^{W}_{W,1-\tau}\right)+\frac{t}{1-\tau}\frac{\bar{I}}{\bar{W}}\left(\bar{\varepsilon}^{I}_{I,1-\tau}+\bar{\xi}^{I}_{I,1-\tau}\right)+\left(\bar{g}^{W}-\bar{g}^{I}\right).\tag{34}$$

The only difference between this expression and the case with constant welfare weights, equation (31) is the final term on the right-hand side of this expression which comprises the difference between the average welfare weight weighted by wealth, and the average welfare weight weighted by income. This equation naturally divides the discussion into two cases. The first is the case where the government is more concerned with income-inequality than with wealth inequality $\bar{g}^I < \bar{g}^W$. In this case, the optimal marginal excess burden of the income tax exceeds the marginal excess burden of the wealth tax. Hence, relative to the case with constant welfare weights, this implies that the government will set higher income tax rates, and lower wealth tax rates. As a result, inequality (32) continues to describe a sufficient condition for a positive optimal income tax $t > 0$. Contrary, inequality (33) no longer describes a sufficient condition for a positive optimal wealth tax. To see this, note that if the left-hand side of (33) exceeds the right-hand side by a small amount δ, this implies that the marginal excess burden of the income tax exceeds that of the wealth tax. However, given that the government is more concerned with income than with wealth inequality, such a difference may not be suboptimal. Finally, by the same reasoning it should be clear that $\hat{\tau}$ continues to describe an upper bound on the wealth tax when inequality (32) is satisfied.

The other case of interest is when the government is more concerned with wealth inequality than income inequality $\bar{g}^I > \bar{g}^W$. In that case, inequality (32) is not sufficient to ensure that $t > 0$, inequality (33) is sufficient to ensure $\tau > 0$, and the optimal wealth tax rate may exceed $\hat{\tau}$ even when

(32) holds. The properties of optimal-tax policy thus, in general, crucially depend on the type of inequality the government is most concerned with.

Nevertheless, there is one case where it is possible to arrive at strong conclusions regarding optimal-tax policy. When the government is strongly inequality-averse such that its preferences can be represented by a Rawlsian welfare function, and the taxable income and wealth of this investor is (approximately) zero, the welfare weight weighted by income and wealth equals zero $\bar{g}^I = \bar{g}^W = 0$.[9] In this case, equation (31) and (34) are equivalent and hence, Proposition 5 applies. Hence, both a government that has no preference for redistribution as described above, and a government with extremely strong preferences for redistribution agree that the optimal-wealth tax should not exceed $\hat{\tau}$ when inequality (32) is satisfied.

5 Extension: lower capital gains taxes in period 2

In our model we have assumed that capital gains are taxed at the same rate in both periods. In reality, many tax systems have features that result in a lower tax rate on postponed capital gains. For instance, in the US long-term capital gains on assets held less than a year are taxed at a maximum rate of 37 percent. Capital gains on assets held for more than a year are taxed at a top rate of 20 percent. Many other OECD countries allow for similar discounts on long-term investments (see e.g. Harding and Marten, 2018 for an overview among OECD countries). In addition, the US tax system exhibits a step-up in which, upon the death of an investor, his heirs are exempted from paying capital-gains taxes obtained over the life-time of the deceased. This further erodes the effective tax rate on long-term capital gains.

In this section we incorporate this feature into our model as an extension by multiplying the capital-gains tax base in period 2 with a factor $0 \leq \kappa \leq 1$. That is, period 2 income is given by:

$$I_2^i \equiv (1 - \phi^i)\kappa((1 + \rho^i)W_1^i - 1) + rB. \tag{35}$$

[9]Note that technically our model does not allow for a taxable wealth of exactly zero, since period-1 wealth equals $W_1 = 1 + R^i$, and the locked-in capital gain R^i is assumed to be non-negative. However, if the wealth level of the poorest investor is very small relative to the wealth level of the average investor, $\bar{g}^W \approx 0$ for Rawlsian preferences.

At the extremes, when $\kappa = 1$ equation (35) coincides with the definition of period 2 income (3). When $\kappa = 0$ long-term capital gains are entirely exempt from taxation.

Reducing the capital-gains tax on long-term investments increases the incentive to postpone the realization of capital gains. This is reflected in the intertemporal budget constraint which, when substituting in the new definition of period-2 income (35), can be written as:

$$\underbrace{C_1^i + \frac{C_2^i}{1+\tilde{r}}}_{\text{NPV Consumption}} \underbrace{+ tR^i \left(\frac{\kappa + \phi^i(1+\tilde{r}-\kappa)}{1+\tilde{r}} \right)}_{\text{NPV tax on } R^i} + \underbrace{\tau W_1^i}_{\text{Initial Wealth tax}} = \underbrace{\left(\frac{1+\tilde{\rho}^i + \phi^i(\tilde{r}-\tilde{\rho}^i) + t(1-\phi^i)(1-\kappa)\rho}{1+\tilde{r}} \right) W_1}_{\text{NPV Asset}} + M, \tag{36}$$

where $\tilde{r}$ and $\tilde{\rho}$ are defined as in the base model. Reducing κ below 1 has two effects on budget constraint (36) which both materialize when the investor keeps his initial investment ($\phi^i = 0$). First, on the left-hand side the net-present value of the tax on the initial capital gain R^i reduces, since reducing κ reduces the effective tax rate on long-run capital gains. Second, on the right-hand side the net-present value of the initial asset increases because reducing κ reduces the tax on the rate of return attained in period 1 (ρ^i).

Taking the derivative with respect to ϕ^i allows us to derive the required rate of return:

$$\rho(R^i) \equiv \frac{(1-t-\tau)(1+(1-t)R^i)r + tR^i(\tau+\kappa-1)}{(1-\kappa t-\tau)(1+R^i)}. \tag{37}$$

Reducing κ below 1 reduces the required rate of return since reducing κ simultaneously reduces the final term of the numerator, while increasing the denominator. Hence, the discount on long-term capital gains results in an reduction in capital-market efficiency. Additionally, when $\kappa < 1$ the required rate of return lies below the market-interest rate even among investors whose initial locked-in capital gain $R^i = 0$. The reason is that for these investors postponing the capital gain reduces the tax on returns attained in period 1.

Solving $\rho(\cdot) = r$ for the wealth tax τ yields the wealth tax that restores capital-market efficiency:

$$\hat{\tau}(R^i) \equiv \frac{(1-\kappa)r}{1+r}\frac{1+R^i}{R^i} + \frac{(1-t)r+1-\kappa}{1+r}. \tag{38}$$

Note that relative to our original model, the wealth tax that eliminates inefficient lock-in now depends on the initial capital gain R^i. Hence, it is no longer possible to eliminate lock-in for all investors, without simultaneously introducing a new distortion in which for some investors the required rate of return exceeds the market-interest rate $\rho(R^i) > r$.

Interestingly, $\hat{\tau}(R^i)$ reduces in R^i. To understand this, remember that the tax benefit associated with selling the initial capital gain is that the tax on the initial capital mechanically reduces taxable wealth in period 2. When $R^i = 0$ this mechanism does not apply, since realization does not create a tax burden in period 1 if there is no initial capital gain. On the other hand, unlike in the original model, investors with $R^i = 0$ do face a fiscal cost when selling the asset since it effectively converts a taxable return, which is taxed at a favorable rate, into taxable interest which is taxed at the regular rate.

Most interesting is perhaps the wealth tax rate that eliminates inefficient lock-in among those with large initial capital gains. The reason is that the efficiency cost associated with inefficient lock-in scale with the size of the lock-in. To find this tax rate we take the limit of (38) as R^i approaches infinity:

$$\lim_{R^i \to \infty} \hat{\tau}(R^i) \equiv \frac{(1-t)r}{1+r} + 1 - \kappa. \tag{39}$$

Relative to the wealth tax that restores capital-market efficiency in the base model (12), reduced capital-gains taxation on long-run capital gains increases the wealth tax by $1-\kappa$. This has a substantial impact. For instance, consider the US where long-run capital-gains are taxed at most 20 percent, and short-run capital gains at most 37 % such that $\kappa = 20/37 = 0.54$. In this case, this last term amounts to $1 - .54 = .46$. Hence, if under uniform taxation the wealth tax that restores capital-market efficiency is 1.6 percent, than the wealth tax that restores capital-market efficiency among those with large capital gains is $1.6 + 43 = 44.6$ percent. Hence, if the capital gains tax contains additional imperfections that diminish its ability to tax long-run capital gains, the optimal wealth-tax rate could be significantly larger than in our baseline setting.

6 Concluding remarks

Our study demonstrates that when realization-based taxes are in effect, a wealth tax can effectively address two key objectives: it has the potential to reduce inequality while simultaneously enhancing capital market efficiency. The efficiency role of the wealth tax is attributed to that it creates an incentive for individuals to realize capital gains, especially when the rate of return on those assets is below the market-rate of return. Therefore, the wealth tax has a role to play even if the government has no social preference for reducing wealth inequality. In general, optimal tax policy balances obtaining tax revenue with efficiency losses related to lock-in and intertemporal distortions.

We show that a wealth tax reduces the cost of paying taxes early on, instead of postponing the tax burden. The reason is that paying taxes immediately reduces net wealth mechanically. Looking beyond the capital-gains tax, this mechanism may also affect efficiency losses associated with other tax instruments such as taxes on real estate.

References

Advani, Arun, and Hannah Tarrant (2021) 'Behavioural responses to a wealth tax.' *Fiscal Studies* 42(3-4), 509–537

Agersnap, Ole, and Owen Zidar (2021) 'The tax elasticity of capital gains and revenue-maximizing rates.' *American Economic Review: Insights* 3(4), 399–416

Alvaredo, Facundo, and Emmanuel Saez (2009) 'Income and wealth concentration in spain from a historical and fiscal perspective.' *Journal of the European Economic Association* 7(5), 1140–1167

Auerbach, Alan J (1988) 'Retrospective capital gains taxation.' NBER Working Paper No. 2792 Cambridge, MA

Bach, Laurent, Antoine Bozio, Brice Fabre, Arthur Guillouzouic, Claire Leroy, and Clément Malgouyres (2021) 'Follow the money! why dividends overreact to flat-tax reforms.' mimeo Paris School of Economics Paris

Barkai, Simcha (2020) 'Declining labor and capital shares.' *The Journal of Finance* 75(5), 2421–2463

Berzins, Janis, Øyvind Bøhren, and Bogdan Stacescu (2019) 'Shareholder illiquidity and firm behavior: Financial and real effects of the personal wealth tax in private firms.' ECGI Working Paper No 646/2019

Boadway, Robin, and Kevin Spiritus (2021) 'Optimal taxation of normal and excess returns to risky assets.' Tinbergen Institute Discussion Paper 2021-025/VI Amsterdam

Bozio, Antoine, Bertrand Garbinti, Jonathan Goupille-Lebret, Malka Guillot, and Thomas Piketty (forthcoming) 'Predistribution vs. redistribution: Evidence from france and the united states.' *American Economic Journal: Applied Economics*

Bruil, Arjan, Céline Van Essen, Wouter Leenders, Arjan Lejour, Jan Möhlmann, and Simon Rabaté (2022) 'Inequality and redistribution in the netherlands.' CPB Discussion Paper No 436 The Hague

Brülhart, Marius, Jonathan Gruber, Matthias Krapf, and Kurt Schmidheiny (2022) 'Behavioral responses to wealth taxes: Evidence from switzerland.' *American Economic Journal: Economic Policy* 14(4), 111–150

Diamond, Peter A (1975) 'A many-person ramsey tax rule.' *Journal of Public Economics* 4(4), 335–342

Eisinger, Jesse, Jeff Ernsthausen, and Paul Kiel (2021) 'The secret irs files: trove of never-before-seen records reveal how the wealthiest avoid income tax.' ProPublica

Ferey, Antoine, Benjamin Lockwood, and Dmitry Taubinsky (2021) 'Sufficient statistics for nonlinear tax systems with preference heterogeneity.' NBER Working Paper No. 29582 Cambridge, MA

Garbinti, Bertrand, Jonathan Goupille-Lebret, Mathilde Muñoz, Stefanie Stantcheva, and Gabriel Zucman (2023) 'Tax design, information, and elasticities: Evidence from the french wealth tax.' Technical Report, National Bureau of Economic Research

Gerritsen, Aart, Bas Jacobs, Alexandra V Rusu, and Kevin Spiritus (2020) 'Optimal taxation of capital income with heterogeneous rates of return.' Tinbergen Institute Discussion Paper 2020-038/VI Amsterdam

Guvenen, Fatih, Gueorgui Kambourov, Burhan Kuruscu, Sergio Ocampo, and Daphne Chen (2023) 'Use it or lose it: Efficiency and redistributional effects of wealth taxation.' *The Quarterly Journal of Economics*

Guyton, John, Patrick Langetieg, Daniel Reck, Max Risch, and Gabriel Zucman (2020) 'Tax evasion by the wealthy: Measurement and implications.' In 'Measuring and Understanding the Distribution and Intra/Inter-Generational Mobility of Income and Wealth' (University of Chicago Press)

Hall, Robert E, and Susan E Woodward (2010) 'The burden of the nondiversifiable risk of entrepreneurship.' *American Economic Review* 100(3), 1163–1194

Harding, Michelle, and Melanie Marten (2018) 'Statutory tax rates on dividends, interest and capital gains: The debt equity bias at the personal level.' OECD Taxation Working Papers No. 34 Paris

Hemel, Daniel (2019) 'Taxing wealth in an uncertain world.' *National Tax Journal* 72(4), 755–776

Jacobs, Bas (2018) 'The Marginal Cost of Public Funds is One.' *International Tax and Public Finance* 25, 883–912

Jakobsen, Katrine, Kristian Jakobsen, Henrik Kleven, and Gabriel Zucman (2018) 'Wealth taxation and wealth accumulation: Theory and evidence from Denmark.' NBER Working Paper 24371 Cambridge, MA

Jordà, Òscar, Katharina Knoll, Dmitry Kuvshinov, Moritz Schularick, and Alan M Taylor (2019) 'The rate of return on everything, 1870–2015.' *The Quarterly Journal of Economics* 134(3), 1225–1298

Londoño-Vélez, Juliana, and Javier Ávila-Mahecha (2021) 'Enforcing wealth taxes in the developing world: Quasi-experimental evidence from colombia.' *American Economic Review: Insights* 3(2), 131–148

Piketty, Thomas (2013) *Capital in the Twenty-First Century* (Cambridge, MA: Harvard University Press)

Piketty, Thomas, Emmanuel Saez, and Gabriel Zucman (2018) 'Distributional national accounts: methods and estimates for the united states.' *The Quarterly Journal of Economics* 133(2), 553–609

_ (2023) 'Rethinking capital and wealth taxation.' *Oxford Review of Economic Policy* 39(3), 575–591

Ring, Marius Alexander Kalleberg (2020) 'Wealth taxation and household saving: Evidence from assessment discontinuities in norway.' SSRN No. 3716257

Saez, Emmanuel (2017) 'Taxing the rich more: Preliminary evidence from the 2013 tax increase.' *Tax Policy and the Economy* 31(1), 71–120

Saez, Emmanuel, and Gabriel Zucman (2019) 'Progressive wealth taxation.' *Brookings Papers on Economic Activity* 2019(2), 437–533

Saez, Emmanuel, Danny Yagan, and Gabriel Zucman (2021) 'Capital gains withholding.' mimeo Berkeley, CA

Sandmo, Agnar (1977) 'Portfolio theory, asset demand and taxation: Comparative statics with many assets.' *The review of economic studies* 44(2), 369–379

Scheuer, Florian, and Joel Slemrod (2021) 'Taxing our wealth.' *Journal of Economic Perspectives* 35(1), 207–230

Seim, David (2017) 'Behavioral responses to wealth taxes: Evidence from sweden.' *American Economic Journal: Economic Policy* 9(4), 395–421

Thoresen, Thor O, Marius AK Ring, Odd E Nygård, and Jon Epland (2022) 'A wealth tax at work.' *CESifo Economic Studies* 68(4), 321–361

Yagan, Danny (2023) 'What is the average federal individual income tax rate on the wealthiest americans?' *Oxford Review of Economic Policy* 39(3), 438–450

Young, Cristobal, Charles Varner, Ithai Z Lurie, and Richard Prisinzano (2016) 'Millionaire migration and taxation of the elite: Evidence from administrative data.' *American Sociological Review* 81(3), 421–446

A Proofs

A.1 Proof to Lemma 1

Proof. Starting with the definition of taxable income and wealth, $I(\cdot)$ and $W(\cdot)$ we rewrite them in terms of $\phi(\cdot), C_1(\cdot)$ using equations (1)-(3) as follows:

$$\begin{aligned} I(\cdot) &= \phi(\cdot)R^i + \frac{(1-\phi(\cdot))((1+\rho^i)W_1^i - 1) + rB(\cdot)}{1+r} \\ &= \begin{cases} \frac{R^i + r\left((1-t)R^i + (1-\tau)(1+R^i) + M - C_1(\cdot)\right)}{1+r} & \text{if} \quad \phi(\cdot) = 1 \\ \frac{(1+\rho^i)R^i + \rho^i + r\left(M - \tau(1+R^i) - C_1(\cdot)\right)}{1+r} & \text{if} \quad \phi(\cdot) = 0 \end{cases}, \end{aligned} \tag{40}$$

$$\begin{aligned} W(\cdot) &= W_1^i + \frac{(1-\phi(\cdot))(1+\rho^i)W_1^i + (1+r)B^i}{1+r} \\ &= \begin{cases} (1+R^i)(2-\tau) - tR^i + M - C_1(\cdot) & \text{if} \quad \phi(\cdot) = 1 \\ (1+R^i)(1-\tau) + M - C_1(\cdot) + \frac{(1+\rho^i)(1+R^i)}{1+r} & \text{if} \quad \phi(\cdot) = 0 \end{cases} \end{aligned} \tag{41}$$

Rewriting in terms of $\phi(\cdot), C_1(\cdot)$ is useful, because i.) equation (8) reveals the left and right-limiting behavior of $\phi(\cdot)$ as ρ^i approaches $\rho(\cdot)$, and ii.) because $C_1(\cdot)$ is continuous around the required rate of return $\rho^i = \rho(\cdot)$. To see the latter point, note first that the indirect utility function $V(\rho^i, R^i)$ cannot exhibit a discontinuity around the point $\rho^i = \rho(\cdot)$ irrespective of the value of R^i, since at $\rho^i = \rho(\cdot)$ investors are by the definition of the required rate of return $\rho(\cdot)$ indifferent with respect to ϕ^i. Note second that for a given indirect utility function $V(\cdot)$, $C_1(\cdot), C_2(\cdot)$ are implicitly determined by the consumption Euler equation (6) and the definition of indirect utility $V(\rho^i, R^i) = U(C_1(\rho^i, R^i), C_2(\rho^i, R^i))$, which are both continuous and monotonic in consumption. Therefore, follows that both $C_1(\cdot), C_2(\cdot)$ are continuous around $\rho^i = \rho(\cdot)$.

Now to resolve the limits in ΔI, (ΔW) evaluate the top and bottom case of equation (40) (equation (41)) at $\rho^i = \rho(\cdot)$, substitute in the definition of $\rho(\cdot)$ (equation (8)) and subtract the bottom case from the top case to arrive at expressions (21) and (22). ■

A.2 Proof to Proposition 3

Proof. First substitute the envelope conditions, $\frac{\partial V}{\partial M} = U_{C_1}$, $\frac{\partial V}{\partial t} = U_{C_1} I(\cdot)$, $\frac{\partial V}{\partial \tau} = U_{C_1} W(\cdot)$, the Slutsky equations:

$$\begin{aligned}
\frac{\partial I}{\partial t} &= \frac{\partial I}{\partial t}|_{U=U_0} - \frac{\partial I}{\partial M} I \\
\frac{\partial I}{\partial \tau} &= \frac{\partial I}{\partial \tau}|_{U=U_0} - \frac{\partial I}{\partial M} W \\
\frac{\partial W}{\partial t} &= \frac{\partial W}{\partial t}|_{U=U_0} - \frac{\partial W}{\partial M} I \\
\frac{\partial W}{\partial \tau} &= \frac{\partial W}{\partial \tau}|_{U=U_0} + \frac{\partial W}{\partial M} W,
\end{aligned}$$

and the definition for welfare weights (23) into the optimal-tax expressions (16)-(18) and rearrange to arrive at (28) and:

$$\begin{aligned}
\mathbb{E}[(g-1)\, I] \;\; &\geq \;\; \mathbb{E}\left[t \frac{\partial I}{\partial t}|_{U=U_0} + \tau \frac{\partial W}{\partial t}|_{U=U_0}\right] \qquad (42) \\
&\quad + \mathbb{E}\left[\frac{\partial \rho}{\partial t} \left(t \Delta I + \tau \Delta W\right) f^{\rho} \;\middle|\; \rho^i = \rho(\cdot)\right], \\
\mathbb{E}[(g-1)\, W] \;\; &\geq \;\; \mathbb{E}\left[t \frac{\partial I}{\partial \tau}|_{U=U_0} + \tau \frac{\partial W}{\partial \tau}|_{U=U_0}\right] \qquad (43) \\
&\quad + \mathbb{E}\left[\frac{\partial \rho}{\partial \tau} \left(t \Delta I + \tau \Delta W\right) f^{\rho} \;\middle|\; \rho^i = \rho(\cdot)\right].
\end{aligned}$$

Now divide both sides of equation (42) by $-\mathbb{E}[I]$ and both sides of (43) by $-\mathbb{E}[W]$, substitute in the elasticities (24)-(25) and rearrange the terms to arrive at:

$$\begin{aligned}
\frac{1}{\mathbb{E}[I]} \mathbb{E}[(1-g)\, I] \;\; &\leq \;\; \frac{t}{1-t} \frac{\mathbb{E}[(\varepsilon_{I,1-t} + \xi_{I,1-t})\, I]}{\mathbb{E}[I]} \qquad (44) \\
&\quad + \frac{\tau \mathbb{E}[W]}{(1-t)\mathbb{E}[I]} \frac{\mathbb{E}[(\varepsilon_{W,1-t} + \xi_{W,1-t})\, W]}{\mathbb{E}[W]}, \\
\frac{1}{\mathbb{E}[W]} \mathbb{E}[(1-g)\, W] \;\; &\leq \;\; \frac{t \mathbb{E}[I]}{(1-\tau)\mathbb{E}[W]} \frac{\mathbb{E}[(\varepsilon_{I,1-\tau} + \xi_{I,1-\tau})\, I]}{\mathbb{E}[I]} \qquad (45) \\
&\quad + \frac{\tau}{1-\tau} \frac{\mathbb{E}[(\varepsilon_{W,1-\tau} + \xi_{W,1-\tau})\, W]}{\mathbb{E}[W]},
\end{aligned}$$

Finally, substitute in the definitions of weighted averages to arrive at (26) and (27).

For the sign-restrictions, note that all compensated intertemporal elasticities $\epsilon_{y,1-\sigma}$ are positive, since increasing the net-of-income (net-of-wealth) tax rate increases the after-tax interest rate $\tilde{r}$ in the Euler equation (6). The sign of the portfolio elasticities is determined by the product $-\frac{\partial\rho(\cdot)}{\partial\sigma}\Delta y$. By taking the derivative of the definition of $\rho(\cdot)$ (equation (8)) with respect to t, τ we find that when $\tau < \hat{\tau}$, $\frac{\partial\rho(\cdot)}{\partial t} < 0$ and $\frac{\partial\rho(\cdot)}{\partial\tau} \geq 0$. Again, imposing $\tau < \hat{\tau}$ we have that $\Delta I > 0$ and $\Delta W \leq 0$. ■

A.3 Proof to Proposition 5

Proof. To prove part 1 we will show that the marginal excess burden of the income tax is strictly smaller than the marginal excess burden of the wealth tax when $\xi_{y,1-\sigma} = 0$, $\rho(\cdot) = r$ and inequality (32) is satisfied. That is, we want to show that under these assumptions the left-hand side of (31) is smaller than the right-hand side for all values t, τ:

$$\frac{t}{1-t}\bar{\varepsilon}^{I}_{I,1-t} + \frac{\tau}{1-t}\frac{\bar{W}}{\bar{I}}\bar{\varepsilon}^{W}_{W,1-t} < \frac{\tau}{1-\tau}\bar{\varepsilon}^{W}_{W,1-\tau} + \frac{t}{1-\tau}\frac{\bar{I}}{\bar{W}}\bar{\varepsilon}^{I}_{I,1-\tau}. \tag{46}$$

To simplify we use the fact that the intertemporal effects of the wealth and income tax base are related. First, intertemporal adjustment is mediated through changes in investment in the risk-free asset $B(\cdot)$. Second, compensated changes of taxation only affect the intertemporal trade-off through their impact on the after-tax interest rate $\tilde{r}$. That is, the following chain-rule relationship must hold:

$$\begin{aligned} \frac{\partial y}{\partial\sigma}|_{U=U_0} &= \frac{\partial y}{\partial B}\frac{\partial\tilde{r}}{\partial\sigma}\frac{\partial B}{\partial\tilde{r}}|_{U=U_0}, \\ \frac{\partial I}{\partial t}|_{U=U_0} = -\frac{r^2}{1+r}\frac{\partial B}{\partial\tilde{r}}|_{U=U_0} \quad &, \quad \frac{\partial I}{\partial\tau}|_{U=U_0} = -r\frac{\partial B}{\partial\tilde{r}}|_{U=U_0}, \\ \frac{\partial W}{\partial t}|_{U=U_0} = -r\frac{\partial B}{\partial\tilde{r}}|_{U=U_0} \quad &, \quad \frac{\partial W}{\partial\tau}|_{U=U_0} = -(1+r)\frac{\partial B}{\partial\tilde{r}}|_{U=U_0} \end{aligned} \tag{47}$$

Substituting (47) together with the definition of the compensated elasticities (24) into (46) we arrive at:

$$\frac{t\frac{r^2}{1+r}\mathbb{E}\left[\frac{\partial B}{\partial\tilde{r}}|_{U=U_0}\right]}{\bar{I}} + \frac{\tau r\mathbb{E}\left[\frac{\partial B}{\partial\tilde{r}}|_{U=U_0}\right]}{\bar{I}} < \frac{\tau(1+r)\mathbb{E}\left[\frac{\partial B}{\partial\tilde{r}}|_{U=U_0}\right]}{\bar{W}} + \frac{tr\mathbb{E}\left[\frac{\partial B}{\partial\tilde{r}}|_{U=U_0}\right]}{\bar{W}}. \tag{48}$$

Cancelling out $\mathbb{E}\left[\frac{\partial B}{\partial\tilde{r}}|_{U=U_0}\right]$ from all terms, and rewriting the right-hand side such that the numerator is equal to the numerator on the left-hand side, we arrive at:

$$\begin{aligned} \frac{t\frac{r^2}{1+r}+\tau r}{\bar{I}} &< \frac{t\frac{r^2}{1+r}+\tau r}{\frac{r\bar{W}}{1+r}}, \\ \bar{I} &> \frac{r\bar{W}}{1+r}. \end{aligned} \tag{49}$$

To further simplify note that average income and wealth, $\bar{I}$ and $\bar{W}$ can be rewritten using the definition of taxable income and wealth:

$$\begin{aligned} \bar{I} &= \mathbb{E}\left[I_1+\frac{I_2}{1+r}\right], \\ &= \mathbb{E}\left[\frac{(1+r)\phi R^i+rB+(1-\phi)\left((1+\rho^i)W_1^i-1\right)}{1+r}\right], \\ &= \int_{R_0}^{R_1}\left(\int_{\rho_0}^{\rho(R^i)}\frac{(1+r)R+rB}{1+r}+\int_{\rho(R^i)}^{\rho^1}\frac{rB+(1+\rho^i)W_1^i-1}{1+r}\right)dF^\rho dF^R \end{aligned} \tag{50}$$

$$\begin{aligned} \bar{W} &= \mathbb{E}\left[W_1^i+\frac{W_2}{1+r}\right], \\ &= \mathbb{E}\left[\frac{(W_1^i+B)(1+r)+(1-\phi)(1+\rho^i)W_1^i}{1+r}\right], \\ &= \int_{R_0}^{R_1}\left(\int_{\rho_0}^{\rho(R^i)}W_1^i+B+\int_{\rho(R^i)}^{\rho^1}\frac{(1+r)(B+W_1^i)+(1+\rho^i)W_1^i}{1+r}\right)dF^\rho dF^R \end{aligned} \tag{51}$$

where in the second step we used equation (2) and (3) to substitute for W_2, I_2, and in the third step we split between investors who sell their initial investors $\rho^i < \rho(R^i)$, and investors who keep their current investment. Substituting (50) and (51) into (48) and, simplifying we arrive at (32).

To prove part 2 we need to show that in our model inequality (32) and inequality (33) are jointly sufficient to guarantee an internal equilibrium. That is, we need to rule out the following 3 corner solutions: 1. $t=\tau=0$, 2. $\tau=0, t>0$ and 3. $t=0, \tau>0$.

Corner solution 1 is ruled out by Assumption 1 since this solution does not satisfy the government budget constraint. We rule out corner solution 2 by showing that when $t=0$ for all values of $\tau>0$ the marginal excess burden of the wealth tax exceeds the marginal excess burden of the income tax:

$$\begin{aligned}\frac{\tau\bar{W}}{\bar{I}}\left(\bar{\varepsilon}^{W}_{W,1-t}+\bar{\xi}^{W}_{W,1-t}\right) &< \frac{\tau}{1-\tau}\left(\bar{\varepsilon}^{W}_{W,1-\tau}+\bar{\xi}^{W}_{W,1-\tau}\right),\\ \frac{\tau\bar{W}}{\bar{I}}\bar{\varepsilon}^{W}_{W,1-t} &< \frac{\tau}{1-\tau}\bar{\varepsilon}^{W}_{W,1-\tau},\end{aligned}$$

where the second step follows from the fact that $t=0$ implies that $\Delta W(\cdot)$ and hence the portfolio elasticity $\xi_{W,1-\sigma}=0$ by equations (22) and (25). Finally, we arrive at inequality (32) by i.) substituting in (50) and (51) for $\bar{I},\bar{W}$, ii.) substituting in (47) for the compensated elasticities $\varepsilon_{W,1-\sigma}$ and iii.) noting that $t=0$ implies $\rho(\cdot)=r$ by equation (8).

We rule out corner solution 3 by showing that when t is chosen optimally conditional on $\tau=0$, the marginal excess burden of the income tax exceeds the marginal excess burden of the wealth tax. Formally, this implies showing that the left-hand side of (31) exceeds the right-hand side evaluated at $(t,\tau)=(t^*,0)$:

$$\frac{t^*}{1-t^*}\left(\bar{\varepsilon}^{I}_{I,1-t}+\bar{\xi}^{I}_{I,1-t}\right)>\frac{t^*\bar{I}}{\bar{W}}\left(\bar{\varepsilon}^{I}_{I,1-\tau}+\bar{\xi}^{I}_{I,1-\tau}\right).$$

Reordering this expression yields inequality (33) which we have assumed to be true.

For part 3 we show that $\tau=\hat{\tau}$ implies that independent of the level of the income tax t, the marginal excess burden of the wealth tax exceeds the marginal excess burden of the income tax(i.e. the right-hand side of equation (31) exceeds the left-hand side):

$$\begin{aligned}&\frac{1}{1-t}\left(t\bar{\xi}^{I}_{I,1-t}+\frac{\hat{\tau}\bar{W}}{\bar{I}}\bar{\xi}^{W}_{W,1-t}\right)-\frac{1}{1-\hat{\tau}}\left(\hat{\tau}\bar{\xi}^{W}_{W,1-\hat{\tau}}+\frac{t\bar{I}}{\bar{W}}\bar{\xi}^{I}_{I,1-\hat{\tau}}\right)<\\ &\frac{1}{1-\hat{\tau}}\left(\hat{\tau}\bar{\varepsilon}^{W}_{W,1-\hat{\tau}}+\frac{t\bar{I}}{\bar{W}}\bar{\varepsilon}^{I}_{I,1-\hat{\tau}}\right)-\frac{1}{1-t}\left(t\bar{\varepsilon}^{I}_{I,1-t}+\frac{\hat{\tau}\bar{W}}{\bar{I}}\bar{\varepsilon}^{W}_{W,1-t}\right)\end{aligned}\quad(52)$$

where we have rewritten (31) to place all portfolio-elasticity terms on the left-hand side, and all compensated-elasticity terms on the right-hand side. We will show that the left-hand side of this expression evaluates to zero, whereas the right-hand side evaluates to a positive number. To see that the left-hand side of (52) evaluates to 0 substitute in equation (25) for the portfolio-elasticities:

$$\frac{1}{1-t}\left(t\bar{\xi}^{I}_{I,1-t}+\frac{\hat{\tau}\bar{W}}{\bar{I}}\bar{\xi}^{W}_{W,1-t}\right)-\frac{1}{1-\hat{\tau}}\left(\hat{\tau}\bar{\xi}^{W}_{W,1-\hat{\tau}}+\frac{t\bar{I}}{\bar{W}}\bar{\xi}^{I}_{I,1-\hat{\tau}}\right)=$$
$$\frac{-\mathbb{E}\left[(t\Delta I+\hat{\tau}\Delta W)\frac{\partial\rho}{\partial t}f^{\rho}\right]}{\bar{I}}+\frac{\mathbb{E}\left[(t\Delta I+\hat{\tau}\Delta W)\frac{\partial\rho}{\partial\tau}f^{\rho}\right]}{\bar{W}}, \tag{53}$$

Equation (53) evaluates to zero because $(t\Delta I+\hat{\tau}\Delta W)=0$ for all investors. To see this, we use equation (21) and (22) for $\Delta I(\cdot)$ and $\Delta W(\cdot)$:

$$\begin{aligned}(t\Delta I+\hat{\tau}\Delta W) &= \frac{t\left(rR^{i}(1-t-\hat{\tau})-t\hat{\tau}R^{i}\right)-\hat{\tau}tR^{i}(1-t)}{(1-t-\hat{\tau})(1+r)},\\ &= \frac{trR^{i}\left((1-t)-t(1-t)-(1-t)^{2}\right)}{(1-t-\hat{\tau})(1+r)^{2}}=0,\end{aligned} \tag{54}$$

where in the second step we substitute in (12) for $\hat{\tau}$ and simplify. This proves that the left-hand side of (52) equals zero.

All that remains is to show that the right-hand side of (52) is positive at $\tau=\hat{\tau}$. Note that this is equivalent to the case where inequality (46) is satisfied at $\tau=\hat{\tau}$. However, we have already shown that (46) holds for all values of τ whenever inequality (32) is satisfied as we have assumed. Hence, the right-hand side of (52) is positive when we assume inequality (32) holds.
■