Optimal Enforcement of Redistribution*

Magnus Eldrup†
Claus Thustrup Kreiner‡

Preliminary Draft: November 2025

Abstract

We incorporate taxpayer non-compliance and government enforcement into the theory of optimal non-linear income taxation and derive sufficient-statistics formulas characterizing the jointly optimal tax and enforcement scheme. The optimal tax formula is familiar, but the elasticity of taxable income becomes endogenous to enforcement. Optimal enforcement equates the marginal revenue from increased income reporting—which depends on the optimal tax rate—with the marginal public enforcement costs and welfare-weighted private compliance costs. Stronger redistributive preferences raise optimal top-income tax rates, increasing the social returns to enforcement, while simultaneously lowering the social cost of imposing compliance on high-income taxpayers. When the planner assigns zero weight to marginal consumption of evaders, optimal enforcement reduces to a simple cost–benefit rule. We calibrate the model to the U.S. economy and provide simulations that illustrate the quantitative importance of enforcement for redistribution and its interactions.

JEL Codes: H21, H26
Keywords: Optimal Income Taxation, Non-Compliance, Tax Enforcement

*We thank James Hines, Henrik Kleven, Wojciech Kopczuk, Etienne lehmann, Jukka Pirttila, Bertil Tungodden and conference participants for comments and suggestions. Funding from the Danish National Research Foundation, grant DNRF134, is gratefully acknowledged.

†University of Copenhagen. Email: magnus.eldrup@econ.ku.dk.

‡University of Copenhagen, CEBI, CESifo, RFBerlin and CEPR. Email: ctk@econ.ku.dk.

1 Introduction

This paper integrates taxpayer non-compliance and government tax enforcement into the theory of optimal non-linear income taxation (Mirrlees 1971; Diamond 1998; Saez 2001). History offers numerous examples of how weak enforcement undermines the revenue-generating and redistributive potential of taxes (Keen and Slemrod 2021),[1] and in developing countries, limited tax capacity continues to constrain revenue generation (Gordon and Li, 2009; Keen et al. 2015). Advancements in tax enforcement—particularly the introduction of third-party information reporting and tax withholding—are key reasons why OECD countries today can sustain high government revenues with elevated marginal tax rates (Kleven et al. 2016; Bagchi and Dusek 2021; Jensen 2022).[2] However, even in these countries, non-compliance remains substantial on income sources where third-party verification is absent, including among the rich, thereby limiting the redistributive capacity of the tax system (Kleven et al. 2011; Slemrod 2019; Alstadsæter et al. 2019; Boas et al. 2024).

The canonical model of redistribution—the Mirleesian optimal income tax framework—highlights the trade-off between equity and efficiency losses arising from labor supply distortions, while assuming that all individuals truthfully report their income to the tax authorities. Introducing non-compliance and enforcement makes the framework more complex, as it requires solving for jointly optimal non-linear tax and enforcement schedules.

In our baseline model, a social planner seeks to redistribute income among individuals with heterogenous skills. Achieving redistribution requires positive marginal taxes, which in turn create incentives to underreport income and evade taxes. To sustain the intended redistribution, the planner must al-

[1] A notable example is the introduction of the window tax in 17th-century UK, intended to shift the tax burden onto wealthier individuals using a verifiable tax base. Instead, it led to unintended consequences, such as bricked-up windows and new houses designed with few windows (Oates and Schwab 2015).

[2] For example, the marginal tax rate of an average production worker without children exceeds 50% in twelve OECD countries (computed as income tax plus employee and employer contributions less cash benefits, but excluding consumption taxes, for a single without children as a share of labor costs in 2023; see Table 3.6 in OECD 2024).

locate resources to enforcement. The framework is sufficiently general to capture tax avoidance behavior (Slemrod and Yitzhaki 2002) and non-pecuniary moral costs of noncompliance (Luttmer and Singhal 2014) and to encompass government efforts to strengthen tax morale and limit avoidance. We further extend the framework to incorporate endogenous labor supply responses; to distinguish between mechanical (tax repayments and fines from audits) and behavioral (deterrence-induced compliance) revenue effects of enforcement (Chetty 2009); and to allow for two-dimensional heterogeneity in earnings capacity and compliance behavior, together with differentiated social welfare weights on marginal consumption of evaders and non-evaders.

Our first contribution is to derive *sufficient-statistics formulas* that broadly characterize the jointly optimal redistribution and enforcement schemes. The formula for the optimal marginal tax rate at a given income level aligns with existing results: it depends on income concentration (the local Pareto parameter), social welfare weights, and the elasticity of taxable income (ETI). In our framework, however, these sufficient statistics are endogenous to enforcement. In particular, the ETI captures the responsiveness of *reported* income, which depends on the enforcement intensity.

The optimal enforcement formula equates, at each income level, the social marginal benefits (SMB) and costs (SMC) of enforcement. SMB consist of the additional tax revenue from higher reported income as well as tax repayments and fines from audits. SMC consist of increased public enforcement expenditures and welfare-weighted private compliance costs of taxpayers. The key behavioral statistic in this formula is the additional reported income generated by a marginal increase in enforcement spending.

Crucially, the optimal tax and enforcement schemes are interdependent: the elasticity of taxable income depends on enforcement, and the marginal benefit of enforcement depends on the tax rate.

Our second contribution is insights on optimal enforcement policy and the interaction between redistribution and enforcement:

Feasibility of redistribution and the role of enforcement: Implementation of a redistributive tax schedule requires that taxpayers' marginal cost (pecuniary

and non-pecuniary) of misreporting rises with *true* income. In the absence of sufficiently strong moral costs to sustain voluntary compliance, this requires enforcement, since tax authorities observe only reported income. Instruments such as audits and third-party information can reveal true incomes which, in addition, can incentivize self-revelation of income by taxpayers. However, if true incomes above a certain level cannot be detected, or detection is too costly, then redistribution becomes infeasible to implement beyond that income level.

No enforcement at the bottom: It may not be socially optimal to enforce taxes at low skill levels. In general, one dollar spent on enforcement must raise reported income by at least one dollar to be socially desirable. Because hidden income is small at the bottom of the skill distribution, the return to enforcement is low, implying that enforcement may not be worthwhile. This is further reinforced by the high welfare weights placed on low-skilled individuals, meaning that their compliance costs generated by enforcement carry substantial weight in social welfare.

Redistribution preferences and the need for enforcement: The need for enforcement is intrinsically linked to redistributive objectives. A lump-sum tax requires no income verification, while redistributive taxation does. Stronger preferences for redistribution raise optimal marginal tax rates, which in turn increase the social benefits from enforcement by magnifying the fiscal consequences of evasion. High marginal taxes therefore require strong enforcement to mitigate distortionary evasion responses.

This increase in enforcement through higher social benefits is complemented by lower social costs at the top of the distribution, reflecting the smaller welfare weights placed on the compliance burden of high-income taxpayers. Hence, both forces imply that enforcement effort becomes more valuable to the planner at higher income levels. Beyond these two direct effects, indirect effects may create a multiplier mechanism where stronger enforcement reduces the reporting elasticity, raising the tax rate and thereby increasing the returns to further enforcement.

Conditions for using a simple cost-benefit criterion for enforcement: If soci-

ety assigns zero welfare weight to the marginal consumption of individuals who evade then optimal enforcement equates marginal tax revenue of enforcement to the marginal enforcement costs. This applies under strong inequality aversion (Rawlsian preferences) but also applies under milder preferences for (vertical) redistribution if the planner places high value on horizontal inequity caused by differences in evasion behavior of same-skill individuals. Finally, it may apply only at the top of the distribution where social welfare weights can approach zero.

Our third contribution is an *application to U.S. tax and enforcement policy* and a set of simulation results illustrating the importance of enforcement and enforcement technology for redistribution policy and tax development. The model is calibrated to the U.S. economy using information on the income distribution, degree of third-party reporting across the income distribution, and recent evidence on the costs and benefits of tax audits from Boning et al. (2024). Our preliminary results reveal strong interactions between optimal tax rates and optimal enforcement intensity, and they suggest that current U.S. tax rates and audit probabilities are not jointly optimal. A decline in audit rates similar to the IRS funding cuts from 2012–2019 reduces the optimal top marginal tax rate by about 3 percentage points. We also simulate a counterfactual economy without modern enforcement technologies—specifically, without extensive third-party reporting—and show that this dramatically reduced redistribution while increasing audit expenditures.

The rest of the paper is organized as follows. Section 2 provides additional description on related literature. Section 3 presents our baseline model and main theoretical results. Section 4 presents the different extensions of the baseline model. Section 5 presents our simulation results. Section 6 offers some concluding remarks.

2 Relationship to Literature

Our main contribution is to characterize the jointly optimal nonlinear tax and enforcement schedules across the income distribution. Optimal income

taxation and optimal tax enforcement are typically studied separately. An exception is Keen and Slemrod (2017), who analyze joint tax and enforcement choices in a representative-agent framework where a benevolent government chooses a linear tax rate and an enforcement level. Our modeling of noncompliance behavior and the effects of enforcement is similar in spirit but allows for heterogeneity across individuals and redistributive taxation.

We employ the standard assumption in optimal tax theory that individuals' abilities are unobservable to the government, but assume that the government can vary enforcement efforts across the ability distribution. For example, verifying eligibility for transfer programs or earned income tax credits, or cracking down on black-market work, primarily targets the lower end of the ability distribution. By contrast, monitoring capital income, foreign income, or having high audit propensities for taxpayers with high third-party reported income levels tend to focus enforcement on those in the top of the ability distribution. A recent example is the increased enforcement spending worldwide to limit offshore tax evasion among the rich (Boas et al. 2024).

The literature on optimal enforcement has studied how audit probabilities should vary with reported income (Reinganum and Wilde 1985; Sánchez and Sobel 1993; Chander and Wilde 1998). A central finding is that the optimal audit intensity decreases with reported income, generating a regressive bias in effective tax rates relative to statutory rates. Scotchmer (1987) notes that such declining audit propensities do not align well with real-world practice, but can be reconciled by recognizing that tax authorities use observable characteristics including third-party information to sort taxpayers into audit classes. Audit intensity may decline with income within an audit class but be dominated in the aggregate by substantial variation across audit classes. Our theory focuses precisely on such across-class variation by assuming that enforcement can be targeted at specific ability levels, and thereby making concealment more costly for those taxpayers.

Our optimality condition for enforcement is not surprising in light of previous work in simpler settings (Slemrod and Yitzhaki 1987; Kleven and Kreiner 2006), but it has not been derived before within the Mirrlesian income tax

setting, thereby demonstrating the assumptions needed to derive it in this setting, and making it possible to study the interaction between optimal redistribution and optimal enforcement.

Previous theoretical work surveyed in Slemrod and Yitzhaki (2002) has examined the interaction between taxation, evasion, and labor supply. Our model with endogenous labor supply accommodates various interaction effects among labor supply, evasion, and enforcement. These interactions influence the elasticity of reported income with respect to the net-of-tax rate and the reporting response to enforcement. However, conditional on these sufficient statistics, the optimal tax and enforcement levels are unaffected.

3 The Baseline Model

Our starting point is an optimal non-linear income tax model similar to Diamond (1998), but with distortionary noncompliance responses to taxes instead of labor supply responses. In our setting, the benevolent government decides non only taxpayments, but also the levels of tax enforcement across the distribution.

3.1 Individuals

We consider a population of individuals with different earnings capacities where w denotes the exogenous level of (true) income earned by an individual. Individuals are distributed on the interval $(\underline{w}, \overline{w})$ according to the density function $f(w)$ and the cumulative distribution function $F(w)$.

The income level reported to the tax authority y may be lower than w implying a tax payment of $T(y)$ and an unreported income of $e \equiv w - y \in (0, 1)$. The individuals maximize the utility function

$$u = w - T(y) - c(y, w, a(w)), \tag{1}$$

where the last term captures private pecuniary and non-pecuniary costs of

income underreporting, which depend on the levels of reported income and true income as well as the tax enforcement effort of the government denoted by $a(w)$. Because our focus is on tax enforcement, we assume for simplicity that preferences are quasi-linear between disposable income $w - T(y)$ and concealment costs $c(\cdot)$,[3] but consider a quite general specification for the costs. We assume that the marginal cost of evasion (reduction in y), which equals $-c_y(\cdot)$, is positive and strictly decreasing in income reported to the tax authority y, i.e. $-c_{yy}(\cdot) < 0$, while it is (weakly) increasing in the enforcement efforts of the government, $-c_{ya}(\cdot) \geq 0$, and in true income $-c_{yw}(\cdot) \geq 0$. Below we discuss the modelling of evasion costs in greater detail.

From eq. (1), we obtain the first order condition

$$t(w) \equiv T'((y(w))) = -c_y(y(w), w, a(w)), \tag{2}$$

which determines the optimal choice $y(w)$ of a type w individual conditional on the schedules for taxes and enforcement.

By differentiating this optimality condition (2), we can derive the income reporting elastiticity with respect to the net-of-tax rate (the ETI in this context).[4] This gives (see Appendix A)

$$\varepsilon(w) \equiv \frac{\partial y(w)/y(w)}{\partial(1-t(w))/(1-t(w))} = \frac{1+c_y(\cdot)}{c_{yy}(\cdot)+T''(\cdot)}\frac{1}{y(w)} \geq 0, \tag{3}$$

where the sign follows from the second order condition, $c_{yy}(\cdot) + T''(\cdot) > 0$, and $1 + c_y(\cdot) = 1 - t(w) \leq 1$ in an interior optimum. The ETI is a standard sufficient statistics measure capturing all behavioral responses of a tax change needed to compute the change in the deadweight loss of the tax (Feldstein

[3]The quasi-linearity between disposable income $w - T(y)$ and concealment costs $c(y, w, a(w))$ disregards income effects on behavior, which is commonly done in the literature on optimal income taxation (e.g., Diamond 1998; Saez 2002; Kleven, Kreiner, and Saez 2009).

[4]In the derivations of the optimality formulas, we follow the generalized approach of Jacquet et al. (2013), which accounts for the non-linearity of the tax system when deriving the behavioral responses. Compared to other studies such as Diamond (1998) and Saez (2001) this implies that the term $T''(\cdot)$ enters in the denominators of the eqs (3) and (5).

1999). However, it is not a structural parameter and can be influenced by tax design and administration as demonstrated by Slemrod and Kopczuk (2002). In our case, the size of elasticity depends on the tax enforcement effort of the government $a(w)$.

3.2 Taxpayer Costs of Noncompliance and Enforcement Technology

Our modelling of income underreporting and enforcement, based on the conceilment cost function in utility (1), is similar to previous studies, e.g. Slemrod (2001) and Keen and Slemrod (2017).[5] However, in our setting, with heterogeneity in true income w, it is important how the costs of noncompliance depends on w, which is unobservable to the government in contrast to reported income y. If the costs only depend on reported income y then it follows from the optimality condition (2) that, for a given tax rate t, there is only one solution y^* for all individuals with $w \geq y^*$. Hence, they would all report the same income and pay the same in taxes and all income above y^* would be evaded. This implies that redistributive taxation above $w \geq y^*$ would be infeasible to implement because of violation of the individuals' incentive compability constraints. Indeed, to ensure that the Spence-Mirrlees single-crossing condition is satisfied, a neccesary condition is a monotonic relationship between w and y:

$$\frac{dy(w)}{dw} = \frac{-c_{yw}(\cdot)}{c_{yy}(\cdot) + T''(\cdot)} + \sigma(w)\frac{da}{dw} > 0, \tag{4}$$

where $\sigma(w)$ measures the marginal effect of enforcement on income reported by a type w individual:[6]

[5] The modelling is different from the classical Allingham and Sandmo (1972) model with probabilistic audits. It is well known that the Allingham-Sandmo model can be converted to a certainty-equivalent model (Mayshar 1991, Slemrod 2001), which is the approach we take here. However, in our case, with non-linear taxation, probalistic realization of income is intractable. We return to this topic in Section 4.2.

[6] We use the notation $\partial y(w)/\partial a(w)$ to denote the (partial) causal impact of an increase in $a(w)$ on $y(w)$. This may be different from $dy(w)/da(w)$, which we define as the change along the optimal path as w change. It may be that $\partial y(w)/\partial a(w) \neq dy(w)/da(w)$

$$\sigma(w) \equiv \frac{\partial y(w)}{\partial a(w)} = \frac{-c_{ya}(\cdot)}{c_{yy}(\cdot) + T''(\cdot)} \geq 0, \tag{5}$$

which is positive due to the second order condition and the assumption $-c_{ya}(\cdot) \geq 0$.

Consider now the case without any tax enforcement, $a = 0$. In this case, it follows from the neccesary condition (4) that implementing positive marginal taxes requires morale costs to be increasing in true income, $-c_{yw}(\cdot) > 0$, to ensure voluntary tax payments increasing in true income. Without this property, implying that the first term in condition (4) is zero, sustaining increasing tax payments requires tax enforcement. Moreover, in this case, it is important that tax enforcement efforts $a(w)$ can raise private evasion costs at higher levels of true income w to ensure increasing income reporting.

Note that the assumption that tax enforcement $a(w)$ can vary with true income w does not imply that the government observes w of an individual, but rather that the enforcement technology enables the government to target ressources on detecting evasion of people at certain places in the true income distribution and making it more costly for these individuals to hide their income.

3.3 Government

We consider a benevolent government that chooses taxes and tax enforcement $(T(y(w)), a(w))_{\underline{w}}^{\overline{w}}$ so as to maximize social welfare

$$\Omega = \int_{\underline{w}}^{\overline{w}} \Psi(V(x)) f(x) dx, \tag{6}$$

where $V(x)$ is the indirect utility function

$$V(w) = w - T(y(w)) - c(y(w), w, a(w)), \tag{7}$$

because y and a can depend directly on w.

while $\Psi(\cdot)$ is a concave function capturing the redistributive preferences of the government. The policy of the government needs to obey the government budget constraint

$$\int_{\underline{w}}^{\overline{w}} \left[T\left(y\left(w\right)\right) - a\left(w\right)\right] f\left(w\right) dw \geq R, \tag{8}$$

where R is a fixed revenue requirement. The optimal policy also needs to take into account the behavioral responses of the individuals (2) and ensure that these choices are incentive compatible. A neccesary condition for this is that the indirect utility function is increasing in w, i.e.,

$$V'\left(w\right) = 1 - c_a\left(y\left(w\right), w, a\left(w\right)\right) a'\left(w\right) - c_w\left(y\left(w\right), w, a\left(w\right)\right) > 0, \tag{9}$$

where we have used the first order condition (2).

We use standard definitions of social welfare weights to capture the redistributive preferences of the benevolent government:

$$g\left(w\right) \equiv \Psi'\left(V\left(x\right)\right)/\lambda, \tag{10}$$

$$G\left(w\right) \equiv \frac{\int_w^{\overline{w}} g\left(w\right) f\left(x\right) dx}{1 - F\left(w\right)}, \tag{11}$$

where the welfare weight $g\left(w\right)$ measures the social marginal utility of consumption of a type w individual relative to the the marginal cost of public funds λ, while $G\left(w\right)$ is the average welfare weight of individuals with $x \geq w$. The assumption of quasi-linear preferences implies $\lambda = \int_{\underline{w}}^{\overline{w}} g\left(w\right) f\left(x\right) dx$ from which is follows that the sum of welfare weights of all individuals sum to one.

3.4 Optimal Redistribution and Enforcement

To derive the optimal policy, we first reformulate the optimization problem of the government to an optimal control problem. Using the indirect utility

function (7), we may rewrite the government budget constraint (8) to

$$\int_{\underline{w}}^{\overline{w}} \left[w - c\left(y\left(w\right), w, a\left(w\right)\right) - V\left(w\right) - a\left(w\right)\right] f\left(w\right) dw = R. \tag{12}$$

We let $a\left(w\right)$ be a state variable with changes according to a control function $r\left(w\right)$:

$$a'\left(w\right) = r\left(w\right). \tag{13}$$

With these expressions, we can reformulate the optimization problem of the government to maximize social welfare (6) with respect to the control functions $\{y\left(w\right), r\left(w\right)\}_{\underline{w}}^{\overline{w}}$ subject to the equations determining the changes in the state variables (9) and (13) and subject to the government budget constraint (12) and without any endpoint conditions, i.e., $V\left(\underline{w}\right)$, $V\left(\overline{w}\right)$, $a\left(\underline{w}\right)$ and $a\left(\overline{w}\right)$ are all free. Using Pontryagin's maximum principle, we can then show:

Proposition 1 *The socially optimal redistribution and enforcement policy at interior solutions are characterized by:*

$$t\left(w\right) = \frac{1 - G\left(w\right)}{1 - G\left(w\right) + \alpha\left(w\right)\varepsilon\left(w\right)}, \tag{14}$$

$$t\left(w\right)\sigma\left(w\right) = 1 + g\left(w\right)c_a\left(\cdot\right), \tag{15}$$

where $\alpha\left(w\right) \equiv \frac{h(y(w))y(w)}{1-H(y(w))}$ *is the (local) Pareto parameter of the distribution of reported income.*

Proof. See Appendix B. ■

This proposition provides sufficient statistics formulas for the social optimum. The formula for the tax rate (14) is similar to previous results in the literature and relates the optimal tax rate at the optimum to the redistributive preferences of the social planner captured by the welfare weights $G\left(\cdot\right)$, income concentration captured by the Pareto parameter $\alpha\left(\cdot\right)$, and behavioral responses to tax rate changes captured by the ETI. In our case, the ETI equals the income reporting elasticity $\varepsilon\left(\cdot\right)$ in eq. (3).

The optimal enforcement formula (15) equates, at each income level, the social marginal benefits (SMB) and costs (SMC) of enforcement. SMB consist of additional revenue from higher reported income, while SMC include increased public enforcement expenditures, normalized to one, and welfare-weighted private compliance costs of taxpayers. The key behavioral statistic in this formula is the additional reported income generated by a marginal increase in enforcement spending, $\sigma(w)$.

The optimal tax and enforcement schemes are interdependent, In particular, the elasticity of taxable income depends on enforcement, and the marginal benefit of enforcement depends on the tax rate.

With standard social preferences, the welfare weights $g(w)$ are decreasing at higher incomes implying that private compliance costs matter less as we move up in the distribution. This implies that we move closer to a simple cost-benefit criterion for enforcement when we move up in the distribution. In the extreme case, with Rawlsian preferences a simple cost-benefit criterion applies throughout the distribution.

From the formulas, it follows that stronger preferences for redistribution increase optimal enforcement through two channels. Consider for example an increase in welfare weights, $g(w)$, below a given income level $y(w)$ and a corresponding decrease in weights above this level. The decrease in $g(w)$ above $y(w)$ decreases also $G(w)$ implying, ceteris paribus, an increase in tax rates at incomes above $y(w)$. This raises the marginal social benefit of enforcement, the LHS of formula (15). At the same time, the social planner assigns less weight to the private compliance costs of individuals with income above $y(w)$. In addition to these two direct effects, there are indirect effects that may generate a multiplier mechanism. A higher enforcement intensity can lower the reporting elasticity, which then increases the optimal tax rate, which further raises the benefits of enforcement, and so forth.

At the very top of the distribution, we get the following limit result

Corollary 1 *Assume that that the factors in (14) and (15) converge asymptotically to finite values* $(\hat{\alpha}, \hat{g}, \hat{\sigma})$ *when* $w \to \infty$. *Asymptotically, we then*

have

$$\begin{aligned} \lim_{w\to\infty} t(w) &= \hat{t} = \frac{1-\hat{g}}{1-\hat{g}+\hat{\alpha}\hat{\varepsilon}} \\ \hat{t}\cdot\hat{\sigma} &= 1+\hat{g}\hat{c}_a \end{aligned}$$

Proof. This follows directly from Proposition 1. ■

In the limit $G(w) \to g(w)$, implying that the social welfare weights enter in the same manner in the optimal tax and enforcement formulas.

4 Analytical Extensions

We extend the baseline framework by including endogenous labor supply, mechanical revenue effects of enforcement reflecting tax repayments and fines from tax audits, and heterogeneity in evasion behavior of taxpayers. Finally, we also study the possibility of a corner solution with no enforcement at low incomes.

4.1 Endogenous Labor Supply

We consider the utility function

$$u = z - T(y) - c(y, z, w, a(w)), \tag{16}$$

where w is the income potential of an individual, z is actual/true income, y is income reported to the tax authority, and $e = z - y$ is evasion. The c-function captures both disutility of working, $c_z(\cdot) > 0$, and costs of noncompliance as before. The endogeneity of labor supply implies a distortionary tax wedge between actual and potential income in addition to the distortionary wedge between reported and actual income. The general formulation of the cost function implies that we allow for evasion and labor supply being substitutes or complements and for enforcement to have direct positive or negative impact on labor supply. Importantly, although the optimization problem

of individuals now has two dimensions, the social planner still faces a one-dimensional screening problem because taxpayers can only deviate from the intended tax burden of the social planner by reducing their reported income (through labor supply or evasion).[7]

Maximizing utility with respect to (z, y) gives the first order conditions

$$t(w) = -c_y(y, z, w, a(w)), \tag{17}$$

$$1 = c_z(y, z, w, a(w)). \tag{18}$$

The elasticity of taxable income and the effect of enforcement on reported income becomes

$$\varepsilon(w) = \frac{1 + c_y(\cdot)}{c_{yy}(\cdot) + T''(\cdot) - (c_{yz}(\cdot))^2 / c_{zz}(\cdot)} \frac{1}{y(w)}, \tag{19}$$

$$\sigma(w) = \frac{-c_{ya}(\cdot) + c_{za}(\cdot) c_{yz}(\cdot) / c_{zz}(\cdot)}{c_{yy}(\cdot) + T''(\cdot) - (c_{yz}(\cdot))^2 / c_{zz}(\cdot)}, \tag{20}$$

where the denominators are positive assuming that the second-order conditions for an individual optimum are fulfilled (See Appendix C.4 for derivations). If labor supply is perfectly inelastic, $c_{zz}(\cdot) \to \infty$, then these equations are identical to eqs (3) and (5) and the model is identical to the baseline model in Section 3. Another special case is $c_{zz} = -c_{yz}$ and $c_{ya} = c_{za} = 0$, in which case a change in taxation gives the same response in reported and actual income and enforcement has no effects. This corresponds to the standard Mirrlesian optimal income tax framework.

With endogenous labor supply, we have

Proposition 2 *The socially optimal redistribution and enforcement policy at interior solutions are characterized by eqs (14) and (15) with behavioral statistics $\varepsilon(w)$ and $\sigma(w)$ given by eqs (19) and (20).*

Proof. See Appendix C. ■

[7] This implies that the incentive compatibility constraints still are one dimensional. Therefore, it does not turn into a more complicated multidimensional screening problem (Kleven et al. 2009).

This result shows that the sufficient statistics formulas in Proposition 1 remain valid when incorporating labor supply responses and allowing for various interaction effects between labor supply, noncompliance, and enforcement. These interaction effects influence the size of the ETI and the reporting response to enforcement, but given these sufficient statistics, the optimal tax and enforcement levels remain unchanged.

4.2 Direct Revenue Effects of Enforcement

In the model, enforcement raises government revenue by deterring evasion behavior. In addition to this behavioral effect, enforcement instruments such as audits generate mechanical revenue effects through recovered taxes and fines (Allingham and Sandmo 1972). In our non-linear tax model, probalistic audits and taxable income is intractable. However, our evasion costs function of the taxpayer $c\left(y, w, a\left(w\right)\right)$ can include $p\left(y, w, a\left(w\right)\right)\left(y - w\right)\left(1 + \theta\right)$ where $p\left(y, w, a\left(w\right)\right)$ is the share of evaded income to be repaid—instead of an audit propability—while θ is a fine in proportion to evaded income. Without loss of generality, we may model this as a share $\gamma\left(w\right)$ of private evasion costs $c\left(y, w, a\left(w\right)\right)$ that reflect tax repayments and fines. This does not change the optimum of the taxpayers who do not care about whether some of the private costs reflect repayment and fines received by the government. However, the government budget constraint (8) needs to include the direct revenue from audits:

$$\int_{\underline{w}}^{\overline{w}} \left[T\left(y\left(w\right)\right) + \gamma\left(w\right) c\left(y, w, a\left(w\right)\right) - a\left(w\right)\right] f\left(w\right) dw \geq R.$$

where $\gamma\left(w\right)$ is the share of private costs that reflects disclosed tax evasion and fines.

In this case, Proposition 1 is replaced by

Proposition 3 *The socially optimal redistribution and enforcement policy*

at interior solutions are characterized by:

$$t(w) = \frac{1 - G(w)}{1 - G(w) + \alpha(w)(1 - \gamma(w))\varepsilon(w)} \tag{21}$$

$$\gamma(w) c_a(\cdot) + (1 - \gamma(w)) t(w) \sigma(w) = 1 + g(w) c_a(\cdot) \tag{22}$$

Proof. See Appendix D. ■

The tax formula shows that accounting for the mechanical enforcement effect adjusts the ETI term downwards because a proportion $\gamma(w)$ of the private cost of non-compliance represents transfers to the government rather than deadweight losses, as in Chetty (2009). Consequently, only $1-\gamma(w)$ of the tax wedge generates efficiency costs, which reduces the welfare loss from taxation. This raises optimal tax rates relative to a setting where this component of enforcement revenue is excluded from the government budget.

The social benefit of enforcement remains equal to its effect on tax revenue, but now includes the mechanical revenue effect along with the deterrence-induced revenue effect. The deterrence effect is scaled by the factor $1-\gamma(w)$, similar to the ETI term in the tax formula, reflecting that only $1 - \gamma(w)$ of this revenue effect corresponds to reduced efficiency costs.

4.3 Heterogeneity in Evasion Behavior

The baseline model focuses on heterogeneity in skills (w), which gives rise to vertical inequality and a motive for redistribution. We now add heterogeneity in evasion behavior (θ) due to differences in opportunity or willingness to evade.[8] This heterogeneity—in combination with the redistribution driven by skill heterogeneity (vertical inequality) creates differences in consumption among individuals with the same skill level (horizontal inequality). We further allow the social planner to assign different welfare weights to inequalities stemming from skill heterogeneity versus heterogeneity in evasion, for exam-

[8] In the World Value Survey (2017), about 60% of US respondents state that it is never justifiable to cheat on taxes (chossing 1 on a 1-10 scale), while about 15% find it at least somewhat justifiable (choosing 5-10).

ple involving a social welfare discount on the marginal utility of consumption of evaders as suggested by Cowell (1990).[9]

The utility function now equals

$$u = w - T(y) - c(y, w, \theta, a(w)), \tag{23}$$

where taxpayers with higher θ have better opportunities or larger willingness to evade, $-c_{y\theta} < 0$. Let $y(w, \theta)$ denote the optimal reported income for a (w, θ) type individual. We assume there exist $\bar{\theta}(w)$ for each w such that the corner solution with no evasion, $y(w, \theta) = w$, is optimal for $\theta < \bar{\theta}(w)$ while the first-order condition applies for $\theta \geq \bar{\theta}(w)$, i.e.,

$$\begin{array}{rl} t(w) < -c_y(w, w, \theta, a(w)) & \text{for } \theta < \bar{\theta}(w) \\ t(y(w, \theta)) = -c_y(y(w, \theta), w, \theta, a(w)) & \text{for } \theta \geq \bar{\theta}(w) \end{array} \tag{24}$$

We define $\sigma(w, \theta) \equiv \partial y(w, \theta) / \partial a(w)$ which equals zero for individuals with $\theta < \bar{\theta}(w)$ while it equals the expression (5) for individuals with $\theta \geq \bar{\theta}(w)$.

The social welfare function and government budget constraint become

$$\Omega = \int\int_{\theta \in \Theta, w \in \mathbb{R}_+^*} \Psi(V(w, \theta), w, \theta) f(w, \theta) \, dw d\theta, \tag{25}$$

$$\int\int_{\theta \in \Theta, w \in \mathbb{R}_+^*} [T(y(w, \theta)) - a(w)] f(w, \theta) \, dw d\theta \geq R, \tag{26}$$

where indirect utility $V(\cdot)$ now depends on the type space (w, θ) which has the joint density distribution $f(w, \theta)$. The welfare weight assigned to type (w, θ) individuals is defined as $g(w, \theta) = \Psi_V(\cdot)/\lambda$.

To solve for the optimality conditions of the social planner, we use the extended tax perturbation approach of Jacquet and Lehmann (2021). To do this, we define $\hat{w}(y, \theta)$ as the skill level w that gives optimum y within

[9] Studies on fairness views on inequality show that they depend on the sources of income differences (Hvidberg et al. 2023). For instance, differences due to effort are seen as fairer than those due to luck (Almås et al. 2010). We are not aware of comparative studies also considering differences due to variation in evasion behavior.

type θ individuals, which is the skill level that solves (24) for a given (y, θ). We may then define on the (y, θ) space, the density distribution $\hat{h}(y, \theta) = f(\hat{w}(y, \theta), \theta)$, welfare weights $\hat{g}(y, \theta) = g(\hat{w}(y, \theta), \theta)$ and reporting elasticities $\hat{\varepsilon}(y, \theta) = \varepsilon(\hat{w}(y, \theta), \theta)$. The average welfare weight and elasticity of individuals with reported income y then equals

$$\begin{aligned} \hat{g}(y) &= \int_{\theta \in \Theta} \hat{g}(y, \theta) \hat{h}_{\theta|y}(\theta) \, d\theta, \\ \hat{\varepsilon}(y) &= \int_{\theta \in \Theta} \hat{\varepsilon}(y, \theta) \hat{h}_{\theta|y}(\theta) \, d\theta, \end{aligned}$$

where $\hat{h}_{\theta|y}(\theta)$ denotes the density distribution of θ conditional on reported income y. The average welfare weight on individuals with income above y then becomes

$$\hat{G}(y) \equiv \frac{\int_y^\infty \hat{g}(x) h(x) \, dx}{1 - H(y)},$$

where $h(y)$ is still the density distribution of reported income, which in this setting equals $h(y) = \int_{\theta \in \Theta} \hat{h}(y, \theta) \, d\theta$. Using these definitions, we can now derive the following result:

Proposition 4 *The socially optimal redistribution and enforcement policy at interior solutions are characterized by:*

$$t(y) = \frac{1 - \hat{G}(y)}{1 - \hat{G}(y) + \alpha(y) \hat{\varepsilon}(y)}, \tag{27}$$

$$\int_{\theta \in \Theta} t(y(w, \theta)) \sigma(w, \theta) f_{\theta|w}(\theta) \, d\theta = 1 + \int_{\theta \in \Theta} g(w, \theta) c_a(\cdot) f_{\theta|w}(\theta) \, d\theta. \tag{28}$$

where $\hat{h}_{\theta|y}(\theta)$ denotes the density distribution of θ conditional on skill w.

Proof. To be added. ■

The sufficient-statistics formulas continue to apply but with averages across individuals with different evasion propensities. The optimal tax formula is the same an in (14) but with averages over θ when computing the welfare weights and the reporting elasticity, similar to Jacquet and Lehmann (2021).

The LHS of the optimal enforcement formula is the marginal increase in tax revenue from additional enforcement, accounting for the variation in initial income levels and responses to enforcement across individuals with different θ. In the same vein, the second term on the RHS is again the welfare-weighted private compliance costs of taxpayers, but now accounting for the variation in private costs and welfare weights across θ.

Importantly, the model includes, at a given reported income level, lower-skilled individuals who report income truthfully and high-skilled individuals who evade taxes. These high-skilled individuals obtain even more consumption than similar high-skilled individuals who did not evade taxes. The tax instrument cannot target this (horizontal) inequality—rather it is a consequence of the tax—but the enforcement instrument can. If the social planner assigns negligible welfare weights on the marginal consumption of evaders and marginal enforcement has only direct costs on evaders then the second term in equation (28) vanishes. In this case, the optimal enforcement levels are determined by a simple cost-benefit criterion, independently of the degree of redistribution.

Enforcement instruments may also have direct welfare effects on non-evaders. For example, audits may also require time use for individuals who do not evade taxes. Oppositely, third-party reporting and prepopulated tax returns may reduce time use and spending on tax prepares of non-evaders while making it more costly for those who do evade.

4.4 No enforcement at the bottom

Our main analysis focuses on interior solutions. Here, we consider the possibility that the social optimum involves no tax enforcement at the bottom of the skill distribution. To analyze this case, assume that, for all $a(w) > 0$, there exists a unique solution to the first-order condition (2), denoted by $y^*(t(w), w, a(w))$ whose derivative with respect to enforcement, $y_a^*(\cdot)$, is given by (4). We further assume that the effect of enforcement $y_a^*(\cdot)$ is decreasing in a and increasing in w, and that reported income converges to zero

as enforcement vanishes, i.e., $\lim_{a(w)\to 0} y^*\left(t\left(w\right), w, a\left(w\right)\right) = 0$. In the absence of enforcement, $a\left(w\right) = 0$, reported income is zero, $y^*\left(t\left(w\right), w, 0\right) = 0$.

We extend the planner's problem with the constraint $a\left(w\right) \geq 0$. This implies that $a\left(w\right) = 0$ whenever

$$ty_a^*\left(t\left(w\right), w, a\left(w\right)\right) < 1 + g\left(w\right) c_a\left(y^*\left(t\left(w\right), w, a\left(w\right)\right), w, a\left(w\right)\right) \text{ for } a\left(w\right) \to 0, \tag{29}$$

meaning that the marginal cost of enforcement exceeds its marginal benefits. From this condition, we obtain the following result:

Corollary 2 $\lim_{a\to 0} y_a^*\left(t\left(\underline{w}\right), \underline{w}, a\left(\underline{w}\right)\right) < 1$ *for all* $t\left(\underline{w}\right) \in (0,1)$ *is a sufficient condition implying that* $a\left(w\right) = 0$ *for a non-trivial interval of abilities* $w \in (\underline{w}, \tilde{w})$ *where* $\tilde{w} > \underline{w}$.

Proof. It follows from condition (29) that $a\left(\underline{w}\right) = 0$ if $\lim_{a\to 0} y_a^*\left(t\left(\underline{w}\right), \underline{w}, a\left(\underline{w}\right)\right) < 1$. By continuity, $a\left(w\right) = 0$ will then also apply for a non-trivial interval above $\underline{w}$. ■

The result states that enforcement will not occur for a non-trivial group of low-skill individuals if one dollar spent on enforcement uncovers less than one dollar of reported income at the bottom of the distribution. More generally, enforcement at any ability level w is socially optimal only if it raises reported income by more than its marginal cost. However, this condition is particularly relevant at low skill levels, where true income is small and the level of hidden income is limited, implying that the enforcement return y_a^* is also low. Introducing enforcement higher up in the distribution has larger effects on reported income because more income will be hidden (implied by $y_{aw}^* > 0$). Hence, if enforcement is not worthwhile for relatively high-skill individuals, it will also not be worthwhile for lower-skill individuals, since enforcement is weakest where hidden income is smallest.

5 Application to US Audit and Tax Policy

In the general formulation of the model, we remain agnostic as to how a government might adjust its enforcement effort. This could occur on many margins, including expanding third-party reporting, technological improvements, public awareness campaigns, or increased auditing intensity. For the purposes of our analysis, considering auditing intensity as the margin for adjusting enforcement is natural, as audit intensity has already been shown to vary across the skill distribution and because we have strong recent evidence on the on the behavioral and fiscal responses to changes to audit policy. Boning et al. (2024) uncovers the return to audits for different groups in the income distribution. Using the results of this study and additional data on the extent of misreporting across the income distribution, we structurally estimate a model that illustrates the interplay between audit and tax policy in the U.S.

5.1 Model Specification

We consider a version of the model that includes both noncompliance and labor supply responses, as well as direct government revenue effects of enforcement (repayment of disclosed evasion and fines). In addition, we distinguish between income subject to third-party reporting and income that is self-reported. Specifically, we introduce two income components: a third-party-reported component z_1, for which labor supply adjustments are the primary behavioral response to taxation, and a self-reported component z_2, for which underreporting is the main behavioral margin. We therefore consider utility functions of the form:

$$u = z_1 + z_2 - T(z_1 + s) - c^z(z_1, w) - c^s(s, w, a),$$

which is optimized by individuals with respect to labor supply z_1 and self-reported income s. True income is $z_1 + z_2$ while reported income is $z_1 + s$.

Akin to Saez (2001), we assume for local linearity of the tax function $T(\cdot)$

in the derivations. Under this assumption, the two optimal policy formulas become:

$$t(w) = \frac{1 - G(w)}{1 - G(w) + \alpha(w)\left(\varepsilon_{z_1,1-t}\frac{z_1(w)}{y(w)} + (1-\theta(w))\,\varepsilon_{s,1-t}\frac{s(w)}{y(w)}\right)},$$

$$\theta(w)\,c_a^s(\cdot) + t(w)\,(1-\theta(w))\,\frac{\partial s(w)}{\partial a(w)} = 1 + g(w)\,c_a^s(\cdot).$$

These formulas correspond to (21) and (22). Incorporating both labor supply responses and the direct revenue effects of enforcement implies that the relevant elasticity in the optimal tax formula is a weighted combination of the labor supply elasticity and the self-reporting elasticity. Since evasion is impossible on the third-party-reported income component, this component does not enter the optimal enforcement condition.

For the calibration and simulations, we consider the following specific functional forms:

$$c^z(z_1, w) = \frac{\eta w \varepsilon}{\varepsilon + 1}\left(\frac{z_1}{\eta w}\right)^{\frac{\varepsilon+1}{\varepsilon}}, \quad c^s(s, w, a) = -s + \frac{z_2 \rho}{\rho + 1}\left(\frac{s}{z_2}\right)^{\frac{\rho+1}{\rho}} + \xi a,$$

where $z_2 = (1-\eta)\,w$ and $\rho \equiv \varepsilon + \gamma/a$. The elasticity of third-party reported income with respect to the net of tax rate is ε, which reflects only labor supply responses. The elasticity of self-reported income with respect to the net of tax rate equal $\varepsilon + \gamma/a$, which depends on tax enforcement. With massive spending on enforcement, it converges to the labor supply elasticity ε. When enforcement disappears, it converges to infinity reflecting that none of this income is reported to the tax authority. The term ξa makes the formulation flexible enough to separate the mechanical private cost of enforcement from its effect on self-reporting behavior.

The first order conditions of the individuals become:

$$\begin{aligned} z_1 &= w\eta\,(1-t)^{\varepsilon}, \\ s &= (1-\eta)\,w\,(1-t)^{\varepsilon+\gamma/a}, \end{aligned}$$

where η determines the relative importance of third-reporting.

5.2 Calibration and Structural Fit

... to be added.

5.3 Policy Simulations

Our preliminary simulation results suggest the following: (*i*) That the marginal social welfare weights implied by the tax system and the audit system in the U.S. are incongruent as the former are (slightly) progressive whereas the latter are regressive. Hence, the current tax and audit system is not jointly optimal. (*ii*) The decline in audit rates following IRS funding cuts from 2012–2019 reduces the optimal top marginal tax rate by about 3 percentage points. (*iii*) This reform was welfare-decreasing, as audit expenditures were already below their optimumal level prior to the funding cuts. (*iv*) The degree of third-party reporting has large implications for optimal policy, allowing for substantially more redistribution combined with lower audit expenditures.

6 Concluding Remarks

... to be added.

Appendix

A Derivation of eqs (3) and (5)

We consider a small change in the marginal tax rate ∂x over a small interval ∂w following the approach of Jacquet et al. (2013). From the first order condition (2), we obtain

$$T''(\cdot)\,\partial y + \partial x = -c_{yy}(\cdot)\,\partial y,$$

which implies

$$\frac{\partial y}{\partial x} = -\frac{1}{c_{yy}(\cdot) + T''(\cdot)}$$

By inserting this in the definition of the elasticity and using the first order condition (2), we get

$$\begin{aligned} \varepsilon(w) &\equiv \frac{\partial y(w)/y(w)}{\partial(1-t(w))/(1-t(w))} = \frac{1-t(\cdot)}{c_{yy}(\cdot)+T''(\cdot)}\frac{1}{y(w)} \\ &= \frac{1+c_y(\cdot)}{c_{yy}(\cdot)+T''(\cdot)}\frac{1}{y(w)} \geq 0 \end{aligned}$$

which is eq. (3).

From the first order condition (2), we obtain

$$T''(\cdot)\frac{\partial y}{\partial a} = -c_{yy}(\cdot)\frac{\partial y}{\partial a} - c_{ya}(\cdot),$$

which after rearranging gives eq. (5).

B Proof of Proposition 1

B.1 Optimality conditions

The Hamiltonian for the optimization problem may be written as

$$\begin{aligned} H &= \Psi(V(w)) f(w) + \mu(w)[1 - c_a(y(w), w, a(w)) r(w) \\ &\quad - c_w(y(w), w, a(w))] + \rho(w) r(w) \\ &\quad + \lambda[w - c(y(w), w, a(w)) - V(w) - a(w)] f(w), \end{aligned} \tag{30}$$

where $\mu(w)$ is the co-state variable associated with the state variable $V(w)$ while $\rho(w)$ is the co-state variable associated with the state variable $a(w)$. The optimality conditions are:

$H_{y(w)} = 0$ gives

$$-\mu(w) c_{ay}(\cdot) r(w) - \mu(w) c_{wy}(\cdot) - \lambda c_y(\cdot) f(w) = 0, \tag{31}$$

$H_{V(w)} = -\mu'(w)$ gives

$$[\Psi'(V(w)) - \lambda] f(w) = -\mu'(w), \tag{32}$$

$H_{r(w)} = 0$ gives

$$-\mu(w) c_a(y(w), w, a(w)) + \rho(w) = 0 \tag{33}$$

$H_{a(w)} = -\rho'(w)$ gives

$$-\mu(w) [c_{aa}(\cdot) r(w) + c_{wa}(\cdot)] - \lambda [1 + c_a(\cdot)] f(w) = -\rho'(w) \tag{34}$$

The transversality conditions are

$$\mu(\underline{w}) = 0, \mu(\overline{w}) = 0, \rho(\underline{w}) = 0, \rho(\overline{w}) = 0. \tag{35}$$

B.2 Derivation of the optimal enforcement levels, eq. (15)

By differentiating the optimality condition (33) with respect to w, we obtain

$$\rho'(w) = \mu'(w) c_a(\cdot) + \mu(w) c_{ay}(\cdot) y'(w) + \mu(w) c_{aw}(\cdot) + \mu(w) c_{aa}(\cdot) r(a)$$

where we have used $a'(w) = r(a)$. After inserting this expression for $\rho'(w)$ into (34), we get

$$\lambda [1 + c_a(\cdot)] f(w) = \mu'(w) c_a(\cdot) + \mu(w) c_{ay}(\cdot) y'(w).$$

By inserting $\mu'(w)$ from condition (32), we obtain

$$\mu(w) c_{ay}(\cdot) y'(w) = [\lambda + \Psi'(\cdot) c_a(\cdot)] f(w). \tag{36}$$

From (31) we can isolate $\mu(w)$:

$$\mu(w) = -\lambda \frac{c_y(\cdot)}{c_{ay}(\cdot) r(w) + c_{wy}(\cdot)} f(w).$$

If we insert this into eq. (36), we get

$$-\frac{c_y(\cdot)\,c_{ay}(\cdot)}{c_{ay}(\cdot)\,r(w)+c_{wy}(\cdot)}y'(w)=1+\frac{\Psi'(\cdot)}{\lambda}c_a(\cdot). \tag{37}$$

By differentiating the first order condition (2) with respect to w, we obtain:

$$c_{ya}(\cdot)\,r(w)=-c_{yy}(\cdot)\,y'(w)-c_{yw}(\cdot)-T''(\cdot)\,y'(w),$$

or

$$y'(w)=-\frac{c_{ya}(\cdot)\,r(w)+c_{yw}(\cdot)}{c_{yy}(\cdot)+T''(\cdot)}, \tag{38}$$

where we have used $a'(w)=r(w)$. By inserting this expression into eq. (37), we obtain

$$c_y(\cdot)\frac{c_{ay}(\cdot)}{c_{yy}(\cdot)+T''(\cdot)}=1+\frac{\Psi'(\cdot)}{\lambda}c_a(\cdot).$$

Finally, we use eqs (2), (5) and (10), to obtain eq. (15).

B.3 Derivation of the optimal tax rates, eq. (14)

By integrating eq. (32) and using the transversality conditions $\mu(\underline{w})=\mu(\overline{w})=0$ from (35), we obtain

$$\mu(w)=\int_{\underline{w}}^{w}\mu'(x)\,dx=-\int_{w}^{\overline{w}}\mu'(x)\,dx=\int_{w}^{\overline{w}}\left[\Psi'(V(x))-\lambda\right]f(x)\,dx.$$

By using the definitions of social weights in (10)-(11), we may rewrite the above condition to

$$\frac{\mu(w)}{\lambda}=(G(w)-1)(1-F(w)).$$

By inserting this into eq. (31), we obtain

$$-(G(w)-1)(1-F(w))(c_{ay}(\cdot)\,r(w)+c_{wy}(\cdot))-c_y(\cdot)\,f(w)=0. \tag{39}$$

Let $H(y(w))$ denote the distribution of taxable income while $F(w)$ is the

distribution of true income. By definition, we have $H(y(w)) = F(w)$. This implies

$$h(y(w))\frac{dy}{dw} = f(w).$$

By substituting eq. (38) for $y'(w)$, we obtain

$$f(w) = -h(y(w))\frac{c_{ya}(\cdot)r(w) + c_{yw}(\cdot)}{c_{yy}(\cdot) + T''(\cdot)}.$$

By inserting this into eq. (39), we get

$$-(G(w) - 1)\frac{1 - F(w)}{h(y(w))} + \frac{c_y(\cdot)}{c_{yy}(\cdot) + T''(\cdot)} = 0.$$

If we further use the first order condition (2), the elasticity definition (3) and the identity $F(w) = H(y(w))$, we get

$$\frac{t(w)}{1 - t(w)} = \frac{1}{\varepsilon(w)}\frac{1 - H(y(w))}{h(y(w))y(w)}(1 - G(w)),$$

which is the standard ABC formula in optimal non-linear taxation (Diamond 1998; Saez 2001), which can also be written as

$$t(y(w)) = \frac{1 - G(y(w))}{1 - G(y(w)) + \alpha(y(w))\varepsilon(y(w))}$$

where $\alpha(w) \equiv \frac{h(y(w))y(w)}{1-H(y(w))}$.

C Proof of Proposition 2

C.1 Optimization problem of the government

The government wishes to maximize $W = \int_{\underline{w}}^{\overline{w}} \Psi(V(x)) f(x)\, dx$. The indirect utility function equals

$$V(w) = z(w) - T(y(w)) - c(y(w), z(w), w, a(w)),$$

implying

$$V'(w) = -c_w(y(w), z(w), w, a(w)) - c_a(y(w), z(w), w, a(w))\, a'(w)$$

The government budget constraint is

$$\int_{\underline{w}}^{\overline{w}} [T(y(w)) - a(w)]\, f(w)\, dw \geq R,$$

which can be written as

$$\int_{\underline{w}}^{\overline{w}} [z(w) - c(y(w), z(w), w, a(w)) - V(w) - a(w)]\, f(w)\, dw \geq R.$$

The Hamiltonian for this optimization problem equals

$$\begin{aligned} H &= \Psi(V(w))\, f(w) + \mu(w)\left[-c_w(y(w), z(w), w, a(w)) - c_a(y(w), z(w), w, a(w))\, r(w)\right] \\ &\quad + \rho(w)\, r(w) + \lambda\left[z(w) - c(y(w), z(w), w, a(w)) - V(w) - a(w)\right] f(w) \end{aligned}$$

The optimality conditions are:

$H_{y(w)} = 0$ gives

$$\mu(w)\left[-c_{yw}(\cdot) - c_{ya}(\cdot)\, r(w)\right] - \lambda c_y(\cdot)\, f(w) = 0 \tag{40}$$

$H_{z(w)} = 0$ gives

$$c_{zw}(\cdot) + c_{za}(\cdot)\, r(w) = 0, \tag{41}$$

where we have used condtiion (18).

$H_{V(w)} = -\mu'(w)$ gives

$$[\Psi'(V(w)) - \lambda]\, f(w) = -\mu'(w), \tag{42}$$

$H_{r(w)} = 0$ gives

$$-\mu(w)\, c_a(\cdot) + \rho(w) = 0 \tag{43}$$

$H_{a(w)} = -\rho'(w)$ gives

$$\mu(w)\left[-c_{wa}(\cdot) - c_{aa}(\cdot) r(w)\right] - \lambda\left(1 + c_a(\cdot)\right) f(w) = -\rho'(w) \tag{44}$$

The transversality conditions are again $\mu(\underline{w}) = 0, \mu(\overline{w}) = 0, \rho(\underline{w}) = 0, \rho(\overline{w}) = 0$.

C.2 Derivation of the optimal enforcement levels

By differentiating the optimality condition (43) with respect to w, we obtain

$$\rho'(w) = \mu'(w) c_a(\cdot) + \mu(w)\left[c_{ay}(\cdot) y'(w) + c_{az}(\cdot) z'(w) + c_{aa}(\cdot) r(w) + c_{aw}(\cdot)\right]$$

where we have used $a'(w) = r(w)$. After inserting this expression for $\rho'(w)$ into (44), we get

$$\lambda\left(1 + c_a(\cdot)\right) f(w) = \mu'(w) c_a(\cdot) + \mu(w)\left[c_{ay}(\cdot) y'(w) + c_{az}(\cdot) z'(w)\right]$$

By inserting $\mu'(w)$ from condition (42), we obtain

$$\mu(w)\left[c_{ay}(\cdot) y'(w) + c_{az}(\cdot) z'(w)\right] = \left[\lambda + \Psi'(\cdot) c_a(\cdot)\right] f(w) \tag{45}$$

From (40) we can isolate $\mu(w)$:

$$\mu(w) = -\lambda \frac{c_y(\cdot)}{c_{yw}(\cdot) + c_{ya}(\cdot) r(w)} f(w)$$

If we insert this into eq. (45), we get

$$-c_y(\cdot) \frac{c_{ya}(\cdot) y'(w) + c_{za}(\cdot) z'(w)}{c_{yw}(\cdot) + c_{ya}(\cdot) r(w)} = 1 + \frac{\Psi'(\cdot)}{\lambda} c_a(\cdot).$$

Appendix C.5 derives $y'(w)$ and $z'(w)$ from the individual optimum conditions (17) and (18). By inserting these derivatives, eqs (50) and (51), into

the above equation, we obtain

$$c_y(\cdot)\frac{c_{zz}(\cdot)c_{ya}(\cdot)-c_{za}(\cdot)c_{yz}(\cdot)}{(c_{yy}(\cdot)+T''(\cdot))c_{zz}(\cdot)-(c_{yz}(\cdot))^2}=1+\frac{\Psi'(\cdot)}{\lambda}c_a(\cdot).$$

or

$$t(w)\frac{\partial y(w)}{\partial a(w)}=1+\frac{\Psi'(\cdot)}{\lambda}c_a(\cdot)$$

where we have used eqs (17) and (20). This corresponds to eq. (15).

C.3 Derivation of optimal tax rate

By integrating eq. (42) and using the transversality conditions $\mu(\underline{w})=\mu(\overline{w})=0$, we obtain

$$\mu(w)=\int_{\underline{w}}^{w}\mu'(x)\,dx=-\int_{w}^{\overline{w}}\mu'(x)\,dx=\int_{w}^{\overline{w}}\left[\Psi'(V(x))-\lambda\right]f(x)\,dx.$$

By using the definitions of social weights in (10)-(11), we may rewrite the above condition to

$$\frac{\mu(w)}{\lambda}=(G(w)-1)(1-F(w)).$$

By inserting this into eq. (40), we obtain

$$(G(w)-1)(1-F(w))(c_{ay}(\cdot)r(w)+c_{wy}(\cdot))+c_y(\cdot)f(w)=0. \qquad (46)$$

The distribution of reported income is denoted by $H(y(w))$ while $F(w)$ is the distribution of potential income. Since $H(y(w))=F(w)$, it follows that

$$\begin{aligned}f(w) &= h(y(w))\frac{dy(w)}{dw}\\ &= -h(y(w))\frac{(c_{yw}(\cdot)+c_{ya}(\cdot)r(w))c_{zz}(\cdot)}{(c_{yy}(\cdot)+T''(\cdot))c_{zz}(\cdot)-(c_{yz}(\cdot))^2}\end{aligned}$$

where we have used eq. (50) derived in Appendix C.5. By substituting this expression and $F(w) = H(y(w))$ into the above equation (46), we obtain

$$(G(w) - 1)\frac{1 - H(w)}{h(y(w))} - c_y(\cdot)\frac{c_{zz}(\cdot)}{(c_{yy}(\cdot) + T''(\cdot))c_{zz}(\cdot) - (c_{yz}(\cdot))^2} = 0.$$

By using the expression for the elasticity of taxable income (19), we obtain

$$(G(w) - 1)\frac{1 - H(w)}{h(y(w))y(w)} - \frac{c_y(\cdot)}{1 + c_y(\cdot)}\varepsilon(w) = 0.$$

Finally, by using the first-order condition (17) and the definition $\alpha(w) \equiv \frac{h(y(w))y(w)}{1-H(y(w))}$, we arrive at

$$t(w) = \frac{1 - G(w)}{1 - G(w) + \alpha(w)\varepsilon(w)}.$$

This corresponds to eq. (14).

C.4 Derivation of eqs (19) and (20)

The second-order conditions associated with the first-order conditions (17) and (18) are

$$-c_{zz}(\cdot) < 0, \quad -c_{yy}(\cdot) - T''(\cdot) < 0 \tag{47}$$

$$\det H = (c_{yy}(\cdot) + T''(\cdot))c_{zz}(\cdot) - (c_{yz}(\cdot))^2 > 0, \tag{48}$$

where H is the Hessian matrix

$$H = \begin{bmatrix} -c_{yy}(\cdot) - T''(\cdot) & -c_{yz}(\cdot) \\ -c_{yz}(\cdot) & -c_{zz}(\cdot) \end{bmatrix}. \tag{49}$$

By differentiating the first-order condition (17) and (18) and using Cramer's rule, we can define the matrices

$$H_{y,a} = \begin{bmatrix} c_{ya}(\cdot) & -c_{yz}(\cdot) \\ c_{za}(\cdot) & -c_{zz}(\cdot) \end{bmatrix}, \quad H_{y,t} = \begin{bmatrix} 1 & -c_{yz}(\cdot) \\ 0 & -c_{zz}(\cdot) \end{bmatrix},$$

and calculate the derivatives

$$\frac{\partial y(w)}{\partial a(w)} = \frac{\det H_{y,a}}{\det H}, \quad \frac{\partial y(w)}{\partial t(w)} = \frac{\det H_{y,t}}{\det H},$$

where $t \equiv T'(y(w))$ denotes the marginal tax rate. By computing the determinants of the matrices and using the definition of the elasticity $\varepsilon(w)$ as well as the first order condition $t(w) = -c_y(\cdot)$, we arrive at the eqs (19) and (20). The denominators are positive due to the second-order conditions (47)-(48).

C.5 Impact of changes in w

Using the first-order conditions (17) and (18) and Cramer's rule, we can calculate the derivaties with respect to w in the following way

$$\frac{dy}{dw} = \frac{\det H_{y,w}}{\det H}, \quad \frac{dz}{dw} = \frac{\det H_{z,w}}{\det H},$$

where H is defined in (49) while

$$\begin{aligned} H_{y,w} &= \begin{bmatrix} c_{yw}(\cdot) + c_{ya}(\cdot) r(w) & -c_{yz}(\cdot) \\ c_{zw}(\cdot) + c_{za}(\cdot) r(w) & -c_{zz}(\cdot) \end{bmatrix}, \\ H_{z,w} &= \begin{bmatrix} -c_{yy}(\cdot) - T''(\cdot) & c_{yw}(\cdot) + c_{ya}(\cdot) r(w) \\ -c_{yz}(\cdot) & c_{zw}(\cdot) + c_{za}(\cdot) r(w) \end{bmatrix}. \end{aligned}$$

By computing the determinants and using $c_{zw}(\cdot) + c_{za}(\cdot) r(w) = 0$ from eq. (41), we get

$$\frac{dy}{dw} = \frac{-(c_{yw}(\cdot) + c_{ya}(\cdot) r(w)) c_{zz}(\cdot)}{(c_{yy}(\cdot) + T''(\cdot)) c_{zz}(\cdot) - (c_{yz}(\cdot))^2}, \tag{50}$$

$$\frac{dz}{dw} = \frac{[c_{yw}(\cdot) + c_{ya}(\cdot) r(w)] c_{yz}(\cdot)}{(c_{yy}(\cdot) + T''(\cdot)) c_{zz}(\cdot) - (c_{yz}(\cdot))^2}. \tag{51}$$

D Proof of Proposition 3

D.1 Optimality conditions

The Hamiltonian for the optimization problem may be written as

$$\begin{aligned} H &= \Psi\left(V\left(w\right)\right) f\left(w\right) + \mu\left(w\right)\left[1 - c_a\left(y\left(w\right), w, a\left(w\right)\right) r\left(w\right)\right. \\ &\quad \left. - c_w\left(y\left(w\right), w, a\left(w\right)\right)\right] + \rho\left(w\right) r\left(w\right) \\ &\quad + \lambda\left[w - \left(1 - \gamma\left(w\right)\right) c\left(y\left(w\right), w, a\left(w\right)\right) - V\left(w\right) - a\left(w\right)\right] f\left(w\right) \end{aligned} \quad (52)$$

where $\mu\left(w\right)$ is the co-state variable associated with the state variable $V\left(w\right)$, $\rho\left(w\right)$ is the co-state variable associated with the state variable $a\left(w\right)$ (we disregard the constraint $a\left(w\right) \geq 0$ in this section). The optimality conditions are:

$H_{y(w)} = 0$ gives

$$-\mu\left(w\right) c_{ay}\left(\cdot\right) r\left(w\right) - \mu\left(w\right) c_{wy}\left(\cdot\right) - \lambda\left(1 - \gamma\left(w\right)\right) c_y\left(\cdot\right) f\left(w\right) = 0, \quad (53)$$

$H_{V(w)} = -\mu'\left(w\right)$ gives

$$\left[\Psi'\left(V\left(w\right)\right) - \lambda\right] f\left(w\right) = -\mu'\left(w\right), \quad (54)$$

$H_{r(w)} = 0$ gives

$$-\mu\left(w\right) c_a\left(y\left(w\right), w, a\left(w\right)\right) + \rho\left(w\right) = 0 \quad (55)$$

$H_{a(w)} = -\rho'\left(w\right)$ gives

$$-\mu\left(w\right)\left[c_{aa}\left(\cdot\right) r\left(w\right) + c_{wa}\left(\cdot\right)\right] - \lambda\left[1 + \left(1 - \gamma\left(w\right)\right) c_a\left(\cdot\right)\right] f\left(w\right) = -\rho'\left(w\right) \quad (56)$$

The transversality conditions are again $\mu\left(\underline{w}\right) = 0, \mu\left(\overline{w}\right) = 0, \rho\left(\underline{w}\right) = 0, \rho\left(\overline{w}\right) = 0$.

D.2 Derivation of the optimal enforcement levels

By differentiating the optimality condition (55) with respect to w, we obtain

$$\rho'(w) = \mu'(w) c_a(\cdot) + \mu(w) c_{ay}(\cdot) y'(w) + \mu(w) c_{aw}(\cdot) + \mu(w) c_{aa}(\cdot) r(a)$$

where we have used $a'(w) = r(a)$. After inserting this expression for $\rho'(w)$ into (56), we get

$$\lambda [1 + (1 - \gamma(w)) c_a(\cdot)] f(w) = \mu'(w) c_a(\cdot) + \mu(w) c_{ay}(\cdot) y'(w).$$

By inserting $\mu'(w)$ from condition (54), we obtain

$$\frac{\mu(w)}{\lambda} c_{ay}(\cdot) y'(w) = \left[1 - \gamma(w) c_a(\cdot) + \frac{\Psi'(\cdot)}{\lambda} c_a(\cdot)\right] f(w). \tag{57}$$

From (53) we can isolate $\mu(w)/\lambda$:

$$\frac{\mu(w)}{\lambda} = -(1 - \gamma(w)) \frac{c_y(\cdot)}{c_{ay}(\cdot) r(w) + c_{wy}(\cdot)} f(w).$$

If we insert this into eq. (57), we get

$$-(1 - \gamma(w)) \frac{c_y(\cdot) c_{ay}(\cdot)}{c_{ay}(\cdot) r(w) + c_{wy}(\cdot)} y'(w) = 1 - \gamma(w) c_a(\cdot) + \frac{\Psi'(\cdot)}{\lambda} c_a(\cdot). \tag{58}$$

By differentiating the first order condition (2) with respect to w, we obtain:

$$c_{ya}(\cdot) r(w) = -c_{yy}(\cdot) y'(w) - c_{yw}(\cdot) - T''(\cdot) y'(w),$$

or

$$y'(w) = -\frac{c_{ya}(\cdot) r(w) + c_{yw}(\cdot)}{c_{yy}(\cdot) + T''(\cdot)}, \tag{59}$$

where we have used $a'(w) = r(w)$. By inserting this expression into eq. (37), we obtain

$$(1 - \gamma(w)) \frac{c_y(\cdot) c_{ay}(\cdot)}{c_{yy}(\cdot) + T''(\cdot)} = 1 - \gamma(w) c_a(\cdot) + \frac{\Psi'(\cdot)}{\lambda} c_a(\cdot).$$

Finally, by using (2), (5) and (10), to obtain the result.

D.3 Derivation of the optimal tax rates, eq. (14)

By integrating eq. (54) and using the transversality conditions $\mu(\underline{w}) = \mu(\overline{w}) = 0$, we obtain

$$\mu(w) = \int_{\underline{w}}^{w} \mu'(x)\,dx = -\int_{w}^{\overline{w}} \mu'(x)\,dx = \int_{w}^{\overline{w}} \left[\Psi'(V(x)) - \lambda\right] f(x)\,dx.$$

By using the definitions of social weights in (10)-(11), we may rewrite the above condition to

$$\frac{\mu(w)}{\lambda} = (G(w) - 1)(1 - F(w)).$$

By inserting this into eq. (53), we obtain

$$-(G(w) - 1)(1 - F(w))(c_{ay}(\cdot) r(w) + c_{wy}(\cdot)) - (1 - \gamma(w)) c_y(\cdot) f(w) = 0. \tag{60}$$

Let $H(y(w))$ denote the distribution of taxable income while $F(w)$ is the distribution of true income. By definition, we have $H(y(w)) = F(w)$. This implies

$$h(y(w))\frac{dy}{dw} = f(w).$$

By substituting eq. (59) for $y'(w)$, we obtain

$$f(w) = -h(y(w))\frac{c_{ya}(\cdot) r(w) + c_{yw}(\cdot)}{c_{yy}(\cdot) + T''(\cdot)}.$$

By inserting this into eq. (60), we get

$$-(G(w) - 1)\frac{1 - F(w)}{h(y(w))} + (1 - \gamma(w))\frac{c_y(\cdot)}{c_{yy}(\cdot) + T''(\cdot)} = 0.$$

If we further use the first order condition (2), the elasticity definition (3) and

the identity $F(w) = H(y(w))$, we get

$$\frac{t(w)}{1-t(w)} = \frac{1}{(1-\gamma(w))\,\varepsilon(w)} \frac{1-H(y(w))}{h(y(w))\,y(w)} (1-G(w)),$$

which can also be written as

$$t(y(w)) = \frac{1-G(y(w))}{1-G(y(w)) + \alpha(y(w))(1-\gamma(w))\,\varepsilon(y(w))}.$$

References

Allingham, Michael G. and Agnar Sandmo (1972). Income Tax Evasion: A Theoretical Analysis. *Journal of Public Economics* 1(3–4), 323-338.

Almås, Ingvild, Alexander W. Cappelen, Erik Ø. Sørensen and Bertil Tungodden (2010). Fairness and the Development of Inequality Acceptance. *Science* 328(5982), 1176-1178.

Alstadsæter, Annette, Niels Johannesen, and Gabriel Zucman (2019). Tax Evasion and Inequality. *American Economic Review* 109 (6), 2073–2103.

Andreoni, James, Brian Erard, and Jonathan Feinstein (1998). Tax Compliance. *Journal of Economic Literature* 36(2), 818-860. X

Bagchi, Sutirtha and Libor Dušek (2021). The effects of introducing withholding and third-party reporting on tax collections: Evidence from the U.S. state personal income tax, *Journal of Public Economics* 204, 104537.

Boas, Hjalte Fejerskov, Niels Johannesen, Claus Thsutrup Kreiner, Lauge Truels Larsen and Gabriel Zucman (2024). Taxing Wealth in a Globalized World: The Compliance Effects of Automatic Information Exchange, NBER Working Paper 32714.

Boning, William C, Nathaniel Hendren, Ben Sprung-Keyser and Ellen Stuart (2024). A Welfare Analysis of Tax Audits Across the Income Distribution. *Quarterly Journal of Economics* 140(1), 63-112.

Chander, Prakash, and Louis L. Wilde (1998). A General Characterization of Optimal Income Tax Enforcement. *Review of Economic Studies*

65(1): 165–181.

Chetty, Raj (2009). Is the Taxable Income Elasticity Sufficient to Calculate Deadweight Loss? The Implications of Evasion and Avoidance. *American Economic Journal: Economic Policy* 1(2), 31–52.

Cowell, Frank Alan (1990). *Cheating the Government: The Economics of Evasion*. Cambridge: MIT Press.

Diamond, Peter. A. (1998). Optimal Income Taxation: An Example with a U-Shaped Pattern of Optimal Marginal Tax Rates. *American Economic Review* 88(1), 83–95.

Feldstein, Martin (1999). Tax Avoidance and the Deadweight Loss of the Income Tax. *Review of Economics and Statistics* 81(4), 674–680.

Hvidberg, Kristoffer B., Claus T Kreiner and Stefanie Stantcheva (2023). Social Positions and Fairness Views on Inequality, *Review of Economic Studies* 90(6), 3083–3118.

Internal Revenue Service (2019). Federal Tax Compliance Research: Tax Gap Estimates for Tax Years 2011–2013. Publication 1415 (Rev. 9-2019). Washington, DC.

Jacquet, Laurence, Etienne Lehmann, and Bruno Van der Linden (2013). Optimal redistributive taxation with both extensive and intensive responses. *Journal of Economic Theory* 148(5), 1770-1805.

Jacquet, Laurence and Etienne Lehmann (2021). Optimal Income Taxation with Composition Effects. *Journal of the European Economic Association* vol. 19(2), 1299-1341.

Jensen, Anders (2022). Employment Structure and the Rise of the Modern Tax System. *American Economic Review* 112(1), 213–234.

Keen, Michael, et al. 2015. Current Challenges in Revenue Mobilization: Improving Tax Compliance . IMF Staff Report. Washington, DC: International Monetary Fund.

Keen, Michael and Joel Slemrod (2017). Optimal tax administration. *Journal of Public Economics* 152, 133-142.

Keen, Michael and Joel Slemrod (2021). *Rebellion, Rascals, and Revenue: Tax Follies and Wisdom through the Ages.* Princeton, NJ: Princeton University Press.

Kleven. Henrik J. and Claus T. Kreiner (2006). The Marginal Cost of Public Funds: Hours of Work Versus Labor Force Participation. *Journal of Public Economics* 90(10-11), 1955-1973.

Kleven, Henrik. J., Claus T. Kreiner, and Emmanuel Saez (2009). The Optimal Income Taxation of Couples. *Econometrica* 77(2), 537–560.

Kleven, Henrik J., Martin B. Knudsen, Claus T. Kreiner, Søren Pedersen, and Emmanuel Saez (2011). Unwilling or Unable to Cheat? Evidence from a Tax Audit Experiment in Denmark. *Econometrica* 79 (3), 651–92.

Kleven, Henrik. J., Claus T. Kreiner, and Emmanuel Saez (2016). Why Can Modern Governments Tax So Much? An Agency Model of Firms as Fiscal Intermediaries. *Economica* 83, 219-246.

Luttmer, Erzo F. P., and Monica Singhal (2014). Tax Morale. *Journal of Economic Perspectives* 28 (4): 149–68.

Mayshar, Joram (1991). Taxation with Costly Administration. *Scandinavian Journal of Economics* 93(1), 75-88.

Mirrlees, James. A. (1971). An Exploration in the Theory of Optimum Income Taxation. *Review of Economic Studies* 38(2), 175–208.

Oates, Wallace E. and Robert M. Schwab (2015). The Window Tax: A Case Study in Excess Burden. *Journal of Economic Perspectives* 29(1), 163–80.

OECD (2024). Taxing Wages 2024: *Tax and Gender through the Lens of the Second Earner.* OECD Publishing: Paris.

Reinganum, Jennifer F. and Louis L. Wilde (1985). Income Tax Compliance in a Principal–Agent Framework. *Journal of Public Economics* 26(1): 1–18.

Saez, Emmanuel (2001). Using Elasticities to Derive Optimal Income Tax

Rates. *Review of Economic Studies* 68(1), 205–229.

Saez, Emmanuel (2002). Optimal Income Transfer Programs: Intensive versus Extensive Labor Supply Responses. *Quarterly Journal of Economics* 117(3), 1039–1073.

Sánchez, Isabel, and Joel Sobel (1993). Hierarchical Design and Enforcement of Income Tax Policies. *Journal of Public Economics* 50(3), 345–369.

Scotchmer, Suzanne (1987). Audit Classes and Tax Enforcement Policy. *American Economic Review* 77(2), 229-233.

Slemrod, Joel (1992). Do Taxes Matter? Lessons from the 1980's. *American Economic Review* 82(2), 250-256.

Slemrod, Joel (2001). A General Model of the Behavioral Response to Taxation. *International Tax and Public Finance* 8, 119–128.

Slemrod, Joel (2019). Tax Compliance and Enforcement. *Journal of Economic Literature* 57(4), 904-54.

Slemrod, Joel and Shlomo Yitzhaki (1987). The Optimal Size of a Tax Collection Agency. *Scandinavian Journal of Economics* 89(2), 183-192.

Slemrod, Joel and Shlomo Yitzhaki (2002). "Tax Avoidance, Evasion, and Administration." In: Auerbach, Alan J. & Feldstein, Martin (eds.), *Handbook of Public Economics*, Vol. 3, Elsevier, Amsterdam & New York, pp. 1423-1470.

Slemrod, Joel and Wojciech Kopczuk (2002). The Optimal Elasticity of Taxable Income. *Journal of Public Economics* 84(1), 91–112.

World Values Survey (WVS) (2017). United States – World Values Survey Wave 7 [Microdata file]. Survey ID: USA_2017_WVS-W7_v01_M. IHSN Catalog.