Sufficient Statistics for Nonlinear Tax Systems with Preference Heterogeneity

Antoine Ferey, Benjamin B. Lockwood, and Dmitry Taubinsky*

August 2021

Abstract

A prominent justification for taxation of capital income, savings, bequests, and certain commodities is that taxes on consumption categories preferred by those with higher earnings ability are efficient even in the presence of nonlinear earnings taxation. This paper derives new formulas for optimal nonlinear tax systems in the presence of correlated preference heterogeneity, using intuitive and empirically-estimable sufficient statistics. We study tax systems that implement the optimal direct-revelation mechanism, as well as simpler tax systems. We characterize the optimum using familiar elasticity concepts and a novel sufficient statistic for preference heterogeneity: the difference between the cross-sectional variation of consumption (or saving) with income, and the causal effect of income on consumption (or saving). The Atkinson-Stiglitz Theorem is a knife-edge case corresponding to a zero difference, and many other celebrated results in optimal tax theory are subsumed as special cases. Our formulas apply in the presence of heterogeneous rates of return on savings, heterogeneous endowments, and income shifting. Our extensions include multidimensional preference heterogeneity and motives that lead the government to encourage individuals to save more. We apply the theoretical results to study optimal taxation of capital interest income in a calibrated model of the U.S. economy.

*We are grateful to Afras Sial for excellent research assistance. This project was supported by the National Institute on Aging, P30 AG-012836-26, the Boettner Center for Pensions and Retirement Security, National Institutes of Health, and the Eunice Shriver Kennedy National Institute of Child Health and Development Population Research Infrastructure Program R24 HD-044964-18, Center for Health Initiatives and Behavioral Economics, all at the University of Pennsylvania. Ferey gratefully acknowledges the financial support of Labex ECODEC at CREST, and of the Deutsche Forschungsgemeinschaft through CRC TRR 190 at LMU Munich. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Aging, the National Institutes of Health, or the University of Pennsylvania. Ferey: LMU Munich (antoine.ferey@econ.lmu.de). Lockwood: Wharton & NBER (ben.lockwood@wharton.upenn.edu). Taubinsky: UC Berkeley & NBER (dmitry.taubinsky@berkeley.edu).

1 Introduction

Most tax systems around the world involve taxes on capital income, estates, inheritances, and on some categories of consumption. The celebrated theorem of Atkinson and Stiglitz (1976) suggests that this could be suboptimal: the theorem states that when preferences are homogeneous and weakly separable, an optimal tax system will involve an income tax only. However, as was understood in contemporaneous work by Mirrlees (1976),[1] and more recently explored by Saez (2002), Diamond and Spinnewijn (2011), Gauthier and Henriet (2018) and others, the Atkinson-Stiglitz Theorem does not generalize when tastes for certain commodities (including future consumption) are correlated with earnings ability. Similar issues arise in the presence of correlated heterogeneity in endowments, rates of return or prices, and ability to engage in income shifting. This is particularly crucial in applications to the tax treatment of savings, as recent work suggests not only covariation of time preferences with earned income (see Lockwood, 2020, for a review), but also endowments and rates of return (Fagereng et al., 2019; Bach et al., 2020).

However, the Atkinson-Stiglitz Theorem remains perhaps the sharpest result about optimal commodity taxation, as existing insights about the limitations of this theorem are mostly qualitative,[2] or rely on models with special assumptions such as two-type distributions and other functional form restrictions. To our knowledge, there is no empirically-implementable formula for optimal commodity and savings taxation that allows for correlated heterogeneity in preferences and other economic characteristics that plausibly covary with earnings ability, and does not make a variety of special assumptions.

This paper addresses this important gap. Our formulas nest the Atkinson-Stiglitz Theorem and other celebrated results as special cases, and characterize optimal taxes using empirically-estimable elasticities analogous to those in the nonlinear income tax derivation of Saez (2001). We provide a characterization of optimal tax systems that implement the optimal direct-revelation mechanism, as well as simpler tax systems such as those that involve a nonlinear earnings tax and a separable tax on capital income (either linear or nonlinear), or those that involve a nonlinear earnings tax and an earnings-dependent but otherwise linear capital income tax.

Our model generalizes frameworks such as those of Saez (2002) and Golosov et al. (2013), where consumers with heterogeneous earning abilities and tastes choose labor supply and a consumption bundle that exhausts their after-tax income. To fix terminology, consider a bundle of consumption and savings. The policymaker chooses a flexible nonlinear tax system that depends on both earnings and savings (non-dependence on consumption is without loss of generality) to maximize a (weighted) utilitarian aggregation of individuals' utilities. As in Mirrlees (1971), the policymaker does not observe earnings abilities or tastes for savings, and thus is restricted to second-best policy tools that must trade off the policymakers redistributive goals against the distortionary effects on labor

[1] See also Konishi (1995).

[2] Saez (2002) answered the qualitative question of when a "small" commodity tax can increase welfare in the presence of preference heterogeneity, but left to future work the task of deriving an expression for the optimal tax, writing "It would of course be extremely useful to obtain optimal commodity tax formulas" in such a general framework.

supply and savings. In the absence of any restrictions on the policymaker's choice of tax systems, our model matches the mechanism design framework of Golosov et al. (2013).

We establish the following key results in this framework.

First, we show that under mild regularity assumptions, the optimal mechanism of Golosov et al. (2013) can be implemented by a smooth tax system. This theorem is perhaps the most mathematically innovative result in this paper. Our techniques differ from those in prior work on decentralization because of the possibility of double deviations, where individuals' preferred deviations are bundles that are not offered in the optimal mechanism. This theorem paves the way for characterizing the optimal mechanism using observable tax elasticities.

Our second key result is a characterization of smooth tax systems that implement the optimal mechanism. Optimal savings tax rates depend on a key sufficient statistic for preference heterogeneity: the difference between the cross-sectional variation of savings s with earnings z, $s'(z)$, and the causal effect of income windfalls on savings, which we denote $s'_{inc}(z)$. The residual, $s'_{pref}(z) := s'(z) - s'_{inc}(z)$ is the sufficient statistic for (local) preference heterogeneity. We show how this statistic can be estimated from existing empirical data and from behavioral responses to policy reforms, avoiding the need to explicitly model the relationship between unobserved preferences and ability. The optimal savings tax rates take a form much like that of the optimal earnings tax rate in Saez (2001), but with earnings z replaced by $s'_{pref}(z)$, and with the elasticity of taxable income replaced by the elasticity of savings with respect to the savings tax rate.[3] This sufficient statistics formulation provides an immediate generalization of the Atkinson-Stiglitz Theorem, as it implies that the optimal savings tax rate is everywhere zero when $s'_{pref}(z) = 0$ for all earnings levels z.

We show that our characterization of optimal nonlinear smooth tax systems does not depend on the structural foundations generating s'_{pref} and the other elasticity concepts. Our sufficient statistics formula continues to hold in the presence of heterogeneous endowments, differential rates of return on investments, and ability to engage in income shifting in the sense of Slemrod (1995). Consequently, our formula also provides a synthesis of a prior literature that studies these additional economic forces in isolation, often using models that make additional restrictive assumptions such as two-type distributions.[4]

Our third set of results explores what we call simple tax systems. We document that across a large number of countries, most taxes on various savings vehicles can be classified as one of three types: (i) a separable linear (SL) savings tax (ii) a separable nonlinear (SN) savings tax, and (iii) a system with a linear earnings-dependent (LED) savings tax, which allows, e.g., for lower-income people to have their savings taxed at a lower rate, as is the case for Long-Term Capital Gains in the U.S. For such systems, we show that the optimal policy can expressed using a sufficient statistics formula similar to that of the optimal smooth unrestricted tax system. A corollary of these formulas is a set of pareto efficiency conditions that facilitate direct tests of whether existing

[3] In the special case where $s'_{pref}(z) = s(z)$, our formula coincides with the formula of Saez and Stantcheva (2018).

[4] For example, we generalize Gahvari and Micheletto (2016) and Gerritsen et al. (2020) who study heterogeneous rates of return; Boadway et al. (2000) and Cremer et al. (2003) on heterogeneous endowments; and Christiansen and Tuomala (2008) on income shifting.

tax systems are optimal without additional assumptions about social marginal welfare weights.

Although these simple systems cannot always implement the optimal mechanism, we show that the sufficient conditions for when the SN and LED systems implement the optimal mechanism are still surprisingly broad. This suggests that the simple types of tax systems found frequently across the world allow for sufficient policy flexibility to achieve the welfare gains available even under much more complicated systems.

Fifth, we provide a number of extensions of our main results. Our first extension is to show that our results about SL, SN, and LED systems generalize tractably to multidimensional heterogeneity and to a potentially suboptimal income tax. Even under these more general assumptions, the causal effect of income on savings, together with the cross-sectional profile of savings across the income distribution, are together sufficient statistics for characterizing the optimal savings tax.

Our second extension is to consider many dimensions of consumption. The key new sufficient statistics with many dimensions of consumption are the cross-price elasticities, which allow us to compute what we call the tax diversion ratio—the impact on taxes collected on goods $j \neq i$ relative to the impact on taxes collected on good i when the price of good i is increased. Our baseline results are a special case that corresponds to a tax diversion ratio of zero.

The third extension considers situations where the government wants individuals to save more. This analysis generalizes Farhi and Werning (2010), who study a model model in which the weight that parents attach to the wellbeing of future generations is too low from the government's perspective. While Farhi and Werning (2010) focus on homogeneous preferences, our methods allow us to consider arbitrary heterogeneity. Our methods also allow us to consider more general and heterogeneous forms of misalignment between the government and individuals, as might result from myopia or present focus, in the sense of Laibson (1997).

The fourth extension builds on Gerritsen et al. (2020) to study cases where there is an additional efficiency rationale to tax savings. This results when the government must collect taxes in each period to meet a revenue requirement in each period, and higher-ability individuals have particularly favorable rates of return that are higher than the ratio of the marginal value of public funds between the second and the first period.

In the final part of our paper, we use our sufficient statistics formulas to study the optimal tax treatment of savings in the U.S. We use the Distributional National Accounts micro-files of Piketty et al. (2018) to calibrate the distribution of savings across the income distribution; we use recent estimates of the medium-run marginal propensity to consume (MPC) of Fagereng et al. (2019) and others to estimate the causal effect of income on savings; we use recent evidence from Agersnap and Zidar (2020) and Jakobsen et al. (2020) to obtain a range of estimates for the savings elasticity; and we use the meta-analysis of Chetty (2012) for the earnings elasticity. These estimates imply a quantitatively important role for preference heterogeneity, which leads us to compute savings tax rates that become substantial at high incomes, with a highly progressive schedule. We find that the savings tax at the low-end of the income distribution is zero or slightly negative, because savings are approximately stable at zero across low incomes. However, savings rates increase with income,

outpacing estimates of the marginal propensity to save out of additional income. This suggests that observed savings heterogeneity cannot be fully explained by causal income effects, implying that $s'_{pref}(z) > 0$, and thus that savings taxes are an efficient means of redistribution at higher incomes.

Under a broad range of parameters, our estimates of optimal savings tax rates are notably higher than those currently in place in the U.S., as well as those obtained in prior work. For example, the literature on dynamic taxation, which assumes homogeneous preferences but derives a theoretically robust role for capital taxation via the inverse Euler equation (e.g., Golosov et al., 2003; Golosov and Tsyvinski, 2006), tends to find optimal savings "wedges" of only several percentage points (see, e.g., Golosov and Tsyvinski, 2015; Golosov et al., 2016; Farhi and Werning, 2013). This suggests that preference heterogeneity in savings tax rates may play a quantitatively much larger role in optimal savings tax policy than do the social insurance motives analyzed in the dynamic taxation work. Moreover, we also find higher optimal savings tax rates than previous work studying preference heterogeneity using a structural approach; see Golosov et al. (2013). One possible reason is that Golosov et al. (2013) study preference heterogeneity by regressing a structural estimate of time preferences on a plausibly noisy proxy of earnings ability (performance on the Armed Forces Qualification Test), which may lead attenuation bias due to a noisy right-hand-side variable. Our directly implementable sufficient statistics formulas provide a complementary approach to quantifying preferences heterogeneity, without the need for structural estimates of time preferences, or the need to rely on possibly noisy proxies for earnings ability.

Our paper contributes most directly to the literature studying optimal commodity and savings taxes in the presence of correlated preference heterogeneity. Saez (2002) derives the conditions under which the optimal linear commodity tax is strictly positive or negative, but his results are purely qualitative. Allcott et al. (2019) derive a sufficient statistics formula for the optimal linear commodity tax in the presence of nonlinear income taxation and correlated preference heterogeneity. Relative to Allcott et al. (2019), we characterize a much broader class of tax systems, including some that can implement the allocation from the optimal mechanism. A focus on more flexible mechanisms is crucial to study optimal taxes on capital, inheritances, and estates in the presence of correlated preference heterogeneity. Golosov et al. (2013) derive conditions characterizing the optimal mechanism in a model like the one we study in our baseline result, but their results are in terms of structural primitives rather than empirically-estimable sufficient statistics. In parallel work, Gerritsen et al. (2020) study necessary conditions for the optimal capital tax in the presence of heterogeneous rates of return on savings. Our paper generalizes this work by allowing other forms of heterogeneity, by expressing optimal tax formulas using estimable sufficient statistics, and by providing conditions under which the tax systems in question implement the optimal mechanism.

Finally, our work relates to the larger work on optimal taxation of capital. The dynamic taxation literature establishes a robust role for savings taxation through the inverse Euler equation, but assumes that savings preferences do not vary with earnings ability. Our work is complementary in relaxing the assumption of homogeneous savings preferences, but limiting to a more static frame-

work. Extensions of our variational methods could be fruitfully applied to more dynamic models. Our work also complements the recent work of Saez and Stantcheva (2018), who derive optimal capital tax formulas by "putting wealth in the utility function." Mathematically, their baseline results and model resemble the special case of our framework where all differences in savings rates arise from preference heterogeneity.

The rest of this paper proceeds as follows. Section 2 presents our model and assumptions. Section 3 shows that smooth tax systems can implement the optimal mechanism, and provides sufficient statistics for optimal smooth tax systems. Section 4 turns to simple tax systems, providing conditions under which they implement the optimal mechanism, and providing a sufficient statistics characterization of these systems. Section 5 presents four key generalizations of our results. Section 6 applies our theoretical results to characterize optimal savings tax rates in the United States. Section 7 concludes. All proofs are gathered in the Appendix.

2 Model and Assumptions

Agents We consider a population of heterogeneous agents where θ denotes the vector of agents' characteristics and may encapsulate several characteristics. While Section 5.1 studies the case with multidimensional heterogeneity, we begin with the common assumption that $\theta \in \Theta \subset \mathbb{R}$, where Θ is compact. We assume that θ has a continuously differentiable cumulative distribution function $F(\theta)$.

Agents choose earnings z, and a consumption bundle (c, s), and derive utility $U(c, s, z; \theta)$ from the allocation. This formulation encompasses several possible applications. One is where c is numeraire consumption and s is another commodity; e.g., s could be sugary drinks consumption as in Allcott et al. (2019). A second interpretation is that c is period-1 consumption and s is the realized savings in period 2, as in Saez (2002), Golosov et al. (2013), and many others. A third interpretation is that c is period-1 consumption by the parents, while s is the wealth bequeathed to their children and consumed in period 2, as in Farhi and Werning (2010).

We assume a linear production technology where with marginal rate of transformation p between s and c. In the savings and inheritance interpretations of the model, $p = 1/R$, where R is the rate of return in a linear savings technology between the two periods.

We impose the following assumption about the utility function.

Assumption 1. *$U(c, s, z; \theta)$ is twice continuously differentiable, increasing and weakly concave in c and s, and decreasing and strictly concave in z. The first derivatives U'_c and U'_s are bounded.*

For example, a frequently-used functional form (e.g. Saez, 2002; Golosov et al., 2013) involves additively separable utility and heterogeneity in agents' productivity w and discount factor δ:

$$U(c, s, z; \theta) = u(c) + \delta(\theta) u(s) - k(z/w(\theta)), \tag{1}$$

with $u(.)$ the utility from consumption and $k(z/w)$ the disutility to work. In this example, there

is preference heterogeneity when the discount factor $\delta(\theta)$ covaries with productivity $w(\theta)$. Our analysis encapsulates this particular case as well as any other type of preference heterogeneity.

Formally, preference heterogeneity relates to individual differences in marginal rates of substitutions between commodities. Let $\mathcal{S}$ be the marginal rate of substitution between consumption c and savings s,

$$\mathcal{S}(c,s,z;\theta) := \frac{U_s'(c,s,z;\theta)}{U_c'(c,s,z;\theta)}, \tag{2}$$

and define preference heterogeneity as follows:

Definition 1. *There is preference heterogeneity if some agents prefer different consumption-savings bundles conditional on having the same earnings level; i.e.,*

$$\exists \theta_0, \forall (c,s,z),\ S_\theta'(c,s,z;\theta_0) \neq 0 \tag{3}$$

where we use the shorthand $S_\theta'(c,s,z;\theta_0) := \frac{\partial}{\partial\theta}\mathcal{S}_\theta(c,s,z;\theta)|_{\theta=\theta_0}$. In the formulation in (1), $S_\theta'(c,s,z;\theta) > 0$ for all whenever $\delta'(\theta) > 0$.

As a last bit of notation, we similarly let $\mathcal{Z}$ denote the marginal rate of substitution between consumption c and earnings z,

$$\mathcal{Z}(c,s,z;\theta) := \frac{U_z'(c,s,z;\theta)}{U_c'(c,s,z;\theta)}. \tag{4}$$

Government An agent's type θ is private information and cannot be observed by the government; only the distribution of types, $F(\theta)$, is known. The government must design a tax and transfer system that only depends on the observable variables (c,s,z).

Without any restrictions on the form of the optimal tax system, the resulting optimal allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_\theta$ must solve the following mechanism design program:

$$\max \int_\theta \alpha(\theta)\, U(c(\theta), s(\theta), z(\theta); \theta)\, dF(\theta) \tag{5}$$

subject to the resource constraint

$$\int_\theta [z(\theta) - ps(\theta) - c(\theta)]\, dF(\theta) \geq E \tag{6}$$

and incentive compatibility constraints

$$\forall\ (\theta, \theta') \in \Theta^2,\ U(c(\theta), s(\theta), z(\theta); \theta) \geq U(c(\theta'), s(\theta'), z(\theta'); \theta). \tag{7}$$

We refer to an allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_\theta$ that maximizes (5) subject to (6) and (7) as the *optimal incentive-compatible allocation.*

3 Optimal Smooth Tax Systems

In this section, we provide two key results about *smooth tax systems*, by which we mean twice continuously differentiable tax functions $\mathcal{T}(s,z)$.[5] First, we show that the optimal incentive-compatible allocation, characterized by the program in (5)-(7) is implementable by a smooth tax system under sensible regularity conditions. Second, we leverage our first result to derive a sufficient statistics characterization of optimal smooth tax systems. We assume that all individuals can trade c for s at rate p because, e.g., they are part of competitive market that sets the relative price of s to p.

We maintain the following assumptions throughout the rest of our analysis.

Assumption 2. *In the optimal incentive-compatible allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_\theta$, c, s, and z are smooth functions of θ; s and z are strictly increasing in θ; any type θ strictly prefers its allocation $(c(\theta), s(\theta), z(\theta))$ to the allocation $(c(\theta'), s(\theta'), z(\theta'))$ of another type $\theta' \neq \theta$.*

Assumptions of smoothness and monotonicity, together with strict dominance, enable us to derive implementability results as well as to apply variational methods and obtain optimal tax formulas in terms of empirically estimable sufficient statistics. The assumption that z is strictly increasing in θ guarantees that $z(\theta)$ is a bijection, and we denote its inverse $\vartheta(z)$. Furthermore, the assumption that s is strictly increasing in θ guarantees the existence of a bijection between s and z. Most of our results hold more generally whenever such a bijection exists; for example, when $s(\theta)$ is everywhere strictly decreasing in θ.

To ensure that $z(\theta)$ is increasing in θ, we impose an assumption that $\mathcal{S}'_\theta$ cannot be too negative; i.e., that the relationship between earnings ability and preferences for s isn't too *negative.*[6] Economically, this is a sensible assumption, as existing work, reviewed in our numerical section, suggests a positive association for the applications we consider. This condition is a generalization of the Spence-Mirrlees single crossing condition that $\mathcal{Z}'_\theta(c,s,z;\theta) \geq 0$ to the case of multiple commodities.

Assumption 3. *Under the optimal incentive-compatible allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_\theta$, the following extended Spence-Mirrlees condition holds for any θ:*

$$z'(\theta)\mathcal{Z}'_\theta(c(\theta), s(\theta), z(\theta); \theta) + s'(\theta)\mathcal{S}'_\theta(c(\theta), s(\theta), z(\theta); \theta) \geq 0. \tag{8}$$

3.1 Implementability with Smooth Tax Systems

Definition 2. *We say that an allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_\theta$ is implementable with a tax system $\mathcal{T}$ if*

1. *$\mathcal{T}$ satisfies type-specific feasibility: $c(\theta) + ps(\theta) + \mathcal{T}(s(\theta), z(\theta)) = z(\theta)$ for all $\theta \in \Theta$, and*

[5] Expressing the tax more generally as $\mathcal{T}(c,s,z)$ is plainly redundant, as given such a tax function, any choice of s and z implies a consumption value given by $\mathcal{C}(s,z) := \text{argmax}\{c | c = z - s - \mathcal{T}(c,s,z)\}$; thus, one can re-express tax burden as $\tilde{\mathcal{T}}(s,z) = \mathcal{T}(\mathcal{C}(s,z), s, z)$.

[6] If s is an inferior good, then analogous assumption would be that $\mathcal{S}'_\theta$ cannot be too positive.

2. *$\mathcal{T}$ satisfies individual optimization: $(c(\theta), s(\theta), z(\theta))$ maximizes $U(c, s, z; \theta)$ for all $\theta \in \Theta$, subject to the constraint $c + ps + \mathcal{T}(s, z) \leq z$.*

Our first result shows that the optimal incentive-compatible allocation is implementable by some smooth tax system.

Theorem 1. *Suppose that assumptions 1, 2 and 3 hold. Then the optimal incentive-compatible allocation is implementable by a smooth tax system. In this smooth tax system, agents' choices are interior (first-order conditions hold), and their local optima are strict (strict second-order conditions).*

Although it is clear that the optimal incentive-compatible allocation $\{(c(\theta), s(\theta), z(\theta))\}_\theta$ can always be implemented by *some* two-dimensional tax system—e.g., by defining $\mathcal{T}(s(\theta), z(\theta)) = z(\theta) - c(\theta) - s(\theta)$ for $\theta \in \Theta$ and $\mathcal{T}(s, z) = \infty$ for $(c, s, z) \notin \{(c(\theta), s(\theta), z(\theta))\}_\theta$—such a tax system is not guaranteed to be smooth. A smooth tax system involves a broader set of incentive compatibility constraints than equation (7). A smooth tax system allows agents to independently adjust s and z locally to points not chosen by any other type, and thus the set of possible deviations is much larger.

Starting from any given allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_\theta\}_\theta$, a smooth tax system can implement the allocation only by satisfying the first-order conditions

$$\mathcal{T}'_s\left(s\left(z\left(\theta\right)\right), z\left(\theta\right)\right) = \mathcal{S}\left(c(\theta), s(\theta), z(\theta); \theta\right) - 1 \tag{9}$$

$$\mathcal{T}'_z\left(s\left(z\left(\theta\right)\right), z\left(\theta\right)\right) = \mathcal{Z}\left(c(\theta), s(\theta), z(\theta); \theta\right) + 1. \tag{10}$$

In the presence of preference heterogeneity, individuals' incentives to deviate from their assigned allocation $(c(\theta), s(\theta), z(\theta))$ are higher under a smooth tax system than under an optimal incentive-compatible mechanism. For example, suppose that higher types θ have a higher taste for savings. If they deviate downward to some other earnings level $z(\theta') < z(\theta)$, then under the optimal mechanism they will be forced to choose savings level $s(\theta')$. Under a smooth tax system, however, the deviating type θ will choose a higher savings level $s' > s(\theta')$ at earnings level $z(\theta')$, making the this *double deviation* more appealing.

Typical tax implementation results about multidimensional consumption bundles and multidimensional tax systems avoid the difficulties associated with double deviations by ruling out the type of heterogeneity that we consider here. Thus, to our knowledge, our proof of Theorem 1 involves a different and more subtle mathematical approach. The proof, contained in Appendix B.1, proceeds in three steps. The first step is to construct a sequence of tax systems $\mathcal{T}_k$ such that each element in the sequence satisfies type-specific feasibility and the first-order conditions above. The sequence is ordered such that each successive element is more more convex around the savings choice $s(\theta)$ that corresponds to an earnings choice $z(\theta)$ in the optimal mechanism.

In the second step of the proof, we show that for each type θ there exists k sufficiently large such that for this type the second-order conditions hold at the point $(c(\theta), s(\theta), z(\theta))$. In other words,

for each type there is a sufficiently large k such that $(c(\theta), s(\theta), z(\theta))$ is a *local* optimum under the tax system $\mathcal{T}_k$. This step requires auxiliary Lemmas B.1 and B.2, which characterize individuals' budget constraints and second derivatives of indirect utility functions for any tax system $\mathcal{T}$ that preserves only the first-order conditions of the optimal mechanism.

In the third step, we show that there exists a sufficiently large k such that $(c(\theta), s(\theta), z(\theta))$ is a *global* optimum for *all* types θ under $\mathcal{T}_k$. We complete this step via a proof by contradiction. Under the assumption that such a k does not exist, there exists an infinite sequence of values k and types θ_k such that type θ_k prefers to deviate from $(c(\theta_k), s(\theta_k), z(\theta_k))$ under $\mathcal{T}_k$. Because the type space is compact, the Bolzano-Weirstrass Theorem allows us to extract a convergent subsequence of types θ_j who all prefer to deviate from the allocation assigned to them under the optimal mechanism. We show that this implies a contradiction because the limit type of this sequence, $\hat{\theta}$, must then prefer to deviate from $(c(\hat{\theta}), s(\hat{\theta}), z(\hat{\theta}))$ to some other allocation $(c(\theta'), s(\theta'), z(\theta'))$ offered in the optimal mechanism.

Theorem 1 is an "existence result," and our proof of the theorem does not offer insight into the structure of an optimal tax system. However, because agents' choices are shown to satisfy first-order and second-order conditions, we can use variational methods to characterize optimal tax systems. We now proceed by deriving optimal tax formulas expressed in terms of empirically-estimable sufficient statistics that transparently highlight the key economic forces governing the optimal tax system.

3.2 Sufficient Statistics for Smooth Tax Systems

3.2.1 Definitions

To define the sufficient statistics, it is helpful to write agents' optimization problem under a tax system $\mathcal{T}(s, z)$ as

$$\max_{z} \left\{ \max_{c,s} U(c, s, z; \theta) \text{ s.t. } c \leq z - s - \mathcal{T}(s, z) \right\} \tag{11}$$

where the inner problem represents the optimal choices of consumption $c(z; \theta)$ and savings $s(z; \theta)$ for a given earnings level z, and the outer problem represents the optimal choice of earnings $z(\theta)$ taking into account endogenous consumption and savings choices.

Earnings responses to tax reforms are captured through ζ_z^c, the compensated elasticity of labor income with respect to the marginal labor income tax rate, and η_z, the income effect parameter. Formally, for each level of earnings $z(\theta)$ chosen by a type θ, we define

$$\zeta_z^c(z(\theta)) := -\frac{1 - \mathcal{T}_z'(s(\theta), z(\theta))}{z(\theta)} \frac{\partial z(\theta)}{\partial \mathcal{T}_z'(s(\theta), z(\theta))}$$
$$\eta_z(z(\theta)) := -(1 - \mathcal{T}_z'(s(\theta), z(\theta))) \frac{\partial z(\theta)}{\partial \mathcal{T}(s(\theta), z(\theta))}$$

where $\mathcal{T}(s(\theta), z(\theta))$ is the tax liability and $\mathcal{T}_z'(s(\theta), z(\theta))$ is the marginal labor income tax rate. Since the earnings choice takes into account endogenous consumption and savings choices, these

elasticity concepts take into account the full sequence of adjustments due to changes in consumption and savings choices, as well as those due to any nonlinearities in the tax system.[7]

Savings responses to tax reforms are captured through $\zeta^c_{s|z}$, the compensated elasticity of savings with respect to the marginal savings tax rate, $\eta_{s|z}$, the income effect parameter, and s'_{inc}, the causal effect on consumption of s from a marginal change in gross pre-tax income z

$$\begin{aligned}
\zeta^c_{s|z}\left(z(\theta)\right) &:= -\frac{1+\mathcal{T}'_s\left(s\left(z;\theta\right),z\right)}{s\left(z;\theta\right)}\frac{\partial s\left(z;\theta\right)}{\partial \mathcal{T}'_s\left(s\left(z;\theta\right),z\right)}\Big|_{z=z(\theta)} \\
\eta_{s|z}\left(z(\theta)\right) &:= -\left(1+\mathcal{T}'_s\left(s\left(z;\theta\right),z\right)\right)\frac{\partial s\left(z;\theta\right)}{\partial \mathcal{T}\left(s\left(z;\theta\right),z\right)}\Big|_{z=z(\theta)} \\
s'_{inc}(z(\theta)) &:= \frac{\partial s\left(z;\theta\right)}{\partial z}\Big|_{z=z(\theta)}
\end{aligned}$$

where $\mathcal{T}'_s\left(s\left(z;\theta\right),z\right)$ is the marginal savings tax rate of an agent of type θ who earns labor income z. These elasticity concepts are conditional on z. They measure responses of consumption and savings to tax reforms and nonlinearities in the tax system, holding labor income z fixed fixed at $z(\theta)$. For all elasticity concepts, we use the "bar" notation, as in $\overline{\zeta^c_{s|z}}$, to denote a population elasticity.

We define the elasticity of s with respect to one plus the marginal tax rate, rather than with respect to $p+\mathcal{T}'_s$ because such a measure is more common in the savings or bequest application. However, defining the elasticity with respect to $p+\mathcal{T}'_s$ may be more natural in applications where s is a commodity; with this alternative definition of elasticities, the key results in Propositions 2 and 2 can be obtained by simply multiplying $\zeta^c_{s|z}$ by $(p+\mathcal{T}'_s)/(1+\mathcal{T}'_s)$.[8]

To quantify correlated preference heterogeneity, we decompose the cross-sectional profile of s consumption. Intuitively, $s'(z)$, the cross-sectional change in s with respect to z, reflects both the causal income effect and the degree to which preferences are changing with earnings z. We thus define our measure of correlated preference heterogeneity, $s'_{pref}(z)$, as the difference between the cross-sectional variation of s along the earnings distribution and the causal income effect $s'_{inc}(z)$:[9]

$$s'_{pref}(z(\theta)) := s'(z(\theta)) - s'_{inc}(z(\theta))$$

To illustrate how $s'_{pref}(z)$ relates to model parameters, suppose that $U = \ln c + \delta(\theta)\ln s - \psi(z/\theta)$, where s is period-2 consumption, δ is the discount factor, and $\mathcal{T}$ is a *simple* tax system (see Section 4) that is separable in s and z. Then a z-earner chooses

$$s(z) = \frac{1}{p}\frac{\delta(z)}{1+\delta(z)}\left(z-\mathcal{T}\right) \tag{12}$$

[7]This corresponds to the type of circular adjustment process described in e.g. Jacquet and Lehmann (2020).

[8]In this case, the only change in Theorem 2 is that the left-hand-side in equation (16) becomes $\frac{\mathcal{T}'_s(s(z),z)}{p+\mathcal{T}'_s(s(z),z)}$, and analogously for Proposition 2 .

[9]Put another way, let $\vartheta(z)$ denote the type θ that earns z. Then $s'(z)$ is a total derivative given by $\frac{d}{dz}s(z,\vartheta(z)) = \frac{\partial}{\partial z}s(z,\theta)|_{\theta=\vartheta(z)} + \frac{\partial}{\partial\theta}s(z,\theta)|_{\theta=\vartheta(z)}$, where the first term measures causal income effects and corresponds to $s'_{inc}(z)$, and the second term measures local preference heterogeneity and corresponds to $s'_{pref}(z)$.

where $\delta(z)$ is used to denote the discount factor of the z-earner. Cross-sectional heterogeneity in s is given by

$$s'(z) = \underbrace{\frac{1}{p}\frac{\delta(z)}{1+\delta(z)}(1-\mathcal{T}'_z)}_{s'_{inc}(z)} + \underbrace{\frac{1}{p}(z-\mathcal{T})\frac{d}{dz}\left(\frac{\delta(z)}{1+\delta(z)}\right)}_{s'_{pref}(z)} \tag{13}$$

The causal income effect s'_{inc} is obtained by differentiating (12) with respect to z while holding the discount factor $\delta(z)$ constant. The local preference heterogeneity term $s'_{pref}(z)$ is obtained by differentiating (12) with respect to z while holding after-tax income $z - \mathcal{T}$ constant. The total derivative of s with respect to z that makes up the cross-sectional profile is the sum of the two partial derivatives. The key insight that facilitates measurement is that $s'(z)$ and $s'_{inc}(z)$ are commonly measured statistics that can be used to indirectly estimate $s'_{pref}(z)$, which is often more tractable than a direct estimate of how time preferences $\delta(\theta)$ vary across the income distribution.

3.2.2 Measurement

Because $s'(z)$ is the cross-sectional variation of savings along the income distribution, it can be directly observed from survey or administrative data featuring both earnings and savings. The result below shows how $s'_{inc}(z)$ can be measured in the data.

Proposition 1. *The sufficient statistic $s'_{inc}(z)$ can be measured as follows.*

- *If agents' preferences are weakly separable and the tax system is separable in s and z, $s'_{inc}(z) = \frac{1-T'_z(z)}{1+T'_s(s(z))}\eta_{s|z}(z)$*
- *If agents wage rates w and hours h are observable, $s'_{inc}(z) = s(z)\frac{\xi^s_w(z)}{w(z)+h(z)\xi^h_w(z)}$*
- *If we savings and earnings changes in response to tax reforms are observable, $s'_{inc}(z) = \frac{s(z)}{z}\frac{\chi^c_s(z)}{\zeta^c_z(z)}$,*

where $\xi^s_w(z)$ is the elasticity of savings with respect to the wage rate, $\xi^h_w(z)$ is the elasticity of hours with respect to the wage rate, and $\chi^c_s(z)$ is the elasticity of savings with respect the marginal net of tax rate on labor income.

This proposition shows how $s'_{inc}(z)$ can be related to empirically estimable elasticity concepts. If, as in example (1) above, agents' preferences are weakly separable between the utility of consumption and savings on the one hand, and the disutility of labor supply on the other hand, then $s'_{inc}(z)$ is proportional to the income effect parameter.

If agents' preferences are not weakly separable but wage rates w and hours h are observable, we can use the fact that labor income is given by the product of hours and the wage rate, $z = h \cdot w$, to leverage wage changes to measure $s'_{inc}(z)$. In that case, $s'_{inc}(z)$ can be related to the elasticity of savings with respect to the wage rate and to the elasticity of hours with respect to the wage rate.

Last, if both earnings and savings changes in response to a change in the marginal tax rate z are observable, $s'_{inc}(z)$ can be recovered from the ratio of savings responses to labor income responses. The intuition is that savings responses to this type of tax reform are induced by changes in labor income.

3.3 Sufficient Statistics Characterization of Optimal Smooth Tax Systems

A key result for our sufficient statistics formulas is the following equivalence lemma for tax reforms affecting marginal tax rates on s versus z. This result is a generalization of Lemma 1 from Saez (2002) to a broader class of tax systems.

Lemma 1. *A small increase $d\tau_s$ in the marginal savings tax rate faced by agent θ at earnings z induces the same earnings change as a small increase $s'_{inc}(z)\, d\tau_s$ in the marginal earnings tax rate.*

Intuitively, an agent who reduces earnings by dz reduces savings by $s'_{inc}dz$. Thus, when savings tax rate increases by $d\tau_s$, reducing earnings by dz reduces tax liability by an additional $s'_{inc}d\tau_s dz$. The Lemma shows that a $d\tau_s$ increase in the marginal savings tax generates a similar reduction in earnings as a $s'_{inc}d\tau_s$ increase in the earnings tax rate.

Lemma 1 relates the labor supply distortions induced by savings taxes to the labor supply distortions induced by income taxes. Specifically, it shows that increasing savings tax rates will be more distortionary when $s'_{inc}(z)$ is higher.

To encode the policymaker's redistributive objective, we define $\hat{g}(z)$ as the social marginal welfare weights augmented with the fiscal impact of income effects. These capture the social value of marginally increasing the disposable income of agents earning z. Formally,

$$\begin{aligned}\hat{g}(z(\theta)) := \frac{\alpha(\theta)}{\lambda} U'_c(c(\theta), s(\theta), z(\theta); \theta) + \mathcal{T}'_z(s(\theta), z(\theta)) \frac{\eta_z(z(\theta))}{1 - \mathcal{T}'_z(s(\theta), z(\theta))} \\ + \mathcal{T}'_s(s(\theta), z(\theta)) \left(\frac{\eta_{s|z}(z(\theta))}{1 + \mathcal{T}'_s(s(\theta), z(\theta))} + s'_{inc}(z(\theta)) \frac{\eta_z(z(\theta))}{1 - \mathcal{T}'_z(s(\theta), z(\theta))} \right),\end{aligned} \tag{14}$$

where the last term captures the fact that income effects on earnings proportional to $\eta_z(z(\theta))$ induce changes in savings proportional to $s'_{inc}(z(\theta))$ affecting savings tax revenues. We use $g(z(\theta))$ to denote the first component, $\frac{\alpha(\theta)}{\lambda} U'_c(c(\theta), s(\theta), z(\theta); \theta)$.

In the results that follow, we use $H(s,z)$ and $h(s,z)$ to denote the cumulative and density functions over (s,z), we let h_s denote the marginal density over s, and we let h_z denote the marginal density over z.

Theorem 2. *At each bundle (c,s,z) chosen by a type θ, an optimal smooth tax system satisfies*

the following conditions on marginal tax rates on z and s, respectively:

$$\frac{\mathcal{T}_z'(s,z)}{1-\mathcal{T}_z'(s,z)} = \frac{1}{z\,\zeta_z^c(z)}\frac{1}{h_z(z)}\int_{x\geq z}(1-\hat{g}(x))\,dH_z(x) - s'_{inc}(z)\frac{\mathcal{T}_s'(s,z)}{1-\mathcal{T}_z'(s,z)} \tag{15}$$

$$\frac{\mathcal{T}_s'(s,z)}{1+\mathcal{T}_s'(s,z)} = s'_{pref}(z)\frac{1}{s\,\zeta_{s|z}^c(z)}\frac{1}{h_z(z)}\int_{x\geq z}(1-\hat{g}(x))\,dH_z(x) \tag{16}$$

Any Pareto efficient smooth tax system satisfies

$$\frac{\mathcal{T}_s'(s,z)}{1+\mathcal{T}_s'(s,z)} = s'_{pref}(z)\frac{z\,\zeta_z^c(z)}{s\,\zeta_{s|z}^c(z)}\frac{\mathcal{T}_z'(s,z)+s'_{inc}(z)\mathcal{T}_s'(s,z)}{1-\mathcal{T}_z'(s,z)} \tag{17}$$

Formula (15) constitutes a familiar condition analogous to Saez (2001), with one modification. Because there is a non-zero savings tax rate, the formula also accounts for the distortionary effects on savings caused by distortions to earnings levels.

Formula (16) presents a transparent generalization of the Atkinson-Stiglitz Theorem. When the sufficient statistic for preference heterogeneity, s'_{pref}, is equal to zero, the condition implies that the optimal tax on s must equal zero as well. When the statistic is $s'_{pref} > 0$, implying that higher earners have a higher taste for s, the condition implies that the optimal tax rate on s must be positive.

Optimal tax rates on s satisfy a formula that is remarkably similar to the standard "ABC" formula for optimal income tax rates. When $s'_{pref} > 0$, the magnitude of the optimal tax rate at point (s,z) is decreasing in the elasticity of s with respect to the tax rate, increasing in the strength of redistributive motives, and decreasing in the relative density of individuals at point (s,z).

We can combine conditions (15) and(16) to derive the Pareto efficiency condition in (17). Because the condition in (17) does not feature social marginal welfare weights, it is an efficiency condition that must hold for any tax system that is not Pareto dominated. Intuitively, it quantifies the efficient balance between taxing s and taxing z, given the measure s'_{pref} of how tastes for s vary with earnings ability. The larger is the association between tastes for s and earnings ability, the more efficient it is to tax s instead of z.

An implication of the Pareto efficiency condition in (17) is that in the absence of preference heterogeneity, positive savings tax rates are Pareto dominated, providing an extension of the Atkinson-Stiglitz Theorem for nonlinear tax systems. On the other hand, any Pareto efficient tax system must feature nonzero savings tax rates in the presence of preference heterogeneity.

3.4 Other Determinants of Tax Elasticities Captured by our Sufficient Statistics Formulas

In practice, the elasticities and measures of preference heterogeneity that we utilize may be shaped by other economic considerations such as different agents facing different prices (Gahvari and Micheletto, 2016), *income shifting* (Slemrod, 1995; Christiansen and Tuomala, 2008), and heterogeneity in endowments (Boadway et al., 2000; Cremer et al., 2003). A key feature of our sufficient

statistics characterization of optimal tax schedules (including the characterization of "simple" tax systems in Section 4 below) is that they apply even in the presence of these other determinants of taxpayer behavior. The economic intuition is in the spirit of the Feldstein (1999) result that the elasticity of taxable income is a sufficient statistic for efficiency losses irrespective of whether it is due to real labor supply responses or costly avoidance behavior.[10] While these other mechanisms have previously been studied in isolation in specialized models that qualitatively address the robustness of the Atkinson-Stiglitz Theorem, our sufficient statistics techniques allow us to account for these considerations in a quantitative and general manner.

Heterogeneous Prices. Suppose that individuals face prices $p(e, z, s, \theta)$ that might depend on effort e to seek out lower prices, on types θ (e.g., higher types are better able to seek out lower prices or higher returns on investment), income z (e.g., being richer generates beneficial network effects that expose individuals to better opportunities), and the choice of s (e.g., higher levels of the commodity or savings allow individuals to lock in better prices or interest rates). Individuals' utility function is now $U(c, s, z, e; \theta)$ and their budget constraint is

$$c + p(e, z, s, \theta)s \leq z - \mathcal{T}(s, z)$$

This economy is then equivalent to an economy where $p \equiv 1$ and where individuals maximize the utility function

$$\tilde{U}(c, s, z; \theta) = \max_{e} U\left(c + (1-p)s, s, z, e; \theta\right) \tag{18}$$

subject to the budget constraint $c + s \leq z - \mathcal{T}(s, z)$. This is because with a price of $p = p'$ instead of $p = 1$, individuals get $(1-p')s$ more consumption c at a given choice s. The feasibility constraint $\int_\theta \mathcal{T}\left(s\left(\theta\right), z\left(\theta\right)\right) dF\left(\theta\right) \geq E$ is independent of the price p because the tax is deducted directly from an individual's earnings z.[11] Thus, it is without loss of generality to assume that $p \equiv 1$, which is a normalization we adopt in our proofs.

Type-dependent heterogeneity in prices will generally lead to $s'_{pref}(z) \neq 0$ and thus deviations from the Atkinson-Stiglitz result. For example, suppose that types θ see prices $p(\theta)$. Then heterogeneity in prices functions much like heterogeneity in discount factors. Suppose again that $U = \ln c + \delta \ln s - \psi(z/\theta)$ and that $\mathcal{T}$ is separable. Then, as in the example with heterogeneous discount rates, $s'_{pref}(z) = \delta\left(z - \mathcal{T}\right) \frac{d}{dz}\left(\frac{1}{p(z)}\right)$, where $p(z)$ is the price seen by z-earners.

Heterogeneity in prices that is due to scale effects, however, does not contribute to s'_{pref} . For example, suppose that $U = \ln c + \delta \ln s - \psi(z/\theta)$ and p is an increasing function of s (but not θ) because higher levels of savings garner higher rates of return. This scale dependence would increase

[10] Chetty (2009) suggests limitations to Feldstein's (1999) results due to some avoidance behaviors generating new types of fiscal externalities, or due to behavioral biases. Variations of these considerations are relevant in our setting as well, as explored in Sections 5.3 and 5.4, respectively.

[11] For example, if s is units of soda purchased, the tax applies to individuals' earnings directly because it is units of dollars. However, if the tax had a two-part structure where the individual must pay T_1 to in units of c to the government and T_2 in units of s to the government then the equivalence above would not apply. We explore this in Section 5.4, extending the Gerritsen et al. (2020) analysis of capital income taxation with heterogeneous rates of return.

s'_{inc} but would leave $s'_{pref} \equiv 0$ in the case where δ is homogeneous.

Income Shifting. Slemrod (1995) points out that some tax system may provide incentives for individuals to "convert ordinary income into preferentially taxed capital gains," "converting corporations into pass-through legal entities such as partnerships, or retaining labor compensation within the corporation." Generalizing the two-type model in Christiansen and Tuomala (2008), suppose that individuals can shift χ units of their labor income to capital income or savings, so that the observable pre-tax labor income that they are taxed on is $z = \tilde{z} - \chi$, where $\tilde{z}$ is actual earned income. Individuals choose z, c, s, χ to maximize $U(c, s, z, \chi; \theta)$ subject to the budget constraint $c + p = s \leq z - \mathcal{T}(s, z) - m(\chi; \theta)$. This formulation allows shifting to both enter the utility function directly—e.g., due to effort or psychic costs—as well as through the budget constraint because it may require financial resources m. This formulation also allows for the natural possibility individuals with higher earnings ability may be better able to engage in income shifting. The individual's budget constraint implies that consumption of c is given by

$$\begin{aligned} c &= \tilde{z} - \mathcal{T}(s, z) - s - m(\chi) \\ &= z + \chi - \mathcal{T}(s, z) - s - m(\chi) \end{aligned}$$

Thus, we can alternatively rewrite this as an optimization problem where individuals maximize

$$\tilde{U}(c, s, z; \theta) = \max_{\chi} U\left(c + \chi - m(\chi, \theta), s, z + \chi, \chi; \theta\right)$$

subject to the budget constraint that $c+s \leq z-\mathcal{T}(s, z)$. The feasibility constraint $\int_\theta \mathcal{T}\left(s\left(\theta\right), z\left(\theta\right)\right) dF\left(\theta\right) \geq E$ is independent of the degree to which agents engage in income shifting because the tax is deducted directly from an individual's observed earnings z.

Heterogeneous Endowments. Suppose that individuals have endowments $y_0(\theta)$, from inheritance or other sources, such that their budget constraint is given by $c + ps \leq y_0(\theta) + z - \mathcal{T}(s, z)$. This economy is then equivalent to an economy where $y_0 \equiv 0$ and where individuals maximize the utility function

$$\tilde{U}(c, s, z; \theta) = U\left(c + y_0(\theta), s, z; \theta\right) \tag{19}$$

subject to the budget constraint $c + s \leq z - \mathcal{T}(s, z)$. This is because at a given choice s, individuals get y_0 more consumption c. Again, the feasibility constraint $\int_\theta \mathcal{T}\left(s\left(\theta\right), z\left(\theta\right)\right) dF\left(\theta\right) \geq E$ is independent of the endowments.

4 Optimal "Simple" Tax Systems

In practice, tax systems must be defined by policymakers and implemented by institutions who face constraints on the degree of complexity that can be accommodated. As a result, typical tax

functions are not fully flexible. Although the details of the restrictions vary across institutions, most can be classified into a few common functional forms.

We consider two questions in this section. First, does the importance of preference heterogeneity, as measured by s'_{pref}, extend to simple tax systems? Second, when can simple tax systems implement the optimal allocation?

4.1 A Taxonomy of Common Simple Savings Tax Systems

Many governments tax both labor income earnings and savings / capital interest income. While these tax systems take a variety of forms, the details of which depend on specifics such as timing, many of these tax policies can be interpreted as a function of earnings and savings, analogous to our function $\mathcal{T}(s, z)$. As Table 2 summarizes, many of these tax systems can be categorized as one of three simple types—separable linear (SL), separable nonlinear (SN), and linear income-dependent (LED)—characterized in Table 1. We categorize the tax policies on each class of savings vehicle for 21 countries, most of which fit one of the three simple tax system types from Table 1. In cases where there is some ambiguity (such as the distinction between short-term and long-term capital gains in the United States) we provide supplementary information in Appendix D.

Type of simple tax system	$\mathcal{T}(s,z)$	$\mathcal{T}'_s(s,z)$	$\mathcal{T}'_z(s,z)$
SL: separable linear	$\tau_s\, s + T_z(z)$	τ_s	$T'_z(z)$
SN: separable nonlinear	$T_s(s) + T_z(z)$	$T'_s(s)$	$T'_z(z)$
LED: linear earnings-dependent	$\tau_s(z)\, s + T_z(z)$	$\tau_s(z)$	$T'_z(z) + \tau'_s(z)\, s$

Table 1: Types of simple tax systems

We impose several simplifications for our interpretation of the tax codes. First, we treat ordinary income as consisting primarily of labor income (earnings), written as z in our notation. Second, we separately consider taxes on five broad categories of savings vehicles: wealth, capital gains, real property, private pensions, and inheritances.[12]

Each of the three tax systems in Table 1 can for instance be found in the United States. Most property taxes (levied at the state and local level) take the form of separable linear taxes, with a flat tax rate (independent of one's labor earnings) applied to the assessed value of the total property. The estate tax takes the form of a separable nonlinear tax: the tax rate rises progressively with the value of the estate, but it does not vary with labor income of the donor or the recipient. Finally, taxes on long-term capital gains and on qualified dividends both take the form of linear earnings-dependent taxes: in 2020, for example, an individual with $50,000 in labor earnings faced a long-term capital gains tax rate of 15%, whereas an individual earning $500,000 faced a rate of 20%.[13]

[12] These categories may overlap—real property is a component of wealth, for example—but we use these groups to reflect the tax instruments that many governments use in practice.

[13] Given our application to taxation of savings, bequests, and capital interest income, where individuals consume c in period 1 and s in period 2, it may be more natural to formalize the tax function as imposing a period-1 labor

4.2 Optimal Simple Tax Systems

We now present optimality conditions for SL, SN and LED tax systems. We focus on marginal tax rates on s in the body of the paper, and present conditions for marginal tax rates on z in Appendix (A.3). The preference heterogeneity statistic s'_{pref} remains the key sufficient statistic for the marginal tax rate on s. For SL systems, it is convenient to introduce the term $s_{pref}(z) = \int_{x=\underline{z}}^{z} s'_{pref}(x)dx$, which integrates local preference heterogeneity to obtain a measure of total preference heterogeneity for with gross earnings of z or less.

Proposition 2. *Consider simple tax systems where agents' optima are unique and agents' first-order and second-order conditions strictly hold. At each bundle (c, s, z) chosen by a type θ, SL, SN, and LED systems satisfy the following optimality conditions on the savings tax rates:*

$$SL: \frac{\tau_s}{1+\tau_s} = \frac{1}{\overline{s}\overline{\zeta^c_{s|z}}} \int_z s'_{pref}(z) \left[\int_{x \geq z} (1 - \hat{g}(x)) dH_z(x) \right] dz \tag{20}$$

$$= -\frac{1}{\overline{s}\overline{\zeta^c_{s|z}}} Cov[s_{pref}(z), \hat{g}(z)] \tag{21}$$

$$SN: \frac{T'_s(s)}{1+T'_s(s)} = \frac{1}{s\,\zeta^c_{s|z}(z)} \frac{1}{h_z(z)} s'_{pref}(z) \int_{x \geq z} (1 - \hat{g}(x))\, dH_z(x) \tag{22}$$

$$LED: \frac{\tau_s(z)}{1+\tau_s(z)} = \frac{1}{s\,\zeta^c_{s|z}(z)} \frac{1}{h_z(z)} s'_{pref}(z) \int_{x \geq z} (1 - \hat{g}(x))\, dH_z(x) \tag{23}$$

Moreover, if a SL/SN/LED tax system is not pareto dominated by another SL/SN/LED system, then it must satisfy the following conditions:

$$SL: \frac{\tau_s}{1+\tau_s} = \frac{1}{\overline{s}\overline{\zeta^c_{s|z}}} \int_{x=\underline{z}}^{\overline{z}} s'_{pref}(x)\; z\zeta^c_z(x)\; \frac{T'_z(x) + s'_{inc}(x)\tau_s}{1 - T'_z(x)}\, dH_z(x) \tag{24}$$

$$SN: \frac{T'_s(s)}{1+T'_s(s)} = s'_{pref}(z)\; \frac{z\zeta^c_z(z)}{s\zeta^c_{s|z}(z)}\; \frac{T'_z(z) + s'_{inc}(z)T'_s(s)}{1 - T'_z(z)} \tag{25}$$

$$LED: \frac{\tau_s(z)}{1+\tau_s(z)} = s'_{pref}(z)\; \frac{z\zeta^c_z(z)}{s\zeta^c_{s|z}(z)}\; \frac{T'_z(z) + \tau'_s(z)\,s + s'_{inc}(z)\tau_s(z)}{1 - T'_z(z) - \tau'_s(z)\,s} \tag{26}$$

The optimal tax formulas and the pareto efficiency conditions for SN and LED systems are

income tax and a period-2 tax on gross savings. Let $1 + r = 1/p$ denote the rate of transformation between c and s, so that r can be interpreted as the interest rate. Now consider separable tax functions where the policymaker imposes a tax $T_1(z)$ in period 1, and imposes a tax $T_2(s_g, z)$ on gross pre-tax savings $s_g = (1 + r)(z - T_1(z) - c)$ in period 2, so that period-2 consumption is given by $s = s_g - T_s(s_g, z)$. In this formulation, a SL structure is one where $T_2(s_g, z) = \tau_s s_g$, a SN structure is one where $T_2(s_g, z)$ is a function of s_g only, and a LED structure is one where $T_2(s_g, z) = \tau_s(z)s_g$. As we show in Appendix (A.5), these formulations of two-period tax systems are equivalent to the ones we study in the body of the paper in two fundamental ways. First, the set of allocations implementable by these systems is identical, as there is a simple and natural translation between the "static" tax function $\mathcal{T}$ we consider and the two-period function. Second, if $T_1(z) + T_2(s_g, z)$ implements the same allocation as $\mathcal{T}(s, z)$, then $T_1(z) + T_2(s_g, z)$ will be SL/SN/LED if and only if $\mathcal{T}(s, z)$ is SL/SN/LED.

essentially identical to the conditions for fully flexible smooth tax systems derived in Theorem (2). This equivalence is guaranteed for conditions under which SN and LED systems can implement the optimal allocation, which we summarize in Section (4.3). However, there are also cases in which SN or LED systems cannot implement the same allocation as the optimal nonlinear tax system. In these cases, Proposition (3) shows that the formulas for the optimum are still similar, even if the resulting allocations are not.

The SL system is the most restrictive, and generically cannot implement the same allocation as the optimal nonlinear tax system. This is because the optimal nonlinear tax system does not generally feature constant marginal tax rates on s. Correspondingly, the optimal tax formula for the SL systems takes a different form. As shown in expression (20), the constant marginal tax rate τ_s for SL systems is in a certain sense an average of the z-earner specific marginal tax rates in expressions (22) and (23). Intuitively, the constant marginal tax rate is a population aggregate of the tax rates that would be optimal for individuals with different earnings levels. Expression (20) is also equivalent to expression (21), which has a familiar form resembling the Diamond (1975) "Many-person Ramsey tax rule." The expression is identical to the Diamond (1975) formula when $s_{pref}(z) \equiv s(z)$; i.e., when all consumption is driven by preference heterogeneity. This illustrates both the relevance of Ramsey taxation principles even in the presence of nonlinear taxation, as well as their limitations. The SL formula when $s_{pref}(z) \equiv 0$ reduces to the original statement of the Atkinson-Stiglitz theorem. More generally, it is useful to note that even for arbitrarily nonlinear taxes on s, the optimal tax rule rate is always inversely proportional to the elasticity $\zeta^c_{s|z}(z)$, consistent with standard Ramsey principles, as long as $s'_{pref}(z) \neq 0$.

4.3 When can Simple Savings Tax Systems Implement the Optimum?

In Appendix A.1 we present sufficient conditions for the optimal allocation to be implementable by a SN or LED tax system. We assume that the optimal allocation satisfies the conditions in Theorem 1, implying that it can be implemented by *some* smooth tax system, so that the question of interest is whether SN or LED systems are among those that implement the optimum. We focus on SN and LED systems because the SL system is plainly too restrictive to implement the optimum. We summarize the results here.

We proceed in three steps to characterize the sufficient conditions. First, we define candidate SN and LED tax systems constructed to satisfy type-specific feasibility and individuals' first-order conditions at the optimal allocation, as in Section 3.1. This completely defines candidate tax systems in terms of agents' marginal rates of substitutions $\mathcal{S}$ and $\mathcal{Z}$, and the optimal incentive-compatible allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_\theta$.

Second, in Proposition A.1 we present sufficient conditions under which these candidate SN and LED tax systems also satisfy the local second-order conditions, implying that each type's assigned allocation is a local optimum. The sufficient conditions for the SN tax system are relatively weak, and are satisfied by many common utility functions used in calibrations of savings and income taxation models, including the example in Equation (1). The sufficient conditions for the LED tax

system are somewhat more restrictive, as they place a constraint on the extent of local preference heterogeneity. A high degree of local preference heterogeneity at income z implies a large positive savings tax rate for any individual choosing earnings z. This large increase in the savings tax liability must then be compensated by a sharp decrease in the earnings tax liability, which can create incentives for deviation.

Third, in Proposition A.2 we derive sufficient conditions under which local optima are ensured to be global optima for SN and LED tax systems. These conditions involve separability of agents' preferences and further concavity requirements on preferences for consumption and savings. Together with Proposition A.1, this produces a characterization of when SN and LED tax systems implement the optimal allocation.

5 Extensions

In this section we provide four key extensions of our results. First, we generalize our results on optimal tax rates on s to cases with multidimensional heterogeneity and potentially suboptimal income taxes. Second, we generalize our results to many dimensions of consumption. Third, we extend our results to the possibility that the government wants people to save more than they are, either because of inter-generational social preferences or because of behavioral biases. Fourth, we consider the case where tax systems have a *multidimensional range* so that, for example, they are collected both in period 1 (in units of c) and in period 2 (in units of s).

5.1 Multidimensional Heterogeneity and Non-Optimal Earnings Tax

In this section, we characterize optimal simple tax systems with multidimensional heterogeneity. We show that s'_{inc} remains the key sufficient statistic. Given the accounting identify $s' = s'_{pref} + s'_{inc}$, this implies that s'_{pref} remains the key sufficient statistics for preference heterogeneity.

We focus on simple tax systems, both because simple tax systems are the most practically relevant, and because the optimal mechanism can feature atoms, discontinuities, and randomization in the presence of multidimensional heterogeneity. Proposition 3 below generalizes Proposition 2 in two key ways. First, it allows for multidimensional heterogeneity, where types θ belong to subset of $\mathbb{R}^n$ for $n \geq 2$, so that the support of the distribution over (s, z) is two-dimensional. Second, it characterizes optimal taxes on s even when the earnings tax $T_z(z)$ is not necessarily optimal. Proposition (A.6) in the Appendix characterizes the optimal earnings tax schedule.

We use $\zeta^c_z(s, z)$ to define the compensated elasticity of s of an individual choosing the bundle (s, z). In the unidimensional case, ζ^c_z was a function of z only by assumption. We utilize the notation $\hat{g}(s, z)$ and $s'_{inc}(s, z)$ analogously.

Proposition 3. *Consider a given (and potentially suboptimal) earnings tax schedule $T_z(z)$, and assume that under each simple tax system agents first-order and second-order conditions strictly hold. Then, at each bundle (c^0, s^0, z^0) chosen some type θ, the marginal tax rates on s in SL/SN/LED systems must satisfy the following optimality conditions:*

$$SL: \frac{\tau_s}{1+\tau_s}\int_{z=\underline{z}}^{\overline{z}}\left\{\mathbb{E}\left[s\zeta^c_{s|z}(s,z)\Big|z\right]\right\}dH_z(z) = \int_{z=\underline{z}}^{\overline{z}}\left\{\mathbb{E}\left[(1-\hat{g}(s,z))\,s\Big|z\right]\right.$$
$$\left.-\,\mathbb{E}\left[\frac{T'_z(z)+s'_{inc}(s,z)\,\tau_s}{1-T'_z(z)}z\zeta^c_z(s,z)\,s'_{inc}(s,z)\Big|z\right]\right\}dH_z(z) \tag{27}$$

$$SN: \frac{T'_s(s^0)}{1+T'_s(s^0)}\int_{z=\underline{z}}^{\overline{z}}\left\{s^0\zeta^c_{s|z}(s^0,z)\right\}h(s^0,z)\,dz = \int_{z=\underline{z}}^{\overline{z}}\left\{\mathbb{E}\Big[1-\hat{g}(s,z)\Big|z,s\geq s^0\Big]\right\}dH_z(z)$$
$$-\int_{z=\underline{z}}^{\overline{z}}\left\{\frac{T'_z(z)+s'_{inc}(s^0,z)\,T'_s(s^0)}{1-T'_z(z)}z\zeta^c_z(s^0,z)s'_{inc}(s^0,z)\right\}h(s^0,z)\,dz \tag{28}$$

$$LED: \mathbb{E}\left[\frac{T'_z(z)+\tau'_s(z)\,s+s'_{inc}(s,z)\,\tau_s(z)}{1-T'_z(z)-\tau'_s(z)\,s}z\zeta^c_z(s,z)s\Big|z^0\right]h_z(z^0) = \int_{z\geq z^0}\mathbb{E}\Big[(1-\hat{g}(s,z))\,s\Big|z\Big]dH_z(z)$$
$$-\int_{z\geq z^0}\left\{\mathbb{E}\left[\frac{\tau_s(z)}{1+\tau_s(z)}s\zeta^c_{s|z}(s,z)\Big|z\right]+\mathbb{E}\left[\frac{T'_z(z)+\tau'_s(z)\,s+s'_{inc}(s,z)\,\tau_s(z)}{1-T'_z(z)-\tau'_s(z)\,s}z\zeta^c_z(s,z)s'_{inc}(s,z)\Big|z\right]\right\}dH_z(z) \tag{29}$$

The main difference between multidimensional and unidimensional heterogeneity is that with unidimensional heterogeneity, the expectation operators $\mathbb{E}[\cdot|z]$ are not needed. In expression (27), $\hat{g}(s,z)$, $\zeta^c_z(s,z)$, and $s'_{inc}(s,z)$ are functions of z only with unidimensional heterogeneity, which implies that the operator $\mathbb{E}[\cdot|z]$ is not needed with unidimensional heterogeneity. In expression (28), the term $\mathbb{E}\Big[1-\hat{g}(s,z)\Big|z,s\geq s^0\Big]$ reduces to $\mathbb{E}\Big[1-\hat{g}(z)\Big|z\geq z^0\Big]$ in the unidimensional case, and the functions s'_{inc} and h can be written as functions of z only. Analogous simplifications apply to expression (29).

When all terms inside the expectation operators $\mathbb{E}[\cdot|z]$ in Proposition A.5 are independent of each other, the expectation can be applied to each separate statistic separately, and thus the unidimensional formulas are similar to the multidimensional formulas provided that all statistics are interpreted as averages conditional on z. For example, the first term in the integral in expression (27) can be written as $\left(1-\overline{\hat{g}(z)}\right)\overline{s(z)}$, where the "bar" notation denotes income-conditional averages. The second term in the integral can be written as

$$\frac{T'_z(z)\,\overline{s'_{inc}(z)}+\tau_s\overline{s'_{inc}(s,z)^2}}{1-T'_z(z)}z\overline{\zeta^c_z(z)}$$

The main new effect is the square of s'_{inc} inside the integral. Because $\int(s'_{inc})^2 > \left(\int s'_{inc}\right)^2$ and because the square enters into the optimal tax expression negatively, this implies that ignoring multidimensional heterogeneity can lead to over-estimates of optimal marginal tax rates on s. The formulas in (28) and (29) also involve squares of s'_{inc}, and thus also imply that multidimensional heterogeneity can lower the optimal tax rate on s through $(s'_{inc})^2$. We quantify the importance of this insight in our empirical application in Section 6. More generally, positive covariances between pairs of statistics inside the expectation operator will tend to lower the optimal tax rate on s, while negative covariances will tend to increase it.

5.2 Multiple Goods

We now extend our analysis to a setting where agents consume $n+1$ goods, c and $\boldsymbol{s} = (s_1, s_2, ..., s_n)$. We consider a tax system $\mathcal{T}(\boldsymbol{s}, z) = \mathcal{T}(s_1, s_2, ..., s_n, z)$, where we retain the normalization that the numeraire c is untaxed. We normalize $\boldsymbol{s} = (s_1, s_2, ..., s_n)$ to measure consumption in units of the numeraire. An agent θ then maximizes $U(c, \boldsymbol{s}, z; \theta)$ subject to the budget constraint $c + \sum_{i=1}^{n} s_i \leq z - \mathcal{T}(\boldsymbol{s}, z)$.

We denote own-price elasticities of goods s_j by $\zeta^c_{s_i|z}(z)$, and we define cross-substitution elasticities by $\xi^c_{s_{j,i}|z}(z) := -\frac{\mathcal{T}'_{s_i}(\boldsymbol{s}(z;\theta),z)}{s_j(z;\theta)} \frac{\partial s_j(z;\theta)}{\partial \mathcal{T}'_{s_i}(\boldsymbol{s}(z;\theta),z)}\big|_{\theta=\vartheta(z)}$, where $s_j(z;\theta)$ denotes agent θ consumption of good j when earning labor income z. We denote causal income effects on good s_j by $s'_{j,inc}(z) := \frac{\partial s_j(z;\theta)}{\partial z}\big|_{\theta=\vartheta(z)}$. We continue using $\hat{g}(z)$ to denote the social marginal welfare effect of increase a z-earner's consumption of c by one unit; the formula in this more general setting is in Appendix B.8.

Proposition 4. *With n taxed goods $s_1, ..., s_n$, for each good i and at each bundle $(c, \boldsymbol{s}, z)$ chosen by a type θ, an optimal smooth tax system satisfies*

$$\frac{\mathcal{T}'_{s_i}(\boldsymbol{s}, z)}{1 + \mathcal{T}'_{s_i}(\boldsymbol{s}, z)} = s'_{i,pref}(z) \frac{1}{s_i \zeta^c_{s_i|z}(z)} \frac{1}{h_z(z)} \int_{x=z}^{\bar{z}} \Big[1 - \hat{g}(x)\Big] dH_z(x) + \underbrace{\sum_{j \neq i} \frac{\mathcal{T}'_{s_j}(\boldsymbol{s}, z)}{\mathcal{T}'_{s_i}(\boldsymbol{s}, z)} \frac{s_j \xi^c_{s_{j,i}|z}(z)}{s_i \zeta^c_{s_i|z}(z)}}_{\textit{Tax diversion ratio}}. \quad (30)$$

Any Pareto-efficient smooth tax system satisfies

$$\frac{\mathcal{T}'_{s_i}(\boldsymbol{s}, z)}{1 + \mathcal{T}'_{s_i}(\boldsymbol{s}, z)} = s'_{i,pref}(z) \frac{z \zeta^c_z(z)}{s_i \zeta^c_{s_i|z}(z)} \frac{\mathcal{T}'_z(\boldsymbol{s}, z) + \sum_{j=1}^{n} s'_{j,inc}(z) \mathcal{T}'_{s_j}(\boldsymbol{s}, z)}{1 - \mathcal{T}'_z(\boldsymbol{s}, z)} + \underbrace{\sum_{j \neq i} \frac{\mathcal{T}'_{s_j}(\boldsymbol{s}, z)}{\mathcal{T}'_{s_i}(\boldsymbol{s}, z)} \frac{s_j \xi^c_{s_{j,i}|z}(z)}{s_i \zeta^c_{s_i|z}(z)}}_{\textit{Tax diversion ratio}}. \quad (31)$$

Proposition 4 features all of the familiar terms of Theorem 2, and includes a novel term that captures the tax implications of substitution effects. Intuitively, substituting from s_i to higher-taxed goods generates positive fiscal externalities that motivate higher marginal tax rates on s_i, while substitution to lower-taxed goods generates negative fiscal externalities that motivate lower marginal tax rates on s_i. These effects are summarized by what we call the tax diversion ratio—the impact on taxes collected on goods $j \neq i$ relative to the impact on taxes collected on good i when the price of good i is increased. The higher is the diversion ratio, the more favorable are the fiscal externalities associated with substitution away from good i, and thus the higher is the optimal tax rate on good i.

5.3 Optimal Taxation when the Government Wants Agents to Save More

Our framework can be interpreted as a bequest model in which parents work and consume in the first period, and leave a bequest to their heirs in the second period. However, our main results

only apply under the assumption that the government's preferred level of bequests corresponds to parents' choices. Farhi and Werning (2010) consider a more general model where the weight that parents attach to the wellbeing of future generations may be too low from a normative perspective. This misalignment introduces a policy motive to encourage bequests.

Following Farhi and Werning (2010), we assume additively separable preferences given by

$$U\big(c,s,z;\theta\big)=u\left(c;\theta\right)-k\left(z;\theta\right)+\beta v\left(s;\theta\right), \tag{32}$$

where $u\left(c;\theta\right)$ is the utility parents derive from consumption c, $k\left(z;\theta\right)$ is the disutility parents incur to obtain earnings z, $v\left(s;\theta\right)$ is the utility heirs derive from a bequest s, and β is the weight parents attach to the wellbeing of their heirs. As in Farhi and Werning (2010), the government maximizes

$$\int_{\theta}\big[U\left(c\left(\theta\right),s\left(\theta\right),z\left(\theta\right);\theta\right)+\nu v\left(s\left(\theta\right);\theta\right)\big]dF\left(\theta\right), \tag{33}$$

where ν parametrizes the degree of misalignment. Farhi and Werning (2010) microfound ν as the Lagrange multiplier associated with a constraint that the future generation attains a required level of well-being $\int_{\theta}\nu v\left(s\left(\theta\right);\theta\right)dF\left(\theta\right)\geq\underline{V}$.

The formal model above can be interpreted more generally beyond the bequest application, and can be used to analyze behavioral biases as a motivation for encouraging savings. In particular, suppose that $v(x;\theta)=\delta(\theta)u(x)$, where δ is the "exponential discount factor" and β is "present focus," as in Laibson (1997). If the government utilizes the "long-run criterion" for welfare, then the degree of misalignment is given by $\nu=(1-\beta)$.[14] More generally, β may be heterogeneous, so that misalignment is type-dependent and given by $\nu(\theta)=(1-\beta(\theta))$. For example, Lockwood (2020) summarizes evidence suggesting that individuals with higher earnings ability have lower degrees of present focus.

Below, we characterize optimal taxation with heterogeneous misalignment, where $\beta(z)$ and $\nu(z)$ denote the parameters corresponding to a z-earner. This generalizes the result in Farhi and Werning (2010) by (i) allowing heterogeneity in preferences for s and (ii) by allowing heterogeneity in the misalignment parameter ν.

Proposition 5. *At each bundle (c,s,z) chosen by a type θ, an optimal smooth tax system satisfies the following marginal tax rate condition*

$$\frac{\mathcal{T}_{s}'\left(s,z\right)}{1+\mathcal{T}_{s}'\left(s,z\right)}=s_{pref}'\left(z\right)\frac{1}{s\zeta_{s|z}^{c}(z)}\frac{1}{h_{z}(z)}\int_{x\geq z}\Big[1-\hat{g}(x)\Big]\,dH_{z}(x)-\frac{\nu(z)}{\beta(z)}g(z). \tag{34}$$

Moreover, any Pareto-efficient smooth tax system satisfies

$$\frac{\mathcal{T}_{s}'\left(s,z\right)}{1+\mathcal{T}_{s}'\left(s,z\right)}=s_{pref}'\left(z\right)\frac{z\zeta_{z}^{c}(z)}{s\zeta_{s|z}^{c}(z)}\left[\frac{\mathcal{T}_{z}'\left(s,z\right)+s_{inc}'(z)\mathcal{T}_{s}'\left(s,z\right)}{1-\mathcal{T}_{z}'\left(s,z\right)}+s_{inc}'(z)\frac{\nu(z)}{\beta(z)}g(z)\right]-\frac{\nu\left(z\right)}{\beta\left(z\right)}g(z) \tag{35}$$

[14] See Bernheim and Taubinsky (2018) for a detailed discussion of such a criterion, as well as alternative normative approaches to studying the implications of present focus.

The main result is an intuitive generalization of Proposition 2, where the key new term is a form of pigovian correction, given by $\frac{\nu(z)}{\beta(z)}g(z)$. As equation (34) shows, the presence of misalignment motivates the government to lower the tax rate on s. The degree by which the government lowers the tax rate depends on the relative degree of misalignment (relative to the discount factor β), and on the social marginal welfare weight. Because social marginal welfare weights decline with z, equation (34) gives the "progressive estate taxation" result of Farhi and Werning (2010)—i.e., savings subsidies that decline with income—under the special assumptions that (i) $s'_{pref} \equiv 0$ and (ii) $\beta(z) \equiv \beta \in \mathbb{R}$, $\nu(z) \equiv \nu \in \mathbb{R}$. More generally, Proposition 5 provides a simple formula for balancing the "corrective motives" studied by Farhi and Werning (2010) with the additional motives to tax s in the presence of correlated preference heterogeneity studied in this paper. If $s'_{pref}(z) > 0$ for all z, $\nu(z)/\beta(z)$ is weakly decreasing with z, and $g(z)$ is decreasing with z, the more general result in Proposition 5 suggestions a progressive tax on s that can feature subsidies at low incomes and taxes at high incomes.

5.4 Multi-Dimensional Tax Range with Heterogeneous Rates of Return

Thus far we have considered tax functions where the tax is always paid in units of the numeraire commodity c. In some applications it is also natural to consider tax systems with *multidimensional range*, which include taxes paid in units of c and also taxes paid in units of s. This is natural, for example, if c and s correspond to period 1 and period 2 consumption, respectively, and taxes must be paid in both periods. The additional richness in the range does not alter the optimal tax implications when the rates of transformation p are homogeneous. However, we shall show that there are new economic effects with heterogeneity in p.

Formally, suppose that the government uses a two-part tax structure, where a person must pay a tax $T_1(z)$ in units of c and a tax $T_2(s,z)$ in units of s. For instance, in a two-period model where s is savings, T_1 would be the earnings tax levied in period 1 and paid in period 1 dollars, T_2 would be the savings tax levied in period 2 and paid in period 2 dollars, and $p = 1/(1+r)$ is a function of the rate of return r. Following Gahvari and Micheletto (2016), we consider heterogeneous $p(z,\theta)$ which are a function of gross earnings and type. For example, wealthier individuals have access to better rates of return on savings or prices of commodities. Alternatively, higher earnings ability may be associated with a better ability to obtain high rates of return or to find better prices. Denoting by $\vartheta(z)$ the type θ of agents who choose earnings z, we slightly abuse notation to define $p(z) := p(z, \vartheta(z))$

Agents maximize $U(c,s,z;\theta)$ subject to the budget constraint $c + p(z,\theta)s \leq z - T_1(z) - p(z,\theta)T_2(s,z)$. The government, as before, maximizes a weighted average of utilities,

$$\int_z \left\{ \alpha(z)\, U\Big(z - T_1(z) - p(z)\big(s(z) + T_2(s(z),z)\big), s(z), z; \vartheta(z)\Big) \right\} dH_z(z),$$

subject to the constraints

$$\int_z T_1(z)dH_z(z) \geq E_1$$
$$\int_z T_2(s(z),z)dH_z(z) \geq E_2,$$

which generate marginal values of public funds λ_1 and λ_2 . In Appendix B.10.3 we define $\hat{g}(z)$, analogously to before, as the impact of a unit of after-tax income on z-earners, normalized by the marginal value of public funds λ_1.

Heterogeneity in p generates two novel considerations in this model. First, for agents for whom $p(z) < \lambda_2/\lambda_1$, it is efficient for the government to decrease T_1 and increase T_2 . This generates an economic efficiency rational to increase marginal tax rates on s when T_2 is restricted to be a function of s only. This efficiency effect is present irrespective of the mechanism for the cross-sectional variation of p with z, and leads to a novel deviation from the Atkinson-Stiglitz theorem explored by Gerritsen et al. (2020) and generalized in our results below.

Second, lump-sum changes in T_2 trigger novel substitution effects. This is because a lump-sum increase dT in T_2 has the same effect on an agent's utility as a $p(z,\theta)dT$ increase in T_1, and thus changes behavior as much as a marginal tax rate increase of $\frac{\partial p}{\partial z}dT$ in T_1 (see Appendix B.10.2). We denote by $\phi(z)$ this novel substitution effect at point z, formally derived in Appendix B.10.3. The impact of a uniform lump sum decrease in T_2 is $\overline{\hat{g}(x)p(x) - \varphi(x)}$, where the "bar" notation is used to denote a population average. Thus, $\lambda_2/\lambda_1 = \overline{\hat{g}(x)p(x) - \varphi(x)}$, as we formally show in Appendix B.10.3.

Proposition 6. *With heterogeneous prices, at each bundle (c,s,z) chosen by a type θ, an optimal smooth two-part tax system $\{T_1(z), T_2(s,z)\}$ satisfies*

$$\frac{\lambda_2/\lambda_1 \frac{\partial T_2}{\partial s}\Big|_z}{1 + p(z)\frac{\partial T_2}{\partial s}\big|_z} = s'_{pref}(z)p(z)\frac{1}{s\zeta^c_{s|z}(z)}\frac{1}{h_z(z)}\int_{x=z}^{\bar{z}}\Big[1-\hat{g}(x)\Big]dH_z(x), \tag{36}$$

An optimal smooth two-part tax system $\{T_1(z), T_2(s)\}$ satisfies

$$\frac{\lambda_2/\lambda_1 \frac{\partial T_2}{\partial s}\Big|_z}{1 + p(z)\frac{\partial T_2}{\partial s}\big|_z} = \frac{1}{s\zeta^c_{s|z}(z)}\frac{1}{h_z(z)}\left\{s'_{pref}(z)p(z)\int_{x=z}^{\bar{z}}\Big[1-\hat{g}(x)\Big]dH_z(x) + s'(z)\Big[\Psi(z)+\Phi(z)\Big]\right\} \tag{37}$$

where

$$\Psi(z) := \Big(1-H_z(z)\Big)\int_{x=\underline{z}}^{z}\hat{g}(x)\Big(p(x)-p(z)\Big)dH_z(x) + H_z(z)\int_{x=z}^{\bar{z}}\hat{g}(x)\Big(p(z)-p(x)\Big)dH_z(x),$$
$$\Phi(z) := \int_{x=z}^{\bar{z}}\Big[\varphi(x) - \overline{\varphi(x)}\Big]dH_z(x).$$

We supplement the results in Proposition 6 with derivations of $\frac{\partial T_1}{\partial z}$ and $\frac{\partial T_2}{\partial z}$ at the optimum

(Appendices B.10.4, B.10.5), as well as pareto efficiency conditions (Appendix B.10.7).

Our first key result in Proposition 6, stated in (36), is that when T_2 is fully flexible, the implications for marginal tax rates on s are largely identical to Theorem 2. The result in Theorem 2 obtains when $\lambda_2 = \lambda_1$ and $p(z) \equiv 1$. The presence of $p(z)$ in the denominator is because an agent's marginal tax rate on s, translated to units of c, is $p(z)\frac{\partial T_2}{\partial s}$, and the presence of λ_2/λ_1 in the numerator is because fiscal externalities of substitution from s must be weighted by the "period 2" marginal value of public funds. Consistent with our earlier results, the implication is that marginal tax rates on s are optimally zero when $s'_{pref} \equiv 0$. As we summarized earlier, $s'_{pref} \equiv 0$ will hold if there is no preference heterogeneity and p does not directly depend on the type θ (but can depend on z).

This first result complements the results of Gerritsen et al. (2020), who suggest that there are additional efficiency reasons for taxing s in period 2 even when p does not directly depend on θ. The reason that Gerritsen et al. (2020) reach a different conclusion is because, in the notation of our framework, Gerritsen et al. (2020) assume that T_2 must depend on s only, but not on z. We consider this case in (37), which provides an empirically-implementable generalization of Gerritsen et al. (2020) that accounts for other sources of heterogeneity that can contribute to $s'_{pref}(z)$.[15] Relative to (36), there are two additional efficiency effects, which are both positive when higher earners have lower prices p (e.g., higher rates of returns on savings) and also exhibit larger absolute substitution effects φ. The first term $\Psi(z)$, corresponds to efficiency effects of increase the marginal tax rate on s accrue because the increased tax burden will fall on agents with an average p that is lower than λ_2/λ_1. The second term, $\Phi(z)$, corresponds to the fact that increasing marginal tax rates on s motivates agents to increase labor supply when $\frac{\partial p}{\partial z} < 0$. We emphasize, however, that these additional efficiency benefits of taxing s are a consequence of a missing instrument problem, because they result when T_2 is constrained to depend on s only.

6 Empirical Application

To illustrate the quantitative implications of preference heterogeneity for the taxation of savings and capital, we first calibrate the relevant sufficient statistics from micro data and empirical studies. We then use the Pareto-efficiency conditions derived in Proposition 2 to compute the efficient SL, SN and LED savings tax rates, taking the earnings tax schedule as given. Our results show that preference heterogeneity calls for (mostly) positive and progressive schedule of savings tax rates, which range from approximately 0% at the bottom of the income distribution up to approximately 20% at the top of the income distribution in our baseline calibration.

Calibration To implement our theoretical formulas in the context of savings taxes in the U.S., we require a calibration of the U.S. economy that can be interpreted through the lens of the joint savings and income tax function $\mathcal{T}(s, z)$ and the three simple tax systems described in Table 1.

[15] In contrast to Gerritsen et al. (2020) , who consider "elasticities" with respect to types θ, our results are stated in terms of empirically-measurable elasticities.

We proceed by calibrating a two-period model economy with a fine grid of incomes, where the first period corresponds to work-life and the second to retirement. In keeping with the data we use for this calibration, we assume these periods are separated by 20 years, with an annual rate of return on savings of 3.8% before taxes. In the model, as is standard in models of commodity taxation, z and $\mathcal{T}(s, z)$ are measured in the same units as consumption, which in a dynamic setting corresponds to "period 1" dollars, while s is measured in "period 2" dollars. In practice, savings taxes are typically imposed after returns, and thus measured in period 2 dollars. We accordingly translate all tax rates into units of period 2 dollars when reporting results. Appendix C.1 discusses details of our calibration and this conversion.

The earnings and savings distribution in the U.S., and the U.S. income tax schedule, are calibrated using the Distributional National Accounts micro-files of Piketty et al. (2018). We use 2019 measures of pretax labor income (*plinc*) and net personal wealth (*hweal*) at the individual level as well as the age category (20 to 44 years old, 45 to 64, and above 65). Discretizing the income distribution into percentiles by age group, our measure of annualized earnings during the working life z at the nth percentile is constructed by averaging earnings at the nth percentile across those aged 20 to 44 and those aged 45 to 64. Our measure of s is the accumulated net-of-tax savings at the time of retirement, at each percentile of the income distribution. We construct this measure by projecting forward to age 65 the savings observed at each percentile in the 45-64 age bucket. We normalize both labor earnings and retirement savings by the number of years worked.

Figure 1 plots our estimate of gross (i.e., before-tax) savings across the income distribution in the U.S. The figure shows that savings upon retirement are approximately zero at low incomes, but increase substantially with income. We convert this to net-of-tax savings using a calibration of savings tax rates across the income distribution in the U.S., derived by computing the weighted average of savings tax rates across incomes using the asset composition reported in Bricker et al. (2019). (Details are described in Appendix C.3.) The convex shape of the savings profile in Figure A2b, which persists after accounting for savings taxes, indicates that the cross-sectional slope $s'(z)$ rises with income.

To calibrate the causal income effect on savings $s'_{inc}(z)$ and our measure of local preference heterogeneity $s'_{pref}(z) = s'(z) - s'_{inc}(z)$, we draw from two sources. The first is the work of Fagereng et al. (2019), which estimates the marginal propensity to consume (MPC) windfall income across the earnings distribution using information on lottery prizes linked with administrative data in Norway. Agents' consumption peaks during the winning year and gradually reverts to their previous value afterwards. Over a 5-year horizon, they estimate that agents consume close to 90% of the prize, which translates into an MPC of 0.9, and a marginal propensity to save of 0.1, which we multiply by the rate of return net of taxes to find the causal effect on net retirement savings (see their Figure 2, aggregate consumption response). The second source is a new, nationally representative survey we fielded using the AmeriSpeak panel in the spring of 2021, in which we asked respondents to report what portion of a hypothetical raise they would save. Results, reported in Appendix C.1.2, suggest a short-run MPC close to that reported in Fagereng et al. (2019), with little variation across

incomes. Figure 2 plots the slope of the cross-sectional profile $s'(z)$ and each of these measures of $s'_{inc}(z)$. This plot indicates a substantial difference $s'(z) - s'_{inc}(z)$, rising across incomes, which pushes toward a progressive schedule of optimal savings taxes.

For the earnings elasticity we take the standard value $\zeta_z^c = 0.33$ from the meta-analysis of Chetty (2012). The value of the savings elasticity $\zeta_{s|z}^c$ is harder to obtain, in part because different savings vehicles may trigger different responses to taxes and because different time horizons can lead to different results. We therefore report results for a range of elasticities spanning $\zeta_{s|z}^c = 0.7$ to $\zeta_{s|z}^c = 3$, approximately corresponding to the range of intertemporal elasticities of substitution considered in Golosov et al. (2013). In Appendix **??**, we discuss two recent studies which can further inform the value of this elasticity: Agersnap and Zidar (2020), which exploits across-state variation capital gains tax rates in the U.S. to estimate an elasticity of capital gains realizations with respect to the tax rate, and Jakobsen et al. (2020), which provides evidence on behavioral responses to wealth taxation in Denmark. In both cases, some translation is necessary to convert from the computed elasticity into the compensated elasticity relevant for our formulas, motivating our decision to report results for a wide range of values.

It is interesting to contrast our sufficient statistics calibration with the structural calibration of Golosov et al. (2013), who measure preference heterogeneity through differences in discount factors across *ability* levels. They infer discount factors from a simple optimization model applied to survey data on individuals' household income paths and net worth, while they use survey respondents' results to the Armed Forces Qualification Test (AFQT) to build a proxy for ability. Because they do not find significant variations in discount factors across their proxy for ability, their calibration implies very little preference heterogeneity.

Results Figure 3 reports the schedule of marginal tax rates for SL, SN and LED tax systems that satisfy the Pareto efficiency formulas in Proposition 2, taking the existing U.S. income tax schedule and income distribution as given. In each case, we translate the tax into a marginal tax rate on gross savings at retirement, measured in period 2 dollars. (Appendix A.5 presents a proof that the simplicity structure of a SL, SN, or LED tax system is preserved under such a translation.) Each panel reports results for a different value of the savings elasticity. For SL tax systems, the linear savings tax rate τ_s is by definition constant across earnings levels. For LED tax systems, the linear savings tax rate $\tau_s(z)$ is earnings dependent and we thus report the linear savings tax rate at each earnings level. For SN tax systems, the nonlinear savings tax schedule $T_s(s)$ depends on the value of savings s, and not on earnings z. We also plot the value of the marginal savings tax rate at the corresponding earnings level, employing our calibrated distribution of savings across earnings $s(z)$.

Because we find that $s'_{pref}(z)$ is positive and increasing in z, we find that savings tax rates are positive and increasing in earnings in the SN and LED tax systems. The magnitudes depend on the value of the savings elasticity parameter. In the case of $\zeta_{s|z}^c = 1$, savings tax rates in SN and LED tax systems range from 0% for annual incomes below $\$50,000$, then steadily increase up to nearly 20% for annual incomes around $\$200,000$, and remain somewhat constant thereafter. These

tax rates are substantially higher than the prevailing savings tax rates in the U.S., which are shown in Figure 1. The efficient linear savings tax rate in a SL tax system is around 6%. Changing the savings elasticity parameter scales the efficient savings tax rates without affecting the overall pattern: preference heterogeneity calls for positive and progressive savings tax rates.

7 Conclusion

This paper presents sufficient statistics formulas characterizing the optimal smooth tax system on earnings and savings (or other dimensions of consumption) in the presence of preference heterogeneity that is correlated with earnings ability. The formulas depend on a key sufficient statistic for preference heterogeneity, $s'_{pref}(z)$, which can be estimated from empirical data on cross sectional earnings and savings, and on observed behavioral responses to tax reforms. We show that such a smooth system can implement the optimal mechanism under weak conditions. Under only slightly stronger conditions, the optimal mechanism can also be implemented by two familiar types of "simple" separable tax systems which combine a nonlinear tax on earnings with a nonlinear tax on savings, or with a linear earnings-dependent savings tax. We derive intuitive sufficient statistics formulas for these separable tax systems, as well as for linear separable savings taxes, and we present sufficient conditions under which these necessary conditions are also sufficient to characterize the optimal simple tax systems. Together, these results provide a practical and general method for quantifying optimal tax systems for saving, inheritances, and other commodities in the presence of correlated preference heterogeneity.

References

Agersnap, O. and Zidar, O. M. (2020). The tax elasticity of capital gains and revenue-maximizing rates. Technical report, National Bureau of Economic Research.

Allcott, H., Lockwood, B. B., and Taubinsky, D. (2019). Regressive sin taxes, with an application to the optimal soda tax. *Quarterly Journal of Economics*, 134(3):1557–1626.

Atkinson, A. B. and Stiglitz, J. E. (1976). The design of tax structure: direct versus indirect taxation. *Journal of Public Economics*, 6(1-2):55–75.

Bach, L., Calvet, L. E., and Sodini, P. (2020). Rich pickings? risk, return, and skill in household wealth. *American Economic Review*, 110(9):2703–2747.

Bernheim, B. D. and Taubinsky, D. (2018). Behavioral public economics. In Bernheim, B. D., DellaVigna, S., and Laibson, D., editors, *The Handbook of Behavioral Economics*, volume 1. Elsevier, New York.

Boadway, R., Marchand, M., and Pestieau, P. (2000). Redistribution with unobservable bequests: A case for taxing capital income. *The Scandinavian Journal of Economics*, 102(2):253–267.

Bourguignon, F. and Spadaro, A. (2012). Tax–benefit revealed social preferences. *The Journal of Economic Inequality*, 10(1):75–108.

Bricker, J., Moore, K. B., and Thompson, J. P. (2019). Trends in household portfolio composition. *Finance and Economics Discussion Series: Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C.*

Chetty, R. (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.

Chetty, R. (2012). Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply. *Econometrica*, 80(3):969–1018.

Christiansen, V. and Tuomala, M. (2008). On taxing capital income with income shifting. *International Tax and Public Finance*, 15:527–545.

Cremer, H., Pestieau, P., and Rochet, J.-C. (2003). Capital income taxation when inherited wealth is not observable. *Capital income taxation when inherited wealth is not observable*, 87:2475–2490.

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

Diamond, P. and Spinnewijn, J. (2011). Capital income taxes with heterogeneous discount rates. *American Economic Journal: Economic Policy*, 3(4):52–76.

Fagereng, A., Holm, M. B., and Natvik, G. J. J. (2019). Mpc heterogeneity and household balance sheets. *Available at SSRN 3399027*.

Farhi, E. and Werning, I. (2010). Progressive estate taxation. *Quarterly Journal of Economics*, 125(2):635–673.

Farhi, E. and Werning, I. (2013). Insurance and taxation over the life cycle. *Review of Economic Studies*, 80:596–635.

Feldstein, M. S. (1999). Tax avoidance and the deadweight loss of the income tax. *Review of Economics and Statistics*, 81(4):674–680.

Gahvari, F. and Micheletto, L. (2016). Capital income taxation and the atkinson–stiglitz theorem. *Economics Letters*, 147:86–89.

Gauthier, S. and Henriet, F. (2018). Commodity taxes and taste heterogeneity. *European Economic Review*, 101:284–296.

Gerritsen, A., Jacobs, B., Rusu, A. V., and Spiritus, K. (2020). Optimal taxation of capital income with heterogeneous rates of return.

Golosov, M., Kocherlakota, N., and Tsyvinski, A. (2003). Optimal indirect and capital taxation. *Review of Economic Studies*, 70(569-587).

Golosov, M., Troshkin, M., and Tsyvinski, A. (2016). Redistribution and social insurance. *American Economic Review*, 2(359-386).

Golosov, M., Troshkin, M., Tsyvinski, A., and Weinzierl, M. (2013). Preference heterogeneity and optimal capital income taxation. *Journal of Public Economics*, 97:160–175.

Golosov, M. and Tsyvinski, A. (2006). New dynamic public finance: a user's guide. *NBER Macro Annual*, 317-379.

Golosov, M. and Tsyvinski, A. (2015). Policy implications of dynamic public finance. *Annual Reviews of Economics*, 7:147–171.

Jacquet, L. and Lehmann, E. (2020). Optimal income taxation with composition effects. *Journal of the European Economic Association*.

Jakobsen, K., Jakobsen, K., Kleven, H., and Zucman, G. (2020). Wealth taxation and wealth accumulation: Theory and evidence from denmark. *The Quarterly Journal of Economics*, 135(1):329–388.

Konishi, H. (1995). A pareto-improving commodity tax reform under a smooth nonlinear income tax. *Journal of Public Economics*, 56(3):413–446.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 112(2):443–478.

Lockwood, B. B. (2020). Optimal income taxation with present bias. *American Economic Journal: Economic Policy*, 12(4):298–327.

Lockwood, B. B. and Weinzierl, M. (2016). Positive and normative judgments implicit in us tax policy, and the costs of unequal growth and recessions. *Journal of Monetary Economics*, 77:30–47.

Milgrom, P. and Segal, I. (2002). Envelope theorems for arbitrary choice sets. *Econometrica*, 70(2):583–601.

Mirrlees, J. A. (1971). An exploration in the theory of optimum income taxation. *The Review of Economic Studies*, 38(2):175–208.

Mirrlees, J. A. (1976). Optimal tax theory: a synthesis. *Journal of Public Economics*, 6:327–358.

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

Saez, E. (2001). Using elasticities to derive optimal income tax rates. *The Review of Economic Studies*, 68(1):205–229.

Saez, E. (2002). The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. *Journal of Public Economics*, 83(2):217–230.

Saez, E. and Stantcheva, S. (2018). A simpler theory of optimal capital taxation. *Journal of Public Economics*, 162:120–142.

Slemrod, J. (1995). Income creation or income shifting? behavioral responses to the tax reform act of 1986. *The American Economic Review*, 85(2):175–180.

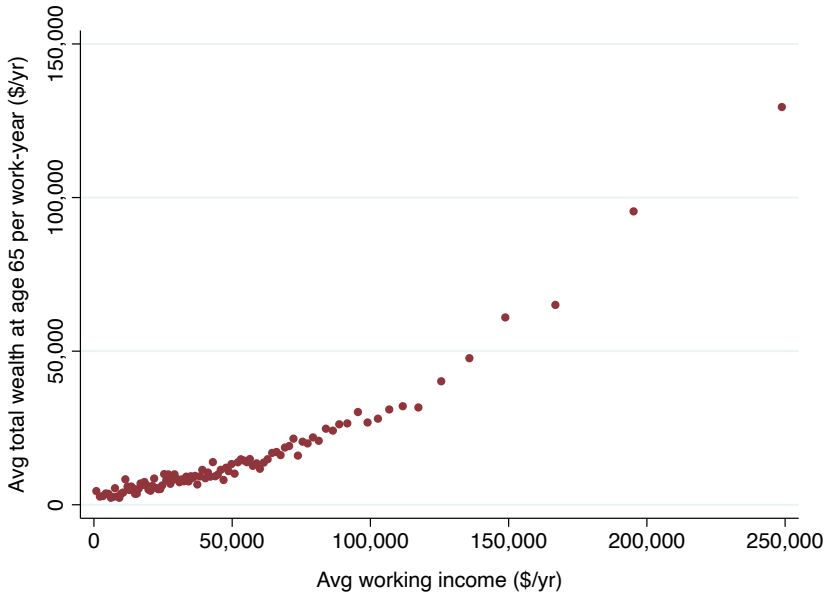
Tables & Figures

Table 2: Tax systems applied to different savings vehicles, by country.

Country	Wealth	Capital Gains	Property	Pensions	Inheritance
Australia	–	Other	SL, SN	SL	–
Austria	–	Other	SL, SN	SN	–
Canada	–	Other	SL	SN	–
Denmark	–	SN	SL, SN	SL, SN	SN
France	–	Other	Other	SL, SN	SN
Germany	–	Other	SL	SN	SN
Ireland	–	SN	SL, SN	SN	SN
Israel	–	Other	Other	SN	–
Italy	SL, SN	SL	SL	SL	SL, SN
Japan	–	SL, SN	SN	SN	SN
Netherlands	SN	SL	SL, SN	SN	SN
New Zealand	–	Other	SN	SL, LED	–
Norway	SN	SL	SL	SN	–
Portugal	–	SL	Other	SN	SL
Singapore	–	Other	SN	SN	–
South Korea	–	SN	SN	SN	SN
Spain	SN	SN	SL, SN	SN	SN
Switzerland	SN	SN	SL, SN	SN	SN
Taiwan	–	SL, SN	SL, SN	SN	SN
United Kingdom	–	Other	SN	SN	SN
United States	–	LED	SL	SN	SN

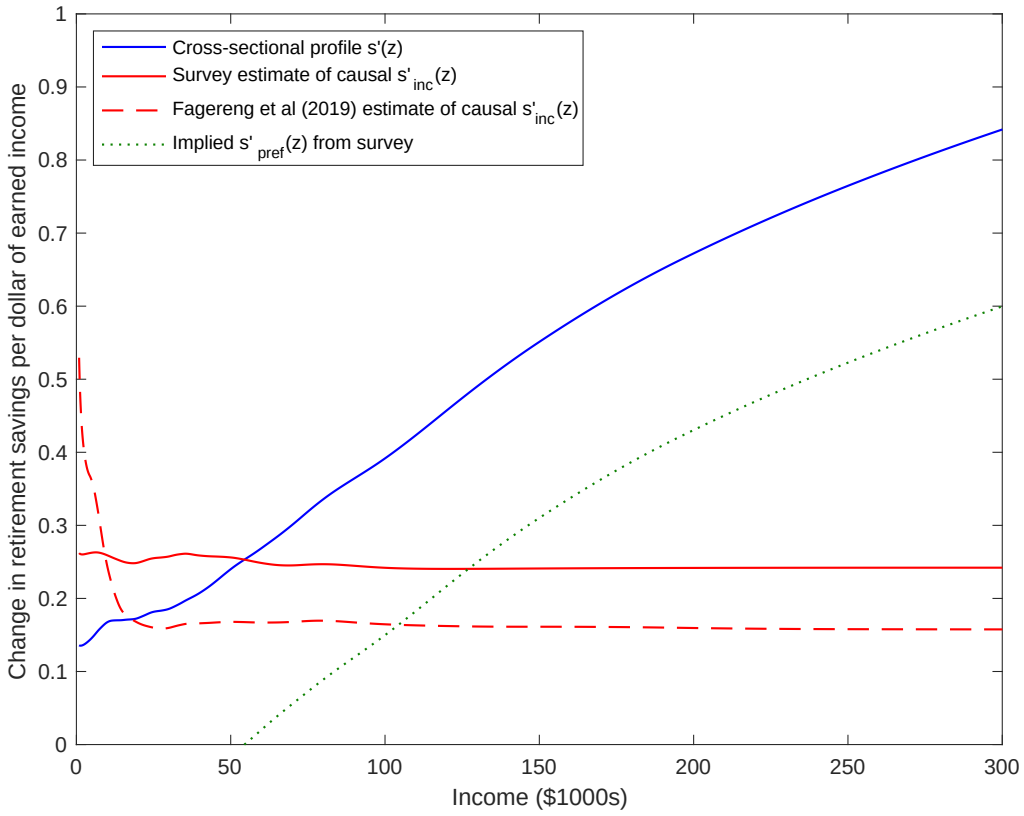
Notes: This table classifies tax systems applied to different savings vehicles across countries in 2020 according to the types in Table 1.

Figure 1: Savings Rates by Income in the United States



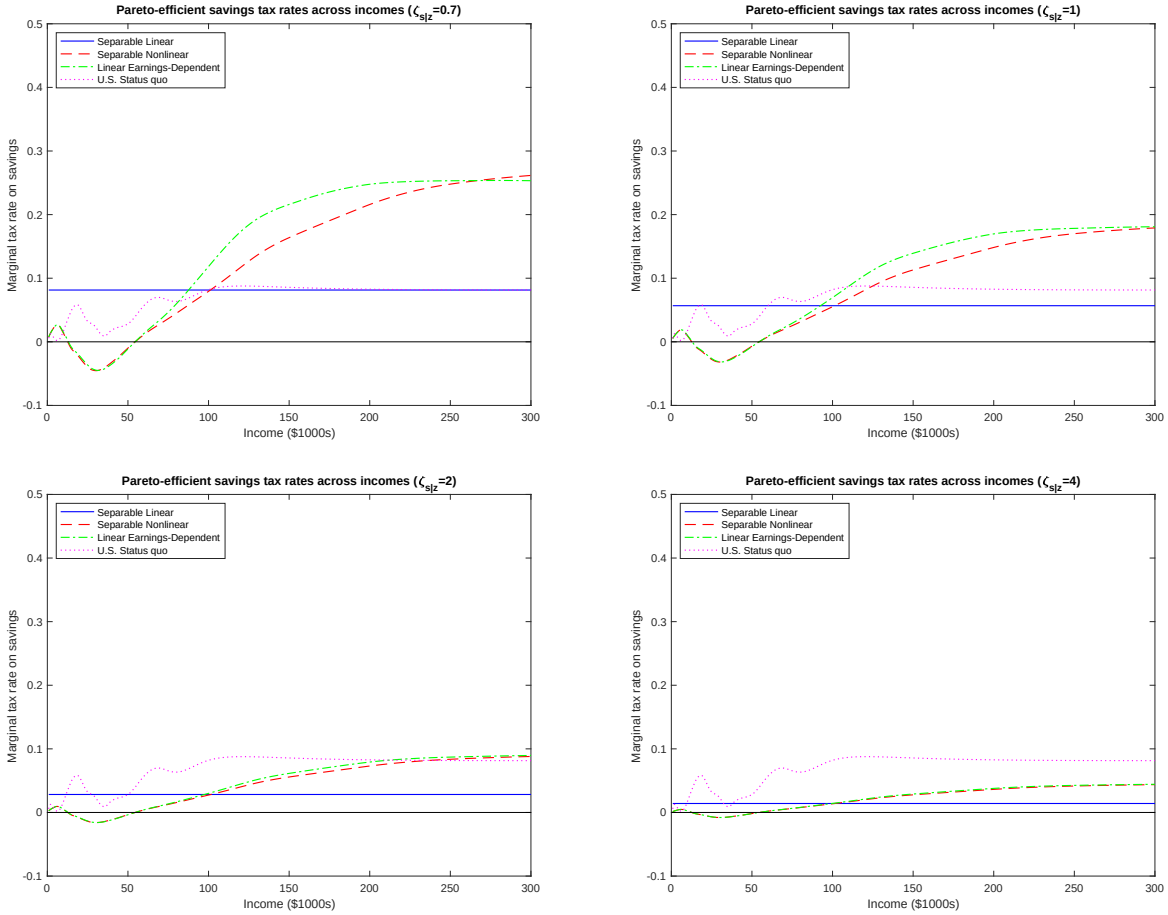
Notes: The earnings and savings distribution in the U.S. is calibrated based on the Distributional National Accounts micro-files of Piketty et al. (2018). We use 2019 measures of pretax income (*plinc*) and net personal wealth (*hweal*) at the individual level, as well as the age category (20 to 44 years old, 45 to 64, and above 65) to impute gross savings at the time of retirement, which we normalize by the number of work years. See Appendix C.1 for further details.

Figure 2: Savings Rates by Income in the United States



Notes: This figure reports the slope of the cross-sectional profile of savings $s'(z)$ (blue), as well as our calibrations of $s'_{inc}(z)$ based on causal income effects, derived from Fagereng et al. (2019) and from a new nationally representative survey. See Section 6 and Appendix C.1 for details.

Figure 3: Savings Tax Rates Implied by Pareto Efficiency Formulas



Notes: This figure presents the marginal savings tax rates values that satisfy the Pareto efficiency formulas in Proposition 2, plotted against the earnings level to which they apply.

Online Appendix

Sufficient Statistics for Nonlinear Tax Systems With Preference Heterogeneity

Antoine Ferey, Benjamin B. Lockwood, and Dmitry Taubinsky

Table of Contents

A Supplementary Theoretical Results **36**
A.1 Implementation Results for Simple Tax Systems 36
A.2 Optimal Tax Rates on s in Simple Tax Systems 38
A.3 Optimal Earnings Tax Rates in Simple Tax Systems 39
A.4 Optimal Earnings tax Rates in Simple Tax Systems with Multidimensional Heterogeneity 40
A.5 Equivalences with Tax Systems Involving Gross Period-2 Savings 40

B Proofs **43**
B.1 Proof of Theorem 1 (Implementation with a Smooth Tax System) 43
B.2 Proof of Proposition A.1 & A.2 (Implementation with Simple Tax Systems) 51
B.3 Proof of Proposition 1 (Measurement of Causal Income Effects) 58
B.4 Proof of Lemma 1 (Earnings Responses to Taxes on s) 59
B.5 Proof of Theorem 2 (Optimal Smooth Tax Systems) 60
B.6 Proof of Propositions 2, A.3, and A.4 (Optimal Simple Tax Systems) 65
B.7 Proof of Proposition 3 (Simple Tax systems and Multidimensional Heterogeneity) . 73
B.8 Proof of Proposition 4 (Many Goods) 80
B.9 Proof of Proposition 5 (Bequest Taxation and Behavioral Agents) 83
B.10 Proof of Proposition 6 (Multi-Dimensional Tax Range with Heterogeneous Prices) 85

C Details on the Empirical Application **91**
C.1 Details of the Calibration of Income, Savings, and Elasticities 92
C.2 Comparison to Golosov et al. (2013) 95
C.3 Estimates of Savings Tax Rates in the United States 96
C.4 Inverse Optimum Approach 97
C.5 Simulations of Optimal Savings Taxes with Multidimensional Heterogeneity 100

D Details of Tax Systems by Country **106**

A Supplementary Theoretical Results

A.1 Implementation Results for Simple Tax Systems

We proceed in three steps to provide sufficient conditions under which some SN and LED tax systems decentralize the optimal incentive-compatible allocation.

First, we define candidate SN and LED tax systems that satisfy type-specific feasibility and agents' first-order conditions at the optimal allocation. Second, in Proposition A.1 we present sufficient conditions under which these SN and LED tax systems also satisfy agents' second-order conditions at the optimal allocation, implying local optimality. Third, in Proposition A.2, we present sufficient conditions under which local optima are ensured to be global optima, implying that the candidate SN and LED systems are indeed implementing the optimal allocation.

Step 1: Defining candidate tax systems. We first define a candidate SN tax system $\mathcal{T}(s,z) = T_s(s)+T_z(z)$, with the nonlinear functions T_s and T_z defined across all savings and earnings bundles of the optimal allocation $\mathcal{A} = (c^*(\theta), s^*(\theta), z^*(\theta))_\theta$ as follows:

$$T_s(s^*(\theta)) := \int_{\vartheta=\underline{\theta}}^{\theta} \left(U_s'(\vartheta)/U_c'(\vartheta) - 1\right) s^{*\prime}(\vartheta) d\vartheta, \tag{38}$$

$$T_z(z^*(\theta)) := z^*(\underline{\theta}) - s^*(\underline{\theta}) - c^*(\underline{\theta}) + \int_{\vartheta=\underline{\theta}}^{\theta} \left(U_z'(\vartheta)/U_c'(\vartheta) + 1\right) s^{*\prime}(\vartheta) d\vartheta \tag{39}$$

where $\underline{\theta}$ denotes the lowest earning type of the compact type space Θ, and the derivatives are evaluated at the bundle assigned in the optimal allocation (e.g., $U_s'(\vartheta) = U_s'(c^*(\vartheta), s^*(\vartheta), z^*(\vartheta); \vartheta)$). Under this tax system, the optimal allocation satisfies by definition each type's first-order conditions for individual optimization given in Equations (9) and (10). By Lemma B.1, this tax system thus satisfies type-specific feasibility.

We similarly define a candidate LED tax system $\mathcal{T}(s,z) = \tau_s(z)\cdot s + T_z(z)$ as follows:

$$\tau_s(z^*(\theta)) := U_s'(\theta)/U_c'(\theta) - 1, \tag{40}$$

$$T_z(z^*(\theta)) := z^*(\underline{\theta}) - s^*(\underline{\theta}) - c^*(\underline{\theta}) + \int_{\vartheta=\underline{\theta}}^{\theta} \left(U_z'(\vartheta)/U_c'(\vartheta) + 1\right) s^{*\prime}(\vartheta) d\vartheta - s^*(z)\cdot\left(\tau_s(z) - \tau_s(z^*(\underline{\theta}))\right). \tag{41}$$

This tax system also satisfies local first-order conditions for individual optimization and type-specific feasibility.

Step 2: Local maxima. We can now derive sufficient conditions under which the above candidate SN and LED tax systems satisfy the second-order conditions for individual optimization, implying that under these conditions assigned bundles are local optima . These conditions can be simply stated in terms of the marginal rates of substitution between consumption and, respectively,

savings $\mathcal{S}(c, s, z; \theta)$ and earnings $\mathcal{Z}(c, s, z; \theta)$. These marginal rates of substitutions are smooth functions of c, s, z, and θ by the smoothness of the allocation and the utility function, and sufficient conditions for local second-order conditions are given by the following proposition.

Proposition A.1. *Suppose that an allocation satisfies the conditions in Theorem 1. Under the SN tax system defined by Equations* (38) *and* (39)*, each agent's assigned choice of savings and earnings is a local optimum if the following conditions hold at each point in the allocation:*

$$\mathcal{S}'_c \geq 0,\ \mathcal{S}'_z \geq 0,\ \mathcal{S}'_\theta \geq 0 \tag{42}$$

and

$$\mathcal{Z}'_c \leq 0,\ \mathcal{Z}'_s \geq 0,\ \mathcal{Z}'_\theta \geq 0. \tag{43}$$

Under the LED tax system defined by Equations (40) *and* (41)*, each agent's assigned choice of savings and earnings is a local optimum if the utility function is additively separable in consumption, savings, and earnings ($U''_{cs} = 0$, $U''_{cz} = 0$, and $U''_{sz} = 0$), and additionally the following conditions hold at each point in the allocation:*

$$\mathcal{S}'_\theta \geq 0,\ \mathcal{S}'_\theta \leq \frac{z^{*\prime}(\theta)}{s^{*\prime}(\theta)}\mathcal{Z}'_\theta,\ \mathcal{S}'_\theta \leq s^{*\prime}(\theta)\left(\mathcal{S}\cdot\mathcal{S}'_c - \mathcal{S}'_s\right). \tag{44}$$

The sufficiency conditions (42) and (43) are quite weak; they are satisfied under many common utility functions used in calibrations of savings and income taxation models, including the simple example function in Equation (1). Conditions $\mathcal{S}'_\theta \geq 0$ and $\mathcal{Z}'_\theta \geq 0$ are single crossing conditions for savings and earnings, while other conditions intuitively relate to the concavity of preferences.

The sufficiency conditions for LED systems are more restrictive. Beyond the single crossing conditions $\mathcal{S}'_\theta \geq 0$ and $\mathcal{Z}'_\theta \geq 0$, they place a constraint on the extent of local preference heterogeneity for savings, as compared with preference heterogeneity in earnings. In words, the preference for savings must not increase "too quickly" across types, or else the second-order condition for earnings may be violated. The intuition for this result can be seen from the definition of the potentially optimal LED system. If the marginal rate of substitution for saving, $\mathcal{S}$, increases very quickly with income at some point in the allocation, then the savings tax rate $\tau_s(z)$ must rise very quickly with z at that point, by Equation (40). Since the savings tax rate $\tau_s(z)$ applies to total savings (including inframarginal savings), this increase in $\tau_s(z)$ must be offset by a sharp decrease in $T_z(z)$ at the same point in the distribution, by Equation (41). Yet a sufficiently steep decrease in $T_z(z)$ will cause the second-order condition for earnings choice—holding fixed savings choice—to be violated.

Step 3: Global maxima. Having presented conditions under which the bundle $(c^*(\theta), s^*(\theta), z^*(\theta))$ assigned to type θ is a local optimum under the candidate SN and LED tax systems, we now present a set of regularity conditions ensuring that these local optima are also global optima.

Proposition A.2. *Assume single crossing conditions for earnings and savings ($\mathcal{Z}'_\theta \geq 0$ and $\mathcal{S}'_\theta \geq 0$), that preferences are weakly separable ($U''_{cz} = 0$ and $U''_{sz} = 0$), and that commodities c and s are*

weak complements ($U''_{cs} \geq 0$). If $\mathcal{A} = \{(c^*(\theta), s^*(\theta), z^*(\theta))\}_\theta$ constitutes a set of local optima for types θ under a smooth tax system $\mathcal{T}$, local optima correspond to global optima when

1. *$\mathcal{T}$ is a SN system, and we have that for all $s > s^*(\theta)$ and θ, $\frac{-U''_{ss}(c(s,\theta),s,z^*(\theta);\theta)}{U'_s(c(s,\theta),s,z^*(\theta);\theta)} > \frac{-T''_{ss}(s)}{1+T'_s(s)}$.*
2. *$\mathcal{T}$ is a LED system, and we have that*
   (a) *for all $s < s^*(\theta)$ and θ, $\frac{-U''_{cc}(c(s,\theta),s,z^*(\theta);\theta)}{U'_c(c(s,\theta),s,z^*(\theta);\theta)} > \frac{1}{1+\tau_s(z^*(\theta))} \frac{\tau'_s(z^*(\theta))}{1-\tau'_s(z^*(\theta))s-T'_z(z^*(\theta))}$,*
   (b) *for all $s > s^*(\theta)$ and θ, $\frac{-U''_{ss}(c(s,\theta),s,z^*(\theta);\theta)}{U'_s(c(s,\theta),s,z^*(\theta);\theta)} > \frac{\tau'_s(z^*(\theta))}{1-\tau'_s(z^*(\theta))s-T'_z(z^*(\theta))}$.*

where $c(s,\theta) := z^*(\theta) - s - \mathcal{T}(s, z^*(\theta))$

In essence, global optimality is ensured under the following assumptions. First, higher types θ derive higher gains from working and allocating those gains to consumption or savings — generalized single-crossing conditions. Second, additive separability of consumption and savings from labor. Third, the utility function U is sufficiently concave in consumption and savings.

For the case of SN tax systems, condition 1 imposes a particular concavity requirement with respect to savings. For the case of LED tax systems, condition 2 imposes particular concavity requirements with respect to both consumption and savings. Notably, these concavity conditions need only be checked when earnings are fixed at each type's assigned earnings level $z^*(\theta)$.

We can naturally apply this result to the candidate SN tax system defined in Equations (38) and (39), and to the candidate LED tax system defined in Equations (40) and (41). Because these candidate tax systems are defined in terms of agents' preferences and optimal allocations, we can then express conditions 1 and 2 fully in terms of agents' preferences and optimal allocations.

A.2 Optimal Tax Rates on s in Simple Tax Systems

We present optimal savings tax formulas for simple tax systems, which characterize the optimal savings tax schedule for *any* given earnings tax schedule—including a potentially suboptimal one. These formulas are derived assuming unidimensional heterogeneity and are similar to those presented in Proposition 3, where heterogeneity is allowed to be multidimensional.

Proposition A.3. *Consider a given (and potentially suboptimal) earnings tax schedule $T_z(z)$, and assume that under each simple tax system agents first-order and second-order conditions strictly hold, and that savings strictly increase with earnings in the cross-section. At each bundle (c, s, z) chosen by a type θ, SL, SN, and LED systems satisfy the following optimality conditions on the*

savings tax rates:

$$SL: \frac{\tau_s}{1+\tau_s}\int_{x=\underline{z}}^{\overline{z}} s(x)\zeta^c_{s|z}(x)dH_z(x) = \int_{x=\underline{z}}^{\overline{z}}\left\{s\left(x\right)\left(1-\hat{g}(x)\right) - \frac{T'_z(x)+s'_{inc}(x)\,\tau_s}{1-T'_z\left(x\right)}x\zeta^c_z(x)\,s'_{inc}(x)\right\}dH_z(x) \tag{45}$$

$$SN: \frac{T'_s(s)}{1+T'_s(s)}s\zeta^c_{s|z}(z)h_z(z) = s'\left(z\right)\int_{x\geq z}\left(1-\hat{g}(x)\right)dH_z(x) - \frac{T'_z(z)+s'_{inc}(z)T'_s(s(z))}{1-T'_z(z)}z\zeta^c_z(z)s'_{inc}(z)h_z(z) \tag{46}$$

$$\begin{aligned} LED: &\frac{T'_z\left(z\right)+\tau'_s\left(z\right)s+s'_{inc}(z)\tau_s(z)}{1-T'_z\left(z\right)-\tau'_s\left(z\right)s}z\zeta^c_z(z)\,s\,h_z(z) + \int_{z^*}^{\bar{z}}\frac{\tau_s(z)}{1+\tau_s(z)}s(z)\zeta^c_{s|z}(z)\,dH_z(z) \\ &= \int_{z^*}^{\bar{z}}\left\{\left(1-\hat{g}(z)\right)s\left(z\right) - \frac{T'_z\left(z\right)+\tau'_s\left(z\right)s\left(z\right)+s'_{inc}(z)\tau_s(z)}{1-T'_z\left(z\right)-\tau'_s\left(z\right)s\left(z\right)}z\zeta^c_z(z)s'_{inc}(z)\right\}dH_z(z) \end{aligned} \tag{47}$$

These optimal savings tax formulas are all different, reflecting differences between the savings tax instruments that we consider. Yet, they share common elements. First, the preference heterogeneity term $s'_{pref}\left(z\right)$ no longer appears in the formulas. The intuition is that outside of the full optimum, it may still be desirable to tax savings in the absence of preference heterogeneity, implying that optimality may clash with Pareto-efficiency when the earnings tax is suboptimal. Second, $s'_{inc}\left(z\right)$ is a key sufficient statistic for optimal savings tax schedules. Indeed, by Lemma 1, a larger $s'_{inc}\left(z\right)$ means that savings tax reforms impose higher distortions on earnings and thus generally calls for lower savings tax rate.

A.3 Optimal Earnings Tax Rates in Simple Tax Systems

We now present optimal earnings tax formulas for simple tax systems.

Proposition A.4. *Consider given (and potentially suboptimal) SL, SN and LED savings tax schedules, and assume that under each simple tax system agents first-order and second-order conditions strictly hold, and that savings strictly increase with earnings in the cross-section. At each bundle (c,s,z) chosen by a type θ, SL, SN, and LED systems satisfy the following optimality conditions on the earnings tax rates:*

$$SL: \frac{T'_z\left(z\right)}{1-T'_z\left(z\right)} = \frac{1}{z\zeta^c_z(z)}\frac{1}{h_z(z)}\int_{x\geq z}\left(1-\hat{g}(x)\right)dH_z(x) - s'_{inc}(z)\frac{\tau_s}{1-T'_z\left(z\right)} \tag{48}$$

$$SN: \frac{T'_z\left(z\right)}{1-T'_z\left(z\right)} = \frac{1}{z\zeta^c_z(z)}\frac{1}{h_z(z)}\int_{x\geq z}\left(1-\hat{g}(x)\right)dH_z(x) - s'_{inc}(z)\frac{T'_s\left(s\right)}{1-T'_z\left(z\right)} \tag{49}$$

$$LED: \frac{T'_z\left(z\right)+\tau'_s\left(z\right)s}{1-T'_z\left(z\right)-\tau'_s\left(z\right)s} = \frac{1}{z\zeta^c_z(z)}\frac{1}{h_z(z)}\int_{x\geq z}\left(1-\hat{g}(x)\right)dH_z(x) - s'_{inc}(z)\frac{\tau_s\left(z\right)}{1-T'_z\left(z\right)-\tau'_s\left(z\right)s} \tag{50}$$

These conditions pinning down the optimal schedule of marginal earnings tax rates are a direct application of Equation (15) presented in Theorem 2 for smooth tax systems. While formulas for SL and SN tax systems look almost identical to the general condition, the formula for LED tax system

looks a bit different. This difference only reflects the fact that for a LED tax system the marginal earnings tax rate is given by $\mathcal{T}'_z(s,z) = T'_z(z) + \tau'_s(z)\,s(z)$, accounting for the earnings-dependent nature of savings taxes.

A.4 Optimal Earnings tax Rates in Simple Tax Systems with Multidimensional Heterogeneity

Proposition A.5. *Consider given (and potentially suboptimal) SL, SN, and LED savings tax schedule, and assume that under each simple tax system agents first-order and second-order conditions strictly hold. Then, at each bundle (c^0, s^0, z^0) chosen by a type θ^0, marginal tax rates on z in SL/SN/LED systems must satisfy the following optimality conditions:*

$$SL: \frac{T'_z(z^0)}{1-T'_z(z^0)}\mathbb{E}\left[\zeta^c_z(s,z)\Big|z^0\right] = \frac{1}{z^0 h_z(z^0)}\int_{z\geq z^0}\left\{\mathbb{E}\left[1-\hat{g}(s,z)\,|z\right]\right\}dH_z(z) - \mathbb{E}\left[s'_{inc}(s,z)\frac{\tau_s}{1-T'_z(z)}\zeta^c_z(s,z)\Big|z^0\right] \tag{51}$$

$$SN: \frac{T'_z(z^0)}{1-T'_z(z^0)}\mathbb{E}\left[\zeta^c_z(s,z)\Big|z^0\right] = \frac{1}{z^0 h_z(z^0)}\int_{z\geq z^0}\left\{\mathbb{E}\left[1-\hat{g}(s,z)\,|z\right]\right\}dH_z(z) - \mathbb{E}\left[s'_{inc}(s,z)\frac{T'_s(s)}{1-T'_z(z)}\zeta^c_z(s,z)\Big|z^0\right] \tag{52}$$

$$LED: \mathbb{E}\left[\frac{T'_z(z)+\tau'_s(z)\,s}{1-T'_z(z)-\tau'_s(z)\,s}\zeta^c_z(s,z)\Big|z^0\right] = \frac{1}{z^0 h_z(z^0)}\int_{z\geq z^0}\left\{\mathbb{E}\left[1-\hat{g}(s,z)\,|z\right]\right\}dH_z(z) - \mathbb{E}\left[s'_{inc}(s,z)\frac{\tau_s(z)}{1-T'_z(z)-\tau'_s(z)\,s}\zeta^c_z(s,z)\Big|z^0\right] \tag{53}$$

These conditions are similar to those presented above for optimal marginal earnings tax rates under unidimensional heterogeneity (Proposition A.4). Indeed, Lemma 1 still applies such that the previous derivations carry over when adding an expectation with respect to savings. Proofs are thus omitted.

A.5 Equivalences with Tax Systems Involving Gross Period-2 Savings

Suppose that there are two periods, and set $1+r = 1/p$. In period 1 the individual earns z and consumes c, and pays income taxes $T_1(z)$. In period 2 the individual realizes *gross pre-tax savings* $s_g = (z-c-T_1(z))(1+r)$, and pays income taxes $T_2(s_g, z)$. The realized savings s are given by $s_g - T_2(s_g, z)$. The total taxes paid in "period 1 dollars" is given by $T_1(z) + T_2(s_g,z)/(1+r)$. The individual maximizes $U(c,s,z)$ subject to the constraint

$$\begin{aligned} & s \leq (z-c-T_1(z))(1+r) - T_2(s_g,z) \\ \Leftrightarrow\; & c + \frac{s}{1+r} \leq z - T_1(z) - \frac{T_2((z-c-T_1(z))(1+r),z)}{1+r} \end{aligned}$$

In our baseline formulation with period-1 tax function $\mathcal{T}(s,z)$, individuals choose s and z to maximize $U(z-s-\mathcal{T}(s,z),s,z;\theta)$. To convert from the formulation with period-2 taxes to our baseline formulation, define a function $s_g^*(s,z)$ implicitly to satisfy the equation

$$s_g^* - T_2(s_g^*, z) = s$$

Note that s_g^* is generally uniquely defined if we have a system with monotonic realized savings s. Then the equivalence in tax schedules is given by

$$\mathcal{T}_s'(s,z) = \frac{1}{1+r}\frac{\partial}{\partial s_g}T_2(s_g,z)|_{s_g=s_g^*}\frac{\partial}{\partial s}s_g^* \tag{54}$$

and $\mathcal{T}_z' = T_z'$. Equation (54) simply computes how a marginal change in s changes the tax burden in terms of period-1 units of money, and the division by $1+r$ is to convert to period-1 units. Now differentiating the definition of s_g^* gives

$$\frac{\partial}{\partial s}s_g^* - \frac{\partial}{\partial s_g}T_2(s_g,z)\frac{\partial}{\partial s}s_g^* = 1$$

and thus

$$\frac{\partial}{\partial s}s_g^* = \frac{1}{1-\frac{\partial}{\partial s_g}T_2(s_g,z)}$$

from which it follows that

$$\mathcal{T}_s'(s,z) = \frac{1}{1+r}\frac{\frac{\partial}{\partial s_g}T_2(s_g,z)|_{s_g=s_g^*}}{1-\frac{\partial}{\partial s_g}T_2(s_g,z)|_{s_g=s_g^*}} \tag{55}$$

We can also start with a schedule $\mathcal{T}$ and converts it to the dynamic tax schedule. Now if s is the realized savings, we can define gross savings in period 2 as $s_g = s + \mathcal{T}(z,s)(1+r) - \mathcal{T}(z,0)$, and we define the function $s^*(s_g,z)$ to satisfy

$$s_g = s^* + (1+r)\left(\mathcal{T}(s^*,z) - \mathcal{T}(0,z)\right)$$

Then

$$\begin{aligned}\frac{\partial}{\partial s_g}T_2(s_g,z) &= (1+r)\mathcal{T}_s'(s^*,z)\frac{\partial}{\partial s_g}s^* \\ &= \frac{(1+r)\mathcal{T}_s'(s^*,z)}{1+(1+r)\mathcal{T}_s'(s^*,z)}\end{aligned} \tag{56}$$

A.5.1 Separable tax systems (SN).

Now if T_2 is a function of s_g only (a separable tax system), then s_g will be a function of s only, and thus $\mathcal{T}_s'$ will only depend on s. Conversely, note that if $\mathcal{T}$ is a separable system, so that $\mathcal{T}_s'$ does not

depend on z, then (56) implies that $\frac{\partial}{\partial s_g}T_2(s_g,z)$ does not depend on z either. Thus, separability is a property preserved under these transformations.

Now if we start with a separable $\mathcal{T}$, then T_2 is given by

$$T_2'(s_g) = (1+r)\frac{\frac{\partial}{\partial s}\mathcal{T}_s'(s))|_{s=s^*}}{1+\frac{\partial}{\partial s}\mathcal{T}_s'(s))|_{s=s^*}}$$

where s^* is the value that solves $s_g = s^* + \mathcal{T}(s^*)$.

A.5.2 Linear tax systems (LED and SL).

If $T_2 = s_g \cdot \tau(z)$, a linear earnings-dependent system, then $s = s_g(1-\tau(z))$ and $s_g = s/(1-\tau(z))$. Moreover, $\frac{\partial}{\partial s}s_g = 1/(1-\tau(z))$, and thus we have that

$$\mathcal{T}_s' = \frac{1}{1+r}\frac{\tau(z)}{1-\tau(z)}$$

which again implies that we have a linear earnings dependent system with rate $\tilde{\tau}(z) = \frac{1}{1+r}\frac{\tau(z)}{1-\tau(z)}$.

Conversely, if we start with a LED system $\mathcal{T}$ with $\mathcal{T}_s' = \tau(z)$, then

$$\frac{\partial}{\partial s_g}T_2(s_g,z) = (1+r)\frac{\tau(z)}{1+\tau(z)}$$

When the tax rates τ are not functions of z, the calculations above show that the conversions preserve not just linearity, but also independence of z.

A.5.3 Analog to Proposition (2)

It is a simple exercise to verify that our proof of Proposition (2) goes through verbatim, provided that all elasticity concepts concern s_g rather than s. Specifically, we define the elasticities as follows:

$$\begin{aligned}
\zeta^c_{s|z}\left(z(\theta)\right) &:= -\frac{1+\mathcal{T}_s'\left(s\left(z;\theta\right),z\right)}{s\left(z;\theta\right)}\frac{\partial s_g\left(z;\theta\right)}{\partial \mathcal{T}_s'\left(s\left(z;\theta\right),z\right)}\Big|_{z=z(\theta)} \\
\eta_{s|z}\left(z(\theta)\right) &:= -\left(1+\mathcal{T}_s'\left(s\left(z;\theta\right),z\right)\right)\frac{\partial s_g\left(z;\theta\right)}{\partial \mathcal{T}\left(s\left(z;\theta\right),z\right)}\Big|_{z=z(\theta)} \\
s'_{pref}(z(\theta)) &:= \Big(\underbrace{\frac{ds_g\left(z;\theta\right)}{dz}}_{s'(z)} - \underbrace{\frac{\partial s_g\left(z;\theta\right)}{\partial z}}_{s'_{inc}(z)}\Big)\Big|_{z=z(\theta)}
\end{aligned}$$

The efficiency conditions for a tax system $T_1(z) + T_2(s_g,z)$ are as follows:

Proposition A.6. *Consider simple tax systems where agents' optima are unique and agents' first-order and second-order conditions strictly hold.*

A SL system $T_2(s_g, z) \equiv \tau_s$ that is not pareto dominated by any other SL system satisfies

$$\frac{\tau_s}{1+r+\tau_s} = \frac{1}{\int_z \zeta^c_{s|z}(z) s_g(z)\, dH_z(z)} \int_z s'_{pref}(z)\, \zeta^c_z(z)\, z\, \frac{T_1'(z) + s'_{inc}(z)\tau_s/(1+r)}{1-T_1'(z)}\, dH_z(z). \quad (57)$$

A SN tax system where $T_2(s_g, z)$ does not depend z that is not pareto dominated by any other SN system satisfies

$$\frac{T_2'(s(z))}{1+r+T_2'(s(z))} = s'_{pref}(z)\, \frac{\zeta^c_z(z) z}{\zeta^c_{s|z}(z) s(z)}\, \frac{T_1'(z) + s'_{inc}(z) T_2'(s(z))/(1+r)}{1-T_1'(z)}. \quad (58)$$

A LED tax system where $T_2(s_g, z) = \tau_s(z) s_g$ that is not pareto dominated by any other LED system satisfies

$$\frac{\tau_s(z)}{1+r+\tau_s(z)} = s'_{pref}(z)\, \frac{\zeta^c_z(z) z}{\zeta^c_{s|z}(z) s(z)}\, \frac{T_1'(z) + (\tau_s'(z)\, s(z) + s'_{inc}(z)\tau_s(z))/(1+r)}{1-T_1'(z) - \tau_s'(z)\, s(z)/(1+r)}. \quad (59)$$

The intuition is simple. First, note that the dynamic system is clearly equivalent to a tax system where $\tilde{T}(s_g, z) = T_1(z) + T_2(s_g, z)/(1+r)$ is paid in period 1. Now consider perturbations of T. Expressions for the impact of a perturbation to the earnings marginal tax rate are clearly identical when written in terms of the respective elasticities. Next, consider a perturbation of the savings tax rates. The expressions for the mechanical effects and distortion to savings are again clearly identical when written in terms of the respective elasticities. Finally, the impact of a perturbation of the savings tax on earnings follows the same logic: a $d\tau_s$ change in the savings tax rate affects earnings as much as a $d\tau_z s'_{inc}$ change in the earnings tax rate, where s'_{inc} is defined with respect to the argument of the tax function—consumption s in our baseline model, or gross savings s_g in the alternative formulation. Thus, expressions for $\frac{d\tilde{T}}{ds_g}$ are identical for the optimal marginal savings tax rates of $\mathcal{T}$ in Proposition (2). To obtain the expressions in Proposition (A.6), we substitute $\frac{dT_2}{ds} = \frac{1}{1+r}\frac{d\tilde{T}}{ds_g}$.

B Proofs

B.1 Proof of Theorem 1 (Implementation with a Smooth Tax System)

In the appendix, we adopt the notation that agent's choices in the optimal mechanism are labeled with a "star"; i.e., $(c^*(\theta), s^*(\theta), z^*(\theta))$. We construct a smooth tax system that implements the optimal incentive-compatible allocation $(c^*(\theta), s^*(\theta), z^*(\theta))$ by introducing penalties for deviations away from these allocations. This proof relies on Lemma B.1 and Lemma B.2, which we derive at the end of this subsection.

By Assumption 2, functions $s^*(\theta)$ and $z^*(\theta)$ are smooth and strictly increasing in θ over the compact set of types $\left[\underline{\theta}, \overline{\theta}\right]$ such that there exists a mapping $s^*(z)$ that denotes the savings level associated with earnings level z at the optimal incentive-compatible allocation. Let $\bar{z} := z^*\left(\overline{\theta}\right)$ and

$\underline{z} := z^*(\underline{\theta})$ denote the maximal and minimal, respectively, earnings levels in the allocation. Let $\overline{s} := s^*(\overline{z})$ and $\underline{s} := s^*(\underline{z})$ denote the corresponding maximal and minimal savings levels.

Step 1: Defining the smooth tax system. We start from a separable and smooth tax system $T_s(s) + T_z(z)$ that satisfies type-specific feasibility and agents' first-order conditions at the optimal incentive-compatible allocation. We then add quadratic penalty terms parametrized by k for deviations from this allocation. For a given deviation, this allows to make the penalty arbitrarily large and enables us to make agents problem locally concave around the optimal incentive-compatible allocation. The other terms that we add are there to guarantee the smoothness of the penalized tax system $\mathcal{T}(s, z; k)$ at the boundaries of the set of optimal allocations.

Formally, $\mathcal{T}_k = \mathcal{T}(s, z; k)$ is defined by:

1. $T_s(\underline{s}) = 0$ and $T_z(\underline{z}) = z^*(\underline{\theta}) - c^*(\underline{\theta}) - s^*(\underline{\theta})$
2. $\forall z \in [\underline{z}; \overline{z}], T'_z(z) = \mathcal{Z}(c^*(\theta_z), s^*(\theta_z), z^*(\theta_z); \theta_z) + 1$ with θ_z such that $z = z^*(\theta_z)$
3. $\forall s \in [\underline{s}; \overline{s}], T'_s(s) = \mathcal{S}(c^*(\theta_s), s^*(\theta_s), z^*(\theta_s); \theta_s) - 1$ with θ_s such that $s = s^*(\theta_s)$
4. $\mathcal{T}(s, z; k) =$
$$\begin{cases}
T_s(s) + T_z(z) + k(s - s^*(z))^2 & \text{if } \underline{z} \leq z \leq \overline{z}, \underline{s} \leq s \leq \overline{s} \\
T_s(\underline{s}) + T_z(z) + k(s - s^*(z))^2 + T'_s(\underline{s})(s - \underline{s}) & \text{if } \underline{z} \leq z \leq \overline{z}, s < \underline{s} \\
T_s(\overline{s}) + T_z(z) + k(s - s^*(z))^2 + T'_s(\overline{s})(s - \overline{s}) & \text{if } \underline{z} \leq z \leq \overline{z}, s > \overline{s} \\
T_s(s) + T_z(\underline{z}) + k(s - \underline{s})^2 + k(z - \underline{z})^2 + T'_z(\underline{z})(z - \underline{z}) & \text{if } z < \underline{z}, \underline{s} \leq s \leq \overline{s} \\
T_s(\underline{s}) + T_z(\underline{z}) + k(s - \underline{s})^2 + k(z - \underline{z})^2 + T'_z(\underline{z})(z - \underline{z}) + T'_s(\underline{s})(s - \underline{s}) & \text{if } z < \underline{z}, s < \underline{s} \\
T_s(\overline{s}) + T_z(\underline{z}) + k(s - \underline{s})^2 + k(z - \underline{z})^2 + T'_z(\underline{z})(z - \underline{z}) + T'_s(\overline{s})(s - \overline{s}) & \text{if } z < \underline{z}, s > \overline{s} \\
T_s(s) + T_z(\overline{z}) + k(s - \overline{s})^2 + k(z - \overline{z})^2 + T'_z(\overline{z})(z - \overline{z}) & \text{if } z > \overline{z}, \underline{s} \leq s \leq \overline{s} \\
T_s(\overline{s}) + T_z(\overline{z}) + k(s - \overline{s})^2 + k(z - \overline{z})^2 + T'_z(\overline{z})(z - \overline{z}) + T'_s(\overline{s})(s - \overline{s}) & \text{if } z > \overline{z}, s > \overline{s} \\
T_s(\underline{s}) + T_z(\overline{z}) + k(s - \overline{s})^2 + k(z - \overline{z})^2 + T'_z(\overline{z})(z - \overline{z}) + T'_s(\underline{s})(s - \underline{s}) & \text{if } z > \overline{z}, s < \underline{s}
\end{cases}$$

Assumptions 1 and 2 guarantee that the separable tax system $T_s(s) + T_z(z)$ is smooth i.e. a twice continuously differentiable function. Our construction of the penalized tax system $\mathcal{T}_k = \mathcal{T}(s, z; k)$ guarantees that it inherits this smoothness property.

Step 2: Local maxima for sufficiently large k. For a given agent θ, we show that the bundle $(c^*(\theta), s^*(\theta), z^*(\theta))$ is a local optimum under the tax system $\mathcal{T}_k = \mathcal{T}(s, z; k)$ for sufficiently large k. To do so, we first establish that type-specific feasibility is satisfied together with the first-order conditions of agent θ maximization problem. We then show that for sufficiently large k, second-order conditions are also satisfied implying that the intended bundle is a local maximum.

The previous definition of the tax system implies

$$\mathcal{T}'_z(s^*(\theta), z^*(\theta); k) = T'_z(z^*(\theta)) = \mathcal{Z}(c^*(\theta), s^*(\theta), z^*(\theta); \theta) + 1$$
$$\mathcal{T}'_s(s^*(\theta), z^*(\theta); k) = T'_s(s^*(\theta)) = \mathcal{S}(c^*(\theta), s^*(\theta), z^*(\theta); \theta) - 1$$

meaning type-specific feasibility is satisfied by Lemma B.1 (see below).

Now, defining

$$V(s,z;\theta,k) := U(z-s-\mathcal{T}(s,z;k),s,z;\theta), \tag{60}$$

first-order conditions of agent θ choice of savings s and earnings z are

$$V'_s(s,z;\theta,k) = -(1+\mathcal{T}'_s(s,z;k))U'_c(z-s-\mathcal{T}(s,z;k),s,z;\theta) + U'_s(z-s-\mathcal{T}(s,z;k),s,z;\theta) = 0$$
$$V'_z(s,z;\theta,k) = (1-\mathcal{T}'_z(s,z;k))U'_c(z-s-\mathcal{T}(s,z;k),s,z;\theta) + U'_z(z-s-\mathcal{T}(s,z;k),s,z;\theta) = 0$$

and they are by construction satisfied at $(s^*(\theta),z^*(\theta))$ for each type θ.

Using Lemma B.2 (see below), second-order conditions at $(s^*(\theta),z^*(\theta))$ are

$$V''_{ss} = \frac{U'_z}{s^{*\prime}(z^*)}\mathcal{S}'_c - \frac{U'_c}{s^{*\prime}(z^*)}\mathcal{S}'_z - \frac{U'_c}{s^{*\prime}(\theta)}\mathcal{S}'_\theta + \frac{U'_c}{s^{*\prime}(z^*)}\mathcal{T}''_{sz} \leq 0 \tag{61}$$

$$V''_{zz} = U'_s s^{*\prime}(z^*)\mathcal{Z}'_c - U'_c s^{*\prime}(z^*)\mathcal{Z}'_s - \frac{U'_c}{z^{*\prime}(\theta)}\mathcal{Z}'_\theta + U'_c s^{*\prime}(z^*)\mathcal{T}''_{sz} \leq 0 \tag{62}$$

$$\begin{aligned}\left(V''_{sz}\right)^2 - V''_{ss}V''_{zz} = \frac{U'_c}{s^{*\prime}(\theta)}\Bigg[&\left(U'_z\mathcal{S}'_c - U'_c\mathcal{S}'_z\right)\mathcal{Z}'_\theta + \left(U'_s\mathcal{Z}'_c - U'_c\mathcal{Z}'_s - U'_c\frac{\mathcal{Z}'_\theta}{s^{*\prime}(\theta)}\right)s^{*\prime}(z^*)\mathcal{S}'_\theta \\ &+ \left(\mathcal{Z}'_\theta + s^{*\prime}(z^*)\mathcal{S}'_\theta\right)U'_c\mathcal{T}''_{sz}\Bigg] \leq 0\end{aligned} \tag{63}$$

where we denote $s^{*\prime}(z^*) := \frac{s^{*\prime}(\theta)}{z^{*\prime}(\theta)}$.

Here, U, $\mathcal{S}$, and $\mathcal{Z}$ are smooth functions implying that their derivatives are continuous functions over compact spaces, and are thus bounded. Now, by definition of $\mathcal{T}_k = \mathcal{T}(s,z;k)$ we have $\mathcal{T}''_{sz} = -2ks^{*\prime}(z)$ which is negative for any $k \geq 0$ and increasing in magnitude with k.

Noting $U'_c \geq 0$ and $s^{*\prime}(z) \geq 0$, it implies that V''_{ss} and V''_{zz} must be negative for sufficiently large k thanks to the last term since other terms are bounded and do not depend on k. By the same reasoning, under the extended Spence-Mirrlees single-crossing assumption that $\mathcal{Z}'_\theta + s^{*\prime}(z^*)\mathcal{S}'_\theta \geq 0$, we also have that $(V''_{sz})^2 - V''_{ss}V''_{zz}$ must be negative for sufficiently large k.

This shows that for a given agent θ, there exists a k_0 such that for all $k \geq k_0$ the allocation $(c^*(\theta),s^*(\theta),z^*(\theta))$ is a local optimum to agent θ maximization problem under the smooth penalized tax system $\mathcal{T}_k = \mathcal{T}(s,z;k)$.

Step 3: Global maxima for sufficiently large k. Let $s_{\mathcal{T}_k}(\theta)$ and $z_{\mathcal{T}_k}(\theta)$ denote the level of savings and earnings, respectively, that a type θ chooses given a smooth penalized tax system $\mathcal{T}_k$. To prove implementability of optimal incentive-compatible allocations, we show that there exists a sufficiently large k such that for all θ, $s_{\mathcal{T}_k}(\theta) = s^*(\theta)$ and $z_{\mathcal{T}_k}(\theta) = z^*(\theta)$.

Let's proceed by contradiction, and suppose that it is not the case. Then, there exists an infinite sequence of types θ_k, choosing savings $s_{\mathcal{T}_k}(\theta_k) \neq s^*(\theta_k)$ and earnings $z_{\mathcal{T}_k}(\theta_k) \neq z^*(\theta_k)$ under tax system $\mathcal{T}_k$, and enjoying utility gains from this "deviation" to allocation $(s_{\mathcal{T}_k}(\theta_k), z_{\mathcal{T}_k}(\theta_k))$.

First, the fact that we impose quadratic penalties for earnings choices outside of $[\,\underline{z}\,;\,\overline{z}\,]$ guar-

antees that for any $\varepsilon > 0$, there exists k_1, such that for all $k \geq k_1$ and for all types θ, agents' earnings choices belong to the compact set $[\,\underline{z} - \varepsilon\,;\,\overline{z} + \varepsilon\,]$. Indeed, starting from a given earnings level $z > \overline{z} + \varepsilon$, the utility change associated with an earnings change is $[(1 - \mathcal{T}'_z)\,U'_c + U'_z]\,dz$. By construction, the marginal earnings tax rate in this region is $\mathcal{T}'_z = 2k(z - \bar{z}) + T'_z(\bar{z})$. Noting that $U'_c > 0$, $U'_z < 0$, and that the type space is compact, we can make for all individuals the utility change from a decrease in earnings arbitrarily positive for sufficiently large k. This shows that all individuals choose earnings $z \leq \overline{z} + \varepsilon$ for sufficiently large k. Symetrically, we can show that all individuals choose earnings $z \geq \underline{z} - \varepsilon$ for sufficiently large k.

Second, since earnings shape agents' disposable income, the fact that earnings belong to a compact set for sufficiently large k implies that agents' allocation choices also belong to a compact set. Indeed, for sufficiently large k agents' allocation choices must belong to the set of (c, s, z) such that $c \geq 0$, $s \geq 0$, $z \in [\,\underline{z} - \varepsilon\,;\,\overline{z} + \varepsilon\,]$, and $c + s \leq z - \mathcal{T}\,(s, z; k)$ where the tax function is smooth and finite. These constraints make the space of allocations compact.

As a result, the infinite sequence $(\theta_k, s_{\mathcal{T}_k}(\theta_k), z_{\mathcal{T}_k}(\theta_k))$ belongs to a compact space of allocations and types, it is thus bounded. By the Bolzano–Weierstrass theorem, this means that there exists a convergent subsequence $\left(\theta_j, s_{\mathcal{T}_j}(\theta_j), z_{\mathcal{T}_j}(\theta_j)\right) \to (\hat{\theta}, \hat{s}, \hat{z})$. Since all types θ_j belong to $\left[\,\underline{\theta}\,;\,\overline{\theta}\,\right]$, we know that the limit type $\hat{\theta}$ must belong to $\left[\,\underline{\theta}\,;\,\overline{\theta}\,\right]$. Now, from the previous paragraph, agents' earnings choices have to be arbitrarily close to $[\,\underline{z}\,;\,\overline{z}\,]$ as the penalty grows. This implies that the limit $\hat{z}$ must belong to $[\,\underline{z}\,;\,\overline{z}\,]$.

Next, we establish that the limit $\hat{s}$ must be such that $\hat{s} = s^*\,(\hat{z})$. Fix an earnings level $z \in [\,\underline{z}\,;\,\overline{z}\,]$, starting from a savings level $s \neq s^*\,(z)$ the utility change associated with a savings change is $[-\,(1 + \mathcal{T}'_s)\,U'_c + U'_s]\,ds$. Assuming without loss of generality that s belongs to $[\,\underline{s}\,;\,\overline{s}\,]$, the marginal savings tax rate in this region is $\mathcal{T}'_s = T'_s(s) + 2k(s - s^*(z))$. Noting that $U'_c > 0$, and that U'_s is bounded, we can make the utility gains from a savings change towards $s^*(z)$ arbitrarily large for sufficiently large k. As a result, for any $\varepsilon > 0$, there exists k_2 such that for all $k \geq k_2$, agent $\hat{\theta}$ chooses savings $s \in [\,s^*\,(z) - \varepsilon\,;\,s^*\,(z) + \varepsilon\,]$ for a fixed z.[16] Since agent $\hat{\theta}$ savings choice can be made arbitrarily close to $s^*\,(z)$ for sufficiently large k, we must have at the limit $s = s^*\,(z)$. Now, because earnings z converge to $\hat{z}$ and the function $s^*\,(z)$ is by assumption continuous, we must have at the limit $\hat{s} = s^*\,(\hat{z})$.

We have thus established that the limit $(\hat{\theta}, \hat{s}, \hat{z})$ is such that $\hat{\theta} \in \left[\,\underline{\theta}\,;\,\overline{\theta}\,\right]$, $\hat{z} \in [\,\underline{z}\,;\,\overline{z}\,]$, and $\hat{s} = s^*\,(\hat{z})$. This means that the limit allocation $(\hat{c}, \hat{s}, \hat{z})$ belongs to the set of optimal incentive-compatible allocations. Given our continuity and monotonicity assumptions, this implies that it is the optimal allocation of some type θ and it has to be by definition that of agent $\hat{\theta}$. Hence, $(\hat{c}, \hat{s}, \hat{z}) = (c^*(\hat{\theta}), s^*(\hat{\theta}), z^*(\hat{\theta}))$.

To complete the proof and find a contradiction, fix a value $k^\dagger$ that is large enough such that second-order conditions hold for type $\hat{\theta}$ at allocation $(s^*(\hat{\theta}), z^*(\hat{\theta}))$ under tax system $\mathcal{T}_{k^\dagger}$ – this $k^\dagger$

[16] A way to see this is that the marginal rate of substitution between consumption and savings $\mathcal{S}$ is continuous on a compact space and thus bounded, whereas the marginal tax rate which is parametrized by k can be made arbitrarily large. As a result, agents' first-order condition can never hold for sufficiently large k, while we can similarly exclude corner solutions for sufficiently large k.

exists by step 2. This implies that there exists an open set N containing $(s^*(\hat{\theta}), z^*(\hat{\theta}))$ such that $V(s, z; \hat{\theta}, k^\dagger)$ is strictly concave over $(s, z) \in N$.

Now, consider types θ^j with $j \geq k^\dagger$. Since these agents belong to the previously defined subsequence, they prefer allocation $\left(s_{\mathcal{T}_j}(\theta_j), z_{\mathcal{T}_j}(\theta_j)\right)$ to allocation $(s^*(\theta_j), z^*(\theta_j))$ under tax system $\mathcal{T}_j$. Because penalties are increasingly large and $j \geq k^\dagger$, this implies that agents θ^j also prefer allocation $\left(s_{\mathcal{T}_j}(\theta_j), z_{\mathcal{T}_j}(\theta_j)\right)$ to allocation $(s^*(\theta_j), z^*(\theta_j))$ under tax system $\mathcal{T}_{k^\dagger}$.

Yet, by continuity, the function $V\left(s, z; \theta_j, k^\dagger\right)$ gets arbitrarily close to the function $V(s, z; \hat{\theta}, k^\dagger)$ for sufficiently large j since $\theta_j \to \hat{\theta}$. For the same reason, $(s^*(\theta_j), z^*(\theta_j)) \to (s^*(\hat{\theta}), z^*(\hat{\theta}))$. Moreover, by definition $\left(s_{\mathcal{T}_j}(\theta_j), z_{\mathcal{T}_j}(\theta_j)\right) \to (\hat{s}, \hat{z})$. As a result, for any open set $N' \subsetneq N$ containing $(s^*(\hat{\theta}), z^*(\hat{\theta}))$, there exists a $j^\dagger \geq k^\dagger$ such that $V(s, z; \theta_{j^\dagger}, k^\dagger)$ is strictly concave over $(s, z) \in N'$ and such that both $\left(s^*(\theta_{j^\dagger}), z^*(\theta_{j^\dagger})\right)$ and $\left(s_{\mathcal{T}_{j^\dagger}}(\theta_{j^\dagger}), z_{\mathcal{T}_{j^\dagger}}(\theta_{j^\dagger})\right)$ belong to the set N'.

Since $V(s, z; \theta_{j^\dagger}, k^\dagger)$ is strictly concave over $(s, z) \in N'$ it has a unique optimum on N'. By definition of $\mathcal{T}_{k^\dagger}$, agent $\theta_{j^\dagger}$ first-order conditions are satisfied at $\left(s^*(\theta_{j^\dagger}), z^*(\theta_{j^\dagger})\right)$. Hence, $\left(s^*(\theta_{j^\dagger}), z^*(\theta_{j^\dagger})\right)$ is agent $\theta_{j^\dagger}$ maximum under the tax system $\mathcal{T}_{k^\dagger}$. This contradicts the fact established above that agent $\theta_{j^\dagger}$ prefers $\left(s_{\mathcal{T}_{j^\dagger}}(\theta_{j^\dagger}), z_{\mathcal{T}_{j^\dagger}}(\theta_{j^\dagger})\right)$ to allocation $\left(s^*(\theta_{j^\dagger}), z^*(\theta_{j^\dagger})\right)$ under tax system $\mathcal{T}_{k^\dagger}$, which completes the proof.

Lemma for type-specific feasibility.

Lemma B.1. *A smooth tax system $\mathcal{T}$ satisfies type-specific feasibility over the compact type space $\left[\underline{\theta}; \overline{\theta}\right]$ if it satisfies the following conditions*

1. $\mathcal{T}(s^*(\underline{\theta}), z^*(\underline{\theta})) = z^*(\underline{\theta}) - c^*(\underline{\theta}) - s^*(\underline{\theta})$
2. $\mathcal{T}'_z(s^*(\theta), z^*(\theta)) = \mathcal{Z}(c^*(\theta), s^*(\theta), z^*(\theta); \theta) + 1$
3. $\mathcal{T}'_s(s^*(\theta), z^*(\theta)) = \mathcal{S}(c^*(\theta), s^*(\theta), z^*(\theta); \theta) - 1$

Proof. Consider the type-specific feasible tax system $T^*_\theta(\theta) = z^*(\theta) - s^*(\theta) - c^*(\theta)$, and note that the lemma amounts to showing that $T^*_\theta(\theta) = \mathcal{T}(s^*(\theta), z^*(\theta))$ for all θ. To that end, note that the first-order condition for truthful reporting of θ under the optimal mechanism implies

$$U'_c \cdot (z'(\theta) - s'(\theta) - T^{*\prime}_\theta(\theta)) + U'_s \cdot s'(\theta) + U'_z \cdot z'(\theta) = 0,$$

with derivatives evaluated at the optimal allocation. This can be rearranged as

$$\begin{aligned} T^{*\prime}_\theta(\theta) &= \left(\frac{U'_s}{U'_c} - 1\right) s'(\theta) + \left(\frac{U'_z}{U'_c} + 1\right) z'(\theta) \\ &= \mathcal{T}'_s(s^*(\theta)) s^{*\prime}(\theta) + \mathcal{T}'_z(z^*(\theta)) z^{*\prime}(\theta). \end{aligned}$$

It thus follows that

$$\begin{aligned}\mathcal{T}(s^*(\theta),z^*(\theta))-\mathcal{T}(s^*(\underline{\theta}),z^*(\underline{\theta}))&=\int_{\vartheta=\underline{\theta}}^{\vartheta=\theta}\left(\mathcal{T}_s'(s^*(\vartheta))s^{*\prime}(\vartheta)+\mathcal{T}_z'(z^*(\vartheta))z^{*\prime}(\vartheta)\right)d\vartheta\\&=T_\theta^*(\theta)-T_\theta^*(\underline{\theta})\end{aligned}$$

Since $\mathcal{T}(s^*(\underline{\theta}),z^*(\underline{\theta}))=T_\theta^*(\underline{\theta})$, this implies that $\mathcal{T}(s^*(\theta),z^*(\theta))=T_\theta^*(\theta)$ for all θ. The smooth tax system $\mathcal{T}$ therefore satisfies type-specific feasibility. □

Lemma on second-order conditions.

Lemma B.2. *Consider a smooth tax system $\mathcal{T}$ satisfying the conditions in Lemma B.1 and define*

$$V(s,z;\theta):=U(z-s-\mathcal{T}(s,z),s,z;\theta). \tag{64}$$

When evaluated at allocation $(c^*(\theta),s^*(\theta),z^*(\theta))$, we show that

$$V_{ss}''=\frac{U_z'}{s^{*\prime}(z^*)}\mathcal{S}_c'-\frac{U_c'}{s^{*\prime}(z^*)}\mathcal{S}_z'-\frac{U_c'}{s^{*\prime}(\theta)}\mathcal{S}_\theta'+\frac{U_c'}{s^{*\prime}(z^*)}\mathcal{T}_{sz}'' \tag{65}$$

$$V_{zz}''=U_s's^{*\prime}(z^*)\mathcal{Z}_c'-U_c's^{*\prime}(z^*)\mathcal{Z}_s'-\frac{U_c'}{z^{*\prime}(\theta)}\mathcal{Z}_\theta'+U_c's^{*\prime}(z^*)\mathcal{T}_{sz}'' \tag{66}$$

$$\begin{aligned}\left(V_{sz}''\right)^2-V_{ss}''V_{zz}''=\frac{U_c'}{s^{*\prime}(\theta)}\Bigg[&\left(U_z'\mathcal{S}_c'-U_c'\mathcal{S}_z'\right)\mathcal{Z}_\theta'+\left(U_s'\mathcal{Z}_c'-U_c'\mathcal{Z}_s'-U_c'\frac{\mathcal{Z}_\theta'}{s^{*\prime}(\theta)}\right)s^{*\prime}(z^*)\mathcal{S}_\theta'\\&+\left(\mathcal{Z}_\theta'+s^{*\prime}(z^*)\mathcal{S}_\theta'\right)U_c'\mathcal{T}_{sz}''\Bigg]\end{aligned} \tag{67}$$

where we denote $s^{*\prime}(z^*):=\frac{s^{*\prime}(\theta)}{z^{*\prime}(\theta)}$.

Proof. First-order derivatives are

$$\begin{aligned}V_s'(s,z;\theta)&=-(1+\mathcal{T}_s'(s,z))U_c'(z-s-\mathcal{T}(s,z),s,z;\theta)+U_s'(z-s-\mathcal{T}(s,z),s,z;\theta)\\V_z'(s,z;\theta)&=(1-\mathcal{T}_z'(s,z))U_c'(z-s-\mathcal{T}(s,z),s,z;\theta)+U_z'(z-s-\mathcal{T}(s,z),s,z;\theta)\end{aligned}$$

Second-order derivatives are

$$V_{ss}''(s,z;\theta)=-\mathcal{T}_{ss}''U_c'-(1+\mathcal{T}_s')\left(-(1+\mathcal{T}_s')U_{cc}''+U_{cs}''\right)-(1+\mathcal{T}_s')U_{cs}''+U_{ss}'' \tag{68}$$

$$V_{zz}''(s,z;\theta)=-\mathcal{T}_{zz}''U_c'+(1-\mathcal{T}_z')\left((1-\mathcal{T}_z')U_{cc}''+U_{cz}''\right)+(1-\mathcal{T}_z')U_{cz}''+U_{zz}'' \tag{69}$$

To obtain the first result of the Lemma, we compute $\mathcal{T}_{ss}''$ by differentiating both sides of $\mathcal{T}_s'(s^*(\theta),z^*(\theta))=\mathcal{S}(c^*(\theta),s^*(\theta),z^*(\theta);\theta)-1$ with respect to θ

$$\begin{aligned}\mathcal{T}_{ss}''s^{*\prime}(\theta)+\mathcal{T}_{sz}''z^{*\prime}(\theta)&=\frac{d}{d\theta}\mathcal{S}(c^*(\theta),s^*(\theta),z^*(\theta);\theta)\\&=\mathcal{S}_c'c^{*\prime}(\theta)+\mathcal{S}_s's^{*\prime}(\theta)+\mathcal{S}_z'z^{*\prime}(\theta)+\mathcal{S}_\theta'\end{aligned}$$

plugging in $c^{*\prime}(\theta) = (1-\mathcal{T}_z')\,z^{*\prime}(\theta) - (1+\mathcal{T}_s')\,s^{*\prime}(\theta)$ and denoting $s^{*\prime}(z^*) := s^{*\prime}(\theta)/z^{*\prime}(\theta)$, the previous expression can be rearranged as

$$\mathcal{T}_{ss}'' = \mathcal{S}_c'\frac{1-\mathcal{T}_z'}{s^{*\prime}(z^*)} - \mathcal{S}_c'(1+\mathcal{T}_s') + \mathcal{S}_s' + \frac{\mathcal{S}_z'}{s^{*\prime}(z^*)} + \frac{\mathcal{S}_\theta'}{s^{*\prime}(\theta)} - \frac{\mathcal{T}_{sz}''}{s^{*\prime}(z^*)}. \tag{70}$$

Moreover, we differentiate Equation (2) to express the derivative of $\mathcal{S}$ with respect to c as

$$\begin{aligned}\mathcal{S}_c'(c^*(\theta), s^*(\theta), z^*(\theta); \theta) &= \frac{U_c'U_{sc}'' - U_s'U_{cc}''}{(U_c')^2} \\ &= \frac{1}{U_c'}\left(-\frac{U_s'}{U_c'}U_{cc}'' + U_{sc}''\right) \\ &= \frac{1}{U_c'}\left(-(1+\mathcal{T}_s')U_{cc}'' + U_{sc}''\right)\end{aligned} \tag{71}$$

and the derivative of $\mathcal{S}$ with respect to s as

$$\begin{aligned}\mathcal{S}_s'(c^*(\theta), s^*(\theta), z^*(\theta); \theta) &= \frac{U_c'U_{ss}'' - U_s'U_{cs}''}{(U_c')^2} \\ &= \frac{1}{U_c'}\left(-\frac{U_s'}{U_c'}U_{cs}'' + U_{ss}''\right) \\ &= \frac{1}{U_c'}\left(-(1+\mathcal{T}_s')U_{cs}'' + U_{ss}''\right).\end{aligned} \tag{72}$$

Substituting equations (70), (71) and (72) into (68), we have

$$\begin{aligned}V_{ss}''\left(s^*(\theta), z^*(\theta); \theta\right) &= -U_c' \cdot \left(\mathcal{S}_c'\frac{1-\mathcal{T}_z'}{s^{*\prime}(z)} - \mathcal{S}_c'(1+\mathcal{T}_s') + \mathcal{S}_s' + \frac{\mathcal{S}_z'}{s^{*\prime}(z)} + \frac{\mathcal{S}_\theta'}{s^{*\prime}(\theta)} - \frac{\mathcal{T}_{sz}''}{s^{*\prime}(z)}\right) - (1+\mathcal{T}_s')U_s'\mathcal{S}_c' + U_c'\mathcal{S}_s' \\ &= -U_c' \cdot \left(\frac{1-T_z'}{s^{*\prime}(z)}\mathcal{S}_c' + \frac{1}{s^{*\prime}(z)}\mathcal{S}_z' + \frac{1}{s^{*\prime}(\theta)}\mathcal{S}_\theta' - \frac{\mathcal{T}_{sz}''}{s^{*\prime}(z)}\right) \\ &= \frac{U_z'}{s^{*\prime}(z)}\mathcal{S}_c' - \frac{U_c'}{s^{*\prime}(z^*)}\mathcal{S}_z' - \frac{U_c'}{s^{*\prime}(\theta)}\mathcal{S}_\theta' + \frac{U_c'}{s^{*\prime}(z^*)}\mathcal{T}_{sz}''.\end{aligned} \tag{73}$$

where we have used $U_z' = -U_c'\left(1-\mathcal{T}_z'\right)$ in the last line.

Similarly, we can obtain the second result of the Lemma by writing $\mathcal{T}_{zz}''$ as

$$\mathcal{T}_{zz}'' = \mathcal{Z}_c'\left(1-\mathcal{T}_z'\right) - \mathcal{Z}_c'\left(1+\mathcal{T}_s'\right)s^{*\prime}(z^*) + \mathcal{Z}_s's^{*\prime}(z^*) + \mathcal{Z}_z' + \frac{\mathcal{Z}_\theta'}{z^{*\prime}(\theta)} - \mathcal{T}_{sz}''s^{*\prime}(z^*). \tag{74}$$

Using

$$\mathcal{Z}_c' = \frac{1}{U_c'}\left(U_{cz}'' + \left(1-\mathcal{T}_z'\right)U_{cc}''\right)$$

as well as

$$\mathcal{Z}_z' = \frac{1}{U_c'}\left(U_{zz}'' + \left(1-\mathcal{T}_z'\right)U_{cz}''\right)$$

we get

$$V''_{zz}\left(s^*\left(\theta\right),z^*\left(\theta\right);\theta\right)=U'_s s^{*\prime}(z^*)\mathcal{Z}'_c-U'_c s^{*\prime}(z^*)\mathcal{Z}'_s-U'_c\frac{\mathcal{Z}'_\theta}{z^{*\prime}(\theta)}+U'_c\mathcal{T}''_{sz}s^{*\prime}(z^*). \tag{75}$$

Finally, to obtain the third result of the Lemma, we must compute $\left(V''_{sz}\right)^2-V''_{ss}V''_{zz}$. Note that the first-order condition $V'_s(s^*(\theta),z^*(\theta);\theta)=0$ holds at every θ by construction. Differentiating with respect to θ we get

$$\frac{d}{d\theta}V'_s(s^*(\theta),z^*(\theta);\theta)=V''_{ss}s^{*\prime}(\theta)+V''_{sz}z^{*\prime}(\theta)+V''_{s\theta}=0 \tag{76}$$

which we can rearrange as

$$-V''_{sz}=V''_{ss}s^{*\prime}(z^*)+\frac{V''_{s\theta}}{z^{*\prime}(\theta)}. \tag{77}$$

Similarly, by totally differentiating the first-order condition $V'_z(s^*(\theta),z^*(\theta);\theta)=0$ and rearranging we find

$$-V''_{sz}=\frac{V''_{zz}}{s^{*\prime}(z^*)}+\frac{V''_{z\theta}}{s^{*\prime}(\theta)}. \tag{78}$$

Writing $\left(V''_{sz}\right)^2$ as the product of the right-hand sides of Equations (77) and (78) yields

$$\begin{aligned}\left(V''_{sz}\right)^2=&\left(V''_{ss}s^{*\prime}(z)+\frac{V''_{s\theta}}{z^{*\prime}(\theta)}\right)\left(\frac{V''_{zz}}{s^{*\prime}(z)}+\frac{V''_{z\theta}}{s^{*\prime}(\theta)}\right)\\ =&V''_{ss}V''_{zz}+\frac{1}{z^{*\prime}(\theta)}V''_{ss}V''_{z\theta}+\frac{1}{s^{*\prime}(\theta)}V''_{zz}V''_{s\theta}+\frac{1}{s^{*\prime}(\theta)z^{*\prime}(\theta)}V''_{s\theta}V''_{z\theta}\end{aligned} \tag{79}$$

Now from the definition $V(s,z;\theta):=U(z-s-\mathcal{T}(s,z),s,z;\theta)$, we can compute

$$\begin{aligned}V''_{s\theta}(s,z;\theta)&=-\left(1+\mathcal{T}'_s(s,z)\right)U''_{c\theta}+U''_{s\theta}\\ V''_{z\theta}(s,z;\theta)&=\left(1-\mathcal{T}'_z(s,z)\right)U''_{c\theta}+U''_{z\theta}\end{aligned}$$

and use the fact that at allocation $\left(c^*\left(\theta\right),s^*\left(\theta\right),z^*\left(\theta\right)\right)$ we have

$$\begin{aligned}\mathcal{S}'_\theta&=\frac{1}{U'_c}\left(U''_{s\theta}-\left(1+\mathcal{T}'_s\right)U''_{c\theta}\right)\\ \mathcal{Z}'_\theta&=\frac{1}{U'_c}\left(U''_{z\theta}+\left(1-\mathcal{T}'_z\right)U''_{c\theta}\right)\end{aligned}$$

to obtain

$$V''_{s\theta}\left(s^*\left(\theta\right),z^*\left(\theta\right);\theta\right)=U'_c\mathcal{S}'_\theta \tag{80}$$

$$V''_{z\theta}\left(s^*\left(\theta\right),z^*\left(\theta\right);\theta\right)=U'_c\mathcal{Z}'_\theta. \tag{81}$$

Substituting these into Equation (79) and rearranging, we have

$$\left(V''_{sz}\right)^2 - V''_{ss}V''_{zz} = \frac{1}{z^{*\prime}(\theta)}V''_{ss}U'_c\mathcal{Z}'_\theta + \frac{1}{s^{*\prime}(\theta)}V''_{zz}U'_c\mathcal{S}'_\theta + \frac{1}{s^{*\prime}(\theta)z^{*\prime}(\theta)}\left(U'_c\right)^2\mathcal{S}'_\theta\mathcal{Z}'_\theta. \tag{82}$$

Expanding V''_{ss} from Equation (73), and V''_{zz} from Equation (75) yields after simplification

$$\left(V''_{sz}\right)^2 - V''_{ss}V''_{zz} = \frac{U'_c}{s^{*\prime}(\theta)}\Bigg[\left(U'_z\mathcal{S}'_c - U'_c\mathcal{S}'_z\right)\mathcal{Z}'_\theta + \left(U'_s\mathcal{Z}'_c - U'_c\mathcal{Z}'_s - U'_c\frac{\mathcal{Z}'_\theta}{s^{*\prime}(\theta)}\right)s^{*\prime}(z^*)\mathcal{S}'_\theta + \left(\mathcal{Z}'_\theta + s^{*\prime}(z^*)\mathcal{S}'_\theta\right)U'_c\mathcal{T}''_{sz}\Bigg],$$

which gives the third result of the Lemma above. □

B.2 Proof of Proposition A.1 & A.2 (Implementation with Simple Tax Systems)

B.2.1 Proof of Proposition A.1

SN tax system. The sufficient conditions for local optimality under the candidate SN tax system follow directly from Lemma B.2 which computes agents' SOCs at the optimal incentive-compatible allocation under a general tax system $\mathcal{T}(s,z)$. Indeed, agents' SOCs are satisfied if Equations (65), (66), and (67) are negative under the SN tax system. Since the cross-partial derivative $\mathcal{T}''_{sz}$ is zero for a SN tax system, it is easy to verify that conditions (42) and (43) on the derivatives of $\mathcal{S}$ and $\mathcal{Z}$, combined with monotonicity ($s^{*\prime}(\theta) > 0$, $s^{*\prime}(z) > 0$) and Assumption 1 on the derivatives of U, jointly imply that each of these three equations is the sum of negative terms. This implies that agents' SOCs are satisfied at the optimal incentive-compatible allocation under the candidate SN tax system.

LED tax system. To derive sufficient conditions for local optimality under the candidate LED tax system, we begin from results obtained in the derivations of Lemma B.2 which computes agents' SOCs at the optimal incentive-compatible allocation. We consider the requirements $V''_{ss} < 0$, $V''_{zz} < 0$, and $V''_{ss}V''_{zz} > (V''_{sz})^2$ in turn.

First, to show that V''_{ss} is negative, note that under a LED tax system, $\mathcal{T}''_{ss} = 0$. Therefore, using the fact that under the candidate LED tax system we have $1 + \mathcal{T}'_s = \frac{U'_s}{U'_c}$ at the optimal incentive-compatible allocation, the general expression for V''_{ss} given in Equation (68) reduces to

$$V''_{ss}(s^*(\theta), z^*(\theta); \theta) = \left(\frac{U'_s}{U'_c}\right)^2 U''_{cc} - 2\frac{U'_s}{U'_c}U''_{cs} + U''_{ss}$$

Therefore when utility is additively separable in c and s (implying $U''_{cs} = 0$), the concavity of preferences ($U''_{cc} \leq 0$ and $U''_{ss} \leq 0$) guarantees that this expression is negative.

Second, to show that V''_{zz} is negative, note that under the candidate LED tax system defined in

Equations (40) and (41) we have

$$\mathcal{T}''_{sz}(s,z) = \tau'_s(z).$$

We can thus find an expression for $\tau'_s(z)$ at any point in the allocation in question by totally differentiating Equation (40) with respect to θ:

$$\begin{aligned}
\tau'_s\left(z^*(\theta)\right) z^{*\prime}(\theta) &= \frac{d}{d\theta}\Big[\mathcal{S}\left(c^*(\theta), s^*(\theta), z^*(\theta); \theta\right)\Big] \\
&= \frac{d}{d\theta}\Big[\mathcal{S}\left(z^*(\theta) - s^*(\theta) - \mathcal{T}\left(s^*(\theta), z^*(\theta)\right), s^*(\theta), z^*(\theta); \theta\right)\Big] \\
&= \mathcal{S}'_c \cdot \left[(1-\mathcal{T}'_z) z^{*\prime}(\theta) - \left(1+\mathcal{T}'_s\right) s^{*\prime}(\theta)\right] + \mathcal{S}'_s s^{*\prime}(\theta) + \mathcal{S}'_z z^{*\prime}(\theta) + \mathcal{S}'_\theta,
\end{aligned}$$

which yields

$$\tau'_s\left(z^*(\theta)\right) = \mathcal{S}'_c \cdot \left(1-\mathcal{T}'_z\right) - \mathcal{S}'_c \cdot \left(1+\mathcal{T}'_s\right) s^{*\prime}(z^*) + \mathcal{S}'_s \cdot s^{*\prime}(z^*) + \mathcal{S}'_z + \frac{\mathcal{S}'_\theta}{z^{*\prime}(\theta)}.$$

Substituting this into the expression for V''_{zz} in (75), we have

$$\begin{aligned}
V''_{zz}\left(s^*(\theta), z^*(\theta); \theta\right) = {} & U'_s s^{*\prime}(z^*)\mathcal{Z}'_c - U'_c s^{*\prime}(z^*)\mathcal{Z}'_s - U'_c \frac{\mathcal{Z}'_\theta}{z^{*\prime}(\theta)} \\
& + U'_c s^{*\prime}(z^*)\left[\mathcal{S}'_c \cdot \left(1-\mathcal{T}'_z\right) - \mathcal{S}'_c \cdot \left(1+\mathcal{T}'_s\right) s^{*\prime}(z^*) + \mathcal{S}'_s \cdot s^{*\prime}(z^*) + \mathcal{S}'_z + \frac{\mathcal{S}'_\theta}{z^{*\prime}(\theta)}\right].
\end{aligned} \tag{83}$$

Now employing the assumption that utility is separable in c, s, and z, (implying both $U''_{cz} = 0$ and $U''_{cs} = 0$) we have

$$\begin{aligned}
U'_s\mathcal{Z}'_c + U'_c\mathcal{S}'_c(1-\mathcal{T}'_z) &= U'_s\mathcal{Z}'_c - U'_z\mathcal{S}'_c \\
&= U'_s \frac{U'_c U''_{cz} - U'_z U''_{cc}}{\left(U'_c\right)^2} - U'_z \frac{U'_c U''_{cs} - U'_s U''_{cc}}{\left(U'_c\right)^2} \\
&= 0.
\end{aligned}$$

Substituting this result into Equation (83), and noting that $\mathcal{Z}'_s = \mathcal{S}'_z = 0$ by separability, yields

$$V''_{zz}(s^*(\theta), z^*(\theta); \theta) = \left(s^{*\prime}(z^*)\right)^2 \left[U'_c\mathcal{S}'_s - U'_s\mathcal{S}'_c\cdot\right] - \frac{U'_c}{z^{*\prime}(\theta)}\left[\mathcal{Z}'_\theta - s^{*\prime}(z^*)\mathcal{S}'_\theta\right] \tag{84}$$

Again employing separability, we have

$$U'_c\mathcal{S}'_s - U'_s\mathcal{S}'_c = U'_c \frac{U'_c U''_{ss} - U'_s U''_{cs}}{\left(U'_c\right)^2} - U'_s \frac{U'_c U''_{cs} - U'_s U''_{cc}}{\left(U'_c\right)^2} = U''_{ss} + \left(\frac{U'_s}{U'_c}\right)^2 U''_{cc} \leq 0,$$

implying that the first term on the right-hand side of Equation (84) is negative. The condition

$\mathcal{Z}'_\theta - s^{*\prime}(z^*)\mathcal{S}'_\theta \geq 0$ from (44) in the Proposition then implies Equation (84) (and thus V''_{zz}) is negative.

Third, to show $V''_{ss}V''_{zz} > (V''_{sz})^2$, we proceed from Equation (67) in Lemma B.2:

$$\begin{aligned}
&\left(V''_{sz}\right)^2 - V''_{ss}V''_{zz} \\
&= \frac{U'_c}{s^{*\prime}(\theta)}\left[\left(U'_z\mathcal{S}'_c - U'_c\mathcal{S}'_z\right)\mathcal{Z}'_\theta + \left(U'_s\mathcal{Z}'_c - U'_c\mathcal{Z}'_s - U'_c\frac{\mathcal{Z}'_\theta}{s^{*\prime}(\theta)}\right)s^{*\prime}(z^*)\mathcal{S}'_\theta + \left(\mathcal{Z}'_\theta + s^{*\prime}(z^*)\mathcal{S}'_\theta\right)U'_c\mathcal{T}''_{sz}\right] \\
&= \left(U'_z\mathcal{S}'_c - U'_c\mathcal{S}'_z\right)\frac{U'_c}{s^{*\prime}(\theta)}\mathcal{Z}'_\theta + \frac{U'_c}{s^{*\prime}(\theta)}\mathcal{Z}'_\theta U'_c\mathcal{T}''_{sz} + \frac{U'_c}{s^{*\prime}(\theta)}\mathcal{S}'_\theta\left(U'_s s^{*\prime}(z^*)\mathcal{Z}'_c - U'_c s^{*\prime}(z^*)\mathcal{Z}'_s - U'_c\frac{\mathcal{Z}'_\theta}{z^{*\prime}(\theta)} + U'_c s^{*\prime}(z^*)\mathcal{T}''_{sz}\right).
\end{aligned}$$

Recognizing that the last bracket term is exactly the expression for V''_{zz} given in Lemma B.2 this gives

$$\left(V''_{sz}\right)^2 - V''_{ss}V''_{zz} = \left(U'_z\mathcal{S}'_c - U'_c\mathcal{S}'_z\right)\frac{U'_c}{s^{*\prime}(\theta)}\mathcal{Z}'_\theta + \frac{U'_c}{s^{*\prime}(\theta)}\mathcal{Z}'_\theta U'_c\mathcal{T}''_{sz} + \frac{U'_c}{s^{*\prime}(\theta)}\mathcal{S}'_\theta V''_{zz}$$

using the previous expression derived for $\mathcal{T}''_{sz} = \tau'_s$, and the fact that separability ensures $\mathcal{S}'_z = 0$, we obtain after simplification

$$\left(V''_{sz}\right)^2 - V''_{ss}V''_{zz} = -\frac{\left(U'_c\right)^2}{s^{*\prime}(\theta)z^{*\prime}(\theta)}\mathcal{Z}'_\theta\left[s^{*\prime}(\theta)\left(\mathcal{S}\cdot\mathcal{S}'_c - \mathcal{S}'_s\right) - \mathcal{S}'_\theta\right] + \frac{U'_c}{s^{*\prime}(\theta)}\mathcal{S}'_\theta V''_{zz}.$$

We have already shown that V''_{zz} is negative. Thus the conditions $\mathcal{S}'_\theta \geq 0$ and $\mathcal{S}'_\theta \leq s^{*\prime}(\theta)\left(\mathcal{S}\cdot\mathcal{S}'_c - \mathcal{S}'_s\right)$ from (44) in the Proposition imply that both terms on the right-hand side are negative, implying that all second-order conditions hold.

B.2.2 Proof of Proposition A.2

We begin with a more general statement, and then derive Proposition A.2 as a corollary. For a fixed type θ, let $c(z,\theta)$ and $s(z,\theta)$ be its preferred consumption and savings choices at earnings z, given the budget constraint induced by $\mathcal{T}(s,z)$

Lemma B.3. *Suppose that $\mathcal{A} = \{(c^*(\theta), s^*(\theta), z^*(\theta))\}_\theta$ constitutes a set of local optima for types θ under a smooth tax system $\mathcal{T}$, where $z^*(\theta)$ is increasing. Individuals' local optima correspond to their global optima when*

1. *$\mathcal{Z} = \frac{U'_z(c,s,z;\theta)}{U'_c(c,s,z;\theta)}$ and $\mathcal{S} = \frac{U'_s(c,s,z;\theta)}{U'_c(c,s,z;\theta)}$ are strictly increasing in θ for all (c,s,z)*
2. *For any two types θ and θ', we cannot have both*

$$\begin{aligned}
&U'_c\Big(c^*(\theta), s^*(\theta), z^*(\theta);\theta\Big)\sigma_c\Big(s^*(\theta), z^*(\theta)\Big) + U'_z\Big(c^*(\theta), s^*(\theta), z;\theta\Big) \qquad (85)\\
&< U'_c\Big(c\left(z^*(\theta),\theta'\right), s\left(z^*(\theta),\theta'\right), z^*(\theta);\theta\Big)\sigma_c\Big(s\left(z^*(\theta),\theta'\right), z^*(\theta)\Big) + U'_z\Big(c\left(z^*(\theta),\theta'\right), s\left(z^*(\theta),\theta'\right), z^*(\theta);\theta\Big)
\end{aligned}$$

and

$$U_s'\Big(c^*(\theta), s^*(\theta), z^*(\theta); \theta\Big)\sigma_c\Big(s^*(\theta), z^*(\theta)\Big) + U_z'\Big(c^*(\theta), s^*(\theta), z; \theta\Big)\Big) \tag{86}$$
$$< U_s'\Big(c\left(z^*(\theta), \theta'\right), s\left(z^*(\theta), \theta'\right), z^*(\theta); \theta\Big)\sigma_s\Big(s\left(z^*(\theta), \theta'\right), z^*(\theta)\Big) + U_z'\Big(c\left(z^*(\theta), \theta'\right), s\left(z^*(\theta), \theta'\right), z^*(\theta); \theta$$

where $\sigma_c(s,z) := 1 - \mathcal{T}_z'(s,z)$ *and* $\sigma_s(s,z) := \frac{1-\mathcal{T}_z'(s,z)}{1+\mathcal{T}_s'(s,z)}$.

Condition 1 corresponds to single crossing assumptions for earnings and savings. Condition 2 is a requirement that if type θ preserves its assigned earnings level $z^*(\theta)$, but chooses some other consumption level s (corresponding to a level that some other type θ' would choose if forced to choose earnings level $z^*(\theta)$), then at this alternative consumption bundle agent θ cannot have both higher marginal utility from increasing its savings through one more unit of work *and* increasing its consumption through one more unit of work. Generally, this condition will hold as long as U is sufficiently concave in consumption and savings when type θ chooses earnings level $z^*(\theta)$.

Proof. To prove agents' local optima are global optima, we want to show that for any given agent θ^*, utility decreases when moving from allocation $(c^*(\theta^*), s^*(\theta^*), z^*(\theta^*))$ to allocation $(c(z,\theta^*), s(z,\theta^*), z)$.

The first step is to compute agent θ^* utility change. The envelope theorem applied to savings choices $s(z,\theta^*)$ implies

$$\begin{aligned}&\frac{d}{dz}U\left(c(z,\theta^*), s(z,\theta^*), z; \theta^*\right)\\ &= U_c'\left(c(z,\theta^*), s(z,\theta^*), z; \theta^*\right)\sigma_c\left(s(z,\theta^*), z\right) + U_z'\left(c(z,\theta^*), s(z,\theta^*), z; \theta^*\right)\end{aligned}$$

where $\sigma_c(s,z) = 1 - \mathcal{T}_z'(s,z)$. Note that, as established by Milgrom and Segal (2002), these equalities hold as long as U is differentiable in z (holding s and c fixed)—differentiability of $c(z,\theta^*)$ or $s(z,\theta^*)$ is actually not required.

Similarly, the envelope theorem applied to consumption choices $c(z,\theta^*)$ implies

$$\begin{aligned}&\frac{d}{dz}U\left(c(z,\theta^*), s(z,\theta^*), z; \theta^*\right)\\ &= U_s'\left(c(z,\theta^*), s(z,\theta^*), z; \theta^*\right)\sigma_s\left(s(z,\theta^*), z\right) + U_z'\left(c(z,\theta^*), s(z,\theta^*), z; \theta^*\right)\end{aligned} \tag{87}$$

where $\sigma_s(s,z) = \frac{1-\mathcal{T}_z'(s,z)}{1+\mathcal{T}_s'(s,z)}$.

Therefore, agent's θ^* utility change when moving from allocation $(c^*(\theta^*), s^*(\theta^*), z^*(\theta^*))$ to allocation $(c(z,\theta^*), s(z,\theta^*), z)$ is

$$\begin{aligned}&U\left(c(z,\theta^*), s(z,\theta^*), z; \theta^*\right) - U\left(c(z^*(\theta^*),\theta^*), s(z^*(\theta^*),\theta^*), z^*(\theta^*); \theta^*\right)\\ &= \int_{x=z^*(\theta^*)}^{x=z}\Bigg[\min\left\{U_c'\left(c(x,\theta^*), s(x,\theta^*), x; \theta^*\right)\sigma_c\left(s(x,\theta^*), x\right), U_s'\left(c(x,\theta^*), s(x,\theta^*), x; \theta^*\right)\sigma_s\left(s(x,\theta^*), x\right)\right\}\\ &\qquad + U_z'\left(c(x,\theta^*), s(x,\theta^*), x; \theta^*\right)\Bigg]dx\end{aligned} \tag{88}$$

where the min operator is introduced without loss of generality given the fact that both terms are equal.

The second step is to show that under our assumptions, agent θ^* utility change (88) is negative. To do so, let θ_x be the type that chooses earnings x. Then, by definition, agent θ_x utility is maximal at earnings x implying both

$$\begin{aligned}
&U'_c\left(c^*(\theta_x), s^*(\theta_x), x; \theta_x\right) \sigma_c\left(s^*(\theta_x), x\right) + U'_z\left(c^*(\theta_x), s^*(\theta_x), x; \theta_x\right) = 0 \\
&U'_s\left(c^*(\theta_x), s^*(\theta_x), x; \theta_x\right) \sigma_s\left(s^*(\theta_x), x\right) + U'_z\left(c^*(\theta_x), s^*(\theta_x), x; \theta_x\right) = 0
\end{aligned}$$

such that

$$\begin{aligned}
&\max\left\{U'_c\left(c^*(\theta_x), s^*(\theta_x), x; \theta_x\right) \sigma_c\left(s^*(\theta_x), x\right), U'_s\left(c^*(\theta_x), s^*(\theta_x), x; \theta_x\right) \sigma_s\left(s^*(\theta_x), x\right)\right\} \\
&+ U'_z\left(c^*(\theta_x), s^*(\theta_x), x; \theta_x\right) = 0
\end{aligned} \tag{89}$$

Now, by condition 2, we either have[17]

$$\begin{aligned}
&U'_c\left(c^*(\theta_x), s^*(\theta_x), x; \theta_x\right) \sigma_c\left(s^*(\theta_x), x\right) + U'_z\left(c^*(\theta_x), s^*(\theta_x), x; \theta_x\right) \\
&\geq U'_c\left(c(x, \theta^*), s(x, \theta^*), x; \theta_x\right) \sigma_c\left(s(x, \theta^*), x\right) + U'_z\left(c(x, \theta^*), s(x, \theta^*), x; \theta_x\right)
\end{aligned}$$

or

$$\begin{aligned}
&U'_s\left(c^*(\theta_x), s^*(\theta_x), x; \theta_x\right) \sigma_s\left(s^*(\theta_x), x\right) + U'_z\left(c^*(\theta_x), s^*(\theta_x), x; \theta_x\right) \\
&\geq U'_s\left(c(x, \theta^*), s(x, \theta^*), x; \theta_x\right) \sigma_s\left(s(x, \theta^*), x\right) + U'_z\left(c(x, \theta^*), s(x, \theta^*), x; \theta_x\right)
\end{aligned}$$

implying that

$$\begin{aligned}
&\max\left\{U'_c\left(c^*(\theta_x), s^*(\theta_x), x; \theta_x\right) \sigma_c\left(s^*(\theta_x), x\right), U'_s\left(c^*(\theta_x), s^*(\theta_x), x; \theta_x\right) \sigma_s\left(s^*(\theta_x), x\right)\right\} \\
&+ U'_z\left(c^*(\theta_x), s^*(\theta_x), x; \theta_x\right) \\
&\geq \min\left\{U'_c\left(c(x, \theta^*), s(x, \theta^*), x; \theta_x\right) \sigma_c\left(s(x, \theta^*), x\right), U'_s\left(c(x, \theta^*), s(x, \theta^*), x; \theta_x\right) \sigma_s\left(s(x, \theta^*), x\right)\right\} \\
&+ U'_z\left(c(x, \theta^*), s(x, \theta^*), x; \theta_x\right)
\end{aligned} \tag{90}$$

But since the maximum is zero, this minimum has to be negative. Hence, we have either

$$\begin{aligned}
&U'_c\left(c(x, \theta^*), s(x, \theta^*), x; \theta_x\right) \sigma_c\left(s(x, \theta^*), x\right) + U'_z\left(c(x, \theta^*), s(x, \theta^*), x; \theta_x\right) \leq 0 \\
&\iff \frac{U'_z\left(c(x, \theta^*), s(x, \theta^*), x; \theta_x\right)}{U'_c\left(c(x, \theta^*), s(x, \theta^*), x; \theta_x\right)} \leq -\sigma_c\left(s(x, \theta^*), x\right)
\end{aligned}$$

[17] Not having $\{a < c \quad and \quad b < c\}$ means having $\{a \geq c \quad \text{or} \quad b \geq d\}$ which implies $\max(a, b) \geq \min(c, d)$

or

$$\begin{aligned} &U'_s\left(c(x,\theta^*),s(x,\theta^*),x;\theta_x\right)\sigma_s\left(s(x,\theta^*),x\right)+U'_z\left(c(x,\theta^*),s(x,\theta^*),x;\theta_x\right)\leq 0\\ \iff &\frac{U'_z\left(c(x,\theta^*),s(x,\theta^*),x;\theta_x\right)}{U'_s\left(c(x,\theta^*),s(x,\theta^*),x;\theta_x\right)}\leq -\sigma_c\left(s(x,\theta^*),x\right)\end{aligned}$$

Suppose that $z > z^*(\theta^*)$ such that $x > z^*(\theta^*)$; the case $z < z^*(\theta^*)$ follows identically. For any $x > z^*(\theta^*)$, the monotonicity of the earnings function means that $\theta_x > \theta^*$. Then, by the single-crossing conditions for $\mathcal{Z} = \frac{U'_z}{U'_c}$ and $\mathcal{S} = \frac{U'_s}{U'_c}$, this means that we have either[18]

$$\frac{U'_z\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)}{U'_c\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)}\leq -\sigma_c\left(s(x,\theta^*),x\right)$$

or

$$\frac{U'_z\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)}{U'_s\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)}\leq -\sigma_c\left(s(x,\theta^*),x\right)$$

implying that for any $x > z^*(\theta^*)$,

$$\begin{aligned} &\min\left\{U'_c\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)\sigma_c\left(s(x,\theta^*),x\right),U'_s\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)\sigma_s\left(s(x,\theta^*),x\right)\right\}\\ &+U'_z\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)\leq 0. \end{aligned} \tag{91}$$

As a result, the right hand-side of Equation (88) is an integral of negative terms, which shows that

$$U\left(c(z,\theta^*),s(z,\theta^*),z;\theta^*\right)-U\left(c^*(\theta^*),s^*(\theta^*),z^*(\theta^*);\theta^*\right)\leq 0 \tag{92}$$

The case with $z < z^*(\theta^*)$ follows identically, proving Lemma B.3. □

Proof of Proposition A.2

We now derive Proposition A.2 as a consequence of Lemma B.3 by deriving assumptions under which condition 2 is met for SN and LED tax systems.

SN systems First, suppose that $s < s^*(\theta)$, then $c > c^*(\theta)$. Noting that $\sigma_c = 1 - T'_z\left(z^*\left(\theta\right)\right)$ is not a function of s, we can use $U''_{cc} \leq 0$ and $U''_{cs} \geq 0$ to obtain

$$U'_c\left(c^*(\theta),s^*(\theta),z^*(\theta);\theta\right)\sigma_c\left(s^*(\theta),z^*(\theta)\right)\geq U'_c\left(c,s,z^*(\theta);\theta\right)\sigma_c\left(s,z^*(\theta)\right).$$

Further relying on the fact that $U''_{cz} = 0$ and $U''_{sz} = 0$, we obtain

$$\begin{aligned} &U'_c\left(c^*(\theta),s^*(\theta),z^*(\theta);\theta\right)\sigma_c\left(s^*(\theta),z^*(\theta)\right)+U'_z\left(c^*(\theta),s^*(\theta),z^*(\theta);\theta\right)\\ &\geq U'_c\left(c,s,z^*(\theta);\theta\right)\sigma_c\left(s,z^*(\theta)\right)+U'_z\left(c,s,z^*(\theta);\theta\right). \end{aligned}$$

[18]Note that having both $\mathcal{Z}$ and $\mathcal{S}$ increasing in θ also implies that $\frac{\mathcal{Z}}{\mathcal{S}} = \frac{U'_z}{U'_s}$ is increasing in θ.

Conversely, suppose that $s > s^*(\theta)$, then $c < c^*(\theta)$. We have

$$\frac{d}{ds}\left[\frac{U'_s(z - T_z(z) - s - T_s(s), s, z^*(\theta);\theta)}{1 + T'_s(s)}\right] = -U''_{cs} + \frac{1}{(1 + T'_s(s))}\left[U''_{ss} - U'_s\frac{T''_{ss}(s)}{1 + T'_s(s)}\right].$$

The condition that $\frac{U''_{ss}(c(s,\theta),s,z^*(\theta);\theta)}{U'_s(c(s,\theta),s,z^*(\theta);\theta)} < \frac{T''_{ss}(s)}{1+T'_s(s)}$, together with $U''_{cs} > 0$, implies that $\frac{U'_s(c(s,\theta),s,z^*(\theta);\theta)}{1+T'_s(s)}$ is decreasing in s and thus that

$$\frac{U'_s(c^*(\theta), s^*(\theta), z^*(\theta);\theta)}{1 + T'_s(s^*(\theta))} \geq \frac{U'_s(c, s, z^*(\theta);\theta)}{1 + T'_s(s)}.$$

Further relying on the fact that $U''_{cz} = 0$ and $U''_{sz} = 0$, and that $\mathcal{T}'_s = T'_z(z)$ is independant of s, we obtain

$$\begin{aligned} &U'_s(c^*(\theta), s^*(\theta), z^*(\theta);\theta)\,\sigma_s(s^*(\theta), z^*(\theta)) + U'_z(c^*(\theta), s^*(\theta), z^*(\theta);\theta) \\ &\geq U'_s(c, s, z^*(\theta);\theta)\,\sigma_s(s, z^*(\theta)) + U'_z(c, s, z^*(\theta);\theta). \end{aligned}$$

LED systems First, consider a type θ' choosing earnings $z = z^*(\theta) > z^*(\theta')$. We have

$$\begin{aligned} &\frac{d}{ds}\left[U'_c(z - s - \tau_s(z^*(\theta))s - T_z(z^*(\theta)), s, z^*(\theta);\theta)\left(1 - T'_z(z^*(\theta)) - \tau'_s(z^*(\theta))s\right)\right] \\ &= U''_{cs}\left(1 - T'_z(z^*(\theta)) - \tau'_s(z^*(\theta))s\right) - U''_{cc}(1 + \tau_s(z^*(\theta)))\left(1 - T'_z(z^*(\theta)) - \tau'_s(z^*(\theta))s\right) - U'_c\tau'_s(z^*(\theta)). \end{aligned}$$

The first term is negative because $U''_{cs} \geq 0$ and $1 - \mathcal{T}'_z = -\mathcal{Z} \geq 0$. Now, the condition that $\mathcal{S} = U'_s/U'_c$ is increasing in θ ensures that a type θ' choosing earnings $z^*(\theta) > z^*(\theta')$ has a desired savings level $s(z^*(\theta), \theta') < s^*(\theta)$. In this case, condition *(a)* of the proposition implies that the remaining terms are negative such that

$$U'_c(z - s - \tau_s(z^*(\theta))s - T(z^*(\theta)), s, z^*(\theta);\theta)\,\sigma_c(s, z^*(\theta))$$

is increasing in s for $s < s^*(\theta)$, where $\sigma_c(s, z^*(\theta)) = 1 - T'_z(z) - \tau'_s(z)s$. As a result,

$$\begin{aligned} &U'_c(c^*(\theta), s^*(\theta), z^*(\theta);\theta)\,\sigma_c(s^*(\theta), z^*(\theta)) \\ &\geq U'_c\left(c(z^*(\theta), \theta'), s(z^*(\theta), \theta'), z^*(\theta);\theta\right)\sigma_c\left(s(z^*(\theta), \theta'), z^*(\theta)\right) \end{aligned}$$

and thus relying on the fact that $U''_{cz} = 0$ and $U''_{sz} = 0$, we have

$$\begin{aligned} &U'_c(c^*(\theta), s^*(\theta), z^*(\theta);\theta)\,\sigma_c(s^*(\theta), z^*(\theta)) + U'_z(c^*(\theta), s^*(\theta), z^*(\theta);\theta) \\ &\geq U'_c\left(c(z^*(\theta), \theta'), s(z^*(\theta), \theta'), z^*(\theta);\theta\right)\sigma_c\left(s(z^*(\theta), \theta'), z^*(\theta)\right) + U'_z\left(c(z^*(\theta), \theta'), s(z^*(\theta), \theta'), z^*(\theta);\theta\right). \end{aligned}$$

Second consider a type θ' choosing $z = z^*(\theta) < z^*(\theta')$. We have

$$\frac{d}{ds}\left[U'_s\left(z-s-\tau_s(z^*(\theta))s-T(z^*(\theta)),s,z^*(\theta);\theta\right)\frac{1-T'_z\left(z^*(\theta)\right)-\tau'_s\left(z^*(\theta)\right)s}{1+\tau_s(z)}\right]$$
$$=-U''_{cs}\left(1-T'_z\left(z^*(\theta)\right)-\tau'_s\left(z^*(\theta)\right)s\right)+U''_{ss}\frac{1-T'_z\left(z^*(\theta)\right)-\tau'_s\left(z^*(\theta)\right)s}{1+\tau_s(z)}+U'_s\frac{\tau'_s\left(z^*(\theta)\right)}{1+\tau_s(z)}$$

The first term is negative because $U''_{cs} \geq 0$ and $1-\mathcal{T}'_z = -\mathcal{Z} \geq 0$. Now, the condition that $\mathcal{S} = U'_s/U'_c$ is increasing in θ ensures that a type θ' choosing earnings $z = z^*(\theta) < z^*(\theta')$ has a desired savings level $s(z^*(\theta),\theta') > s^*(\theta)$. Hence, condition *(b)* of the proposition implies that the remaining terms are negative such that

$$U'_s\left(z-s-\tau_s(z^*(\theta))s-T(z^*(\theta)),s,z^*(\theta);\theta\right)\sigma_s\left(s,z^*(\theta)\right)$$

is decreasing in s for $s > s^*(z)$, where $\sigma_s\left(s,z^*(\theta)\right) = \frac{1-T'_z(z)-\tau'_s(z)s}{1+\tau_s(z)}$. This ensures that

$$U'_s\left(c^*(\theta),s^*(\theta),z^*(\theta);\theta\right)\sigma_c\left(s^*(\theta),z^*(\theta)\right)$$
$$\geq U'_s\left(c(z^*(\theta),\theta'),s(z^*(\theta),\theta'),z^*(\theta);\theta\right)\sigma_c\left(s(z^*(\theta),\theta'),z^*(\theta)\right)$$

and thus relying on the fact that $U''_{cz} = 0$ and $U''_{sz} = 0$, we have

$$U'_s\left(c^*(\theta),s^*(\theta),z^*(\theta);\theta\right)\sigma_c\left(s^*(\theta),z^*(\theta)\right)+U'_z\left(c^*(\theta),s^*(\theta),z^*(\theta);\theta\right)$$
$$\geq U'_s\left(c(z^*(\theta),\theta'),s(z^*(\theta),\theta'),z^*(\theta);\theta\right)\sigma_c\left(s(z^*(\theta),\theta'),z^*(\theta)\right)+U'_z\left(c(z^*(\theta),\theta'),s(z^*(\theta),\theta'),z^*(\theta);\theta\right).$$

B.3 Proof of Proposition 1 (Measurement of Causal Income Effects)

We here derive the different expressions of the sufficient statistic $s'_{inc}(z)$ can be measured empirically.

Case 1. If agents preferences are weakly separable between the utility of consumption $u(.)$ and the disutility to work $k(.)$, agent θ problem writes

$$\max_{c,s,z}\ u(c,s;\theta)-k\left(z/w(\theta)\right)\ \ s.t.\ c \leq z-s-\mathcal{T}(s,z)$$

meaning that conditional on earnings z, savings $s(z;\theta)$ is defined as the solution to

$$-\left(1+\mathcal{T}'_s(s,z)\right)u'_c\left(z-s-\mathcal{T}(s,z),s;\theta\right)+u'_s\left(z-s-\mathcal{T}(s,z),s;\theta\right)=0.$$

Differentiating this equation with respect to savings s and earnings z yields

$$\frac{\partial s}{\partial z} = -\frac{\left[-\mathcal{T}''_{sz}u'_c-\left(1+\mathcal{T}'_s\right)\left(1-\mathcal{T}'_z\right)u''_{cc}+\left(1-\mathcal{T}'_z\right)u''_{cs}\right]}{\left[-\mathcal{T}''_{ss}u'_c+\left(1+\mathcal{T}'_s\right)^2u''_{cc}-2\left(1+\mathcal{T}'_s(s,z)\right)u''_{cs}+u''_{ss}\right]}.$$

Differentiating this equation with respect to savings s and disposable income y yields

$$\frac{\partial s}{\partial y} = -\frac{\left[-\left(1+\mathcal{T}_s'\right)u_{cc}'' + u_{cs}''\right]}{\left[-\mathcal{T}_{ss}''u_c' + \left(1+\mathcal{T}_s'\right)^2 u_{cc}'' - 2\left(1+\mathcal{T}_s'\left(s,z\right)\right)u_{cs}'' + u_{ss}''\right]}.$$

Hence, if $\mathcal{T}_{sz}'' = 0$, we get

$$s_{inc}'\left(z\right) := \frac{\partial s\left(z;\theta\right)}{\partial z} = \left(1-\mathcal{T}_z'\right)\frac{\partial s}{\partial y} = \left(1-\mathcal{T}_z'\right)\frac{\eta_{s|z}\left(z(\theta)\right)}{1+\mathcal{T}_s'},$$

where the last equality follows from the definition of $\eta_{s|z}\left(z(\theta)\right)$. The intuition behind this result is that with separable preferences, savings s depend on earnings z only through disposable income $y = z - s - \mathcal{T}\left(s,z\right)$.

Case 2. If agents wage rates w and hours h are observable, and earnings z are given by $z = w \cdot h$, we can infer s_{inc}' from changes in wages through

$$\begin{aligned}
&\frac{\partial s}{\partial w} = \frac{\partial s\left(w\cdot h;\theta\right)}{\partial w} = \frac{\partial s\left(z;\theta\right)}{\partial z}\left(1+\frac{\partial h}{\partial w}\right) \\
&\iff \frac{\partial s\left(z;\theta\right)}{\partial z} = \frac{\frac{\partial s}{\partial w}}{1+\frac{\partial h}{\partial w}} = s\frac{\frac{w}{s}\frac{\partial s}{\partial w}}{w + h\frac{w}{h}\frac{\partial h}{\partial w}} \\
&\iff s_{inc}'\left(z\right) = s\left(z\right)\frac{\xi_w^s\left(z\right)}{w\left(z\right) + h\left(z\right)\xi_w^h\left(z\right)}
\end{aligned}$$

where $\xi_w^s\left(z\right) \equiv \frac{w(z)}{s(z)}\frac{\partial s(z)}{\partial w(z)}$ is individuals' elasticity of savings with respect to their wage rate, and $\xi_w^h\left(z\right) \equiv \frac{w(z)}{h(z)}\frac{\partial h(z)}{\partial w(z)}$ is individuals' elasticity of hours with respect to their wage rate.

Case 3. Otherwise, if we can measure the elasticity of savings s and earnings z upon a compensated change in the marginal earnings tax rate $\mathcal{T}_z'$, respectively denoted $\chi_s^c := -\frac{1-\mathcal{T}_z'}{s}\frac{\partial s}{\partial \mathcal{T}_z'}$ and $\zeta_z^c := -\frac{1-\mathcal{T}_z'}{z}\frac{\partial z}{\partial \mathcal{T}_z'}$, we then have

$$\begin{aligned}
&\frac{\partial s}{\partial \mathcal{T}_z'} = \frac{\partial s\left(z;\theta\right)}{\partial z}\frac{\partial z}{\mathcal{T}_z'} \\
&\iff \left(-\frac{s}{1-\mathcal{T}_z'}\chi_s^c\right) = s_{inc}'\left(z\right)\left(-\frac{z}{1-\mathcal{T}_z'}\zeta_z^c\right) \\
&\iff s_{inc}'\left(z\right) = \frac{s\left(z\right)}{z}\frac{\chi_s^c\left(z\right)}{\zeta_z^c\left(z\right)}.
\end{aligned}$$

B.4 Proof of Lemma 1 (Earnings Responses to Taxes on s)

Let

$$V\left(\mathcal{T}(.,z),z;\theta\right) = \max_s \; U\left(z - s - \mathcal{T}(s,z), s, z; \theta\right)$$

be agent θ indirect utility function at earnings z.

Consider a savings tax reform $d\mathcal{T}_s(s)$ that consists in a small increase $d\tau_s$ in the marginal savings tax rate in a bandwidth of savings $\left[s^0, s^0 + ds\right]$, with $d\tau_s$ much smaller than ds:

$$d\mathcal{T}_s(s) = \begin{cases} 0 & if\ s \leq s^0 \\ d\tau_s(s - s^0) & if\ s \in [s^0, s^0 + ds] \\ d\tau_s ds & if\ s \geq s^* + ds \end{cases}$$

The idea is to construct for each type θ a perturbation of the earnings tax $d\mathcal{T}_z^\theta(z)$ that induces the same earnings response as the initial perturbation $d\mathcal{T}_s(s)$.

Suppose we define this perturbation for each type θ such that at all earnings z,

$$V(\mathcal{T}(.,z) + d\mathcal{T}_s(.), z; \theta) = V(\mathcal{T}(.,z) + d\mathcal{T}_z^\theta(.), z; \theta)$$

then, by construction, the perturbation $d\mathcal{T}_z^\theta(.)$ induces the same earnings response dz as the initial perturbation $d\mathcal{T}_s(s)$.

Moreover, both tax reforms must induce the same utility change for agent θ. Applying the envelope theorem to agents θ for whom $s(z;\theta) \in [s^0, s^0 + ds]$, yields

$$\begin{aligned} & U_c' \cdot d\tau_s \left(s(z;\theta) - s^0\right) = U_c' \cdot d\mathcal{T}_z^\theta(z) \\ \iff & d\mathcal{T}_z^\theta(z) = \left(s(z;\theta) - s^0\right) d\tau_s. \end{aligned}$$

As can be seen differentiating both sides by earnings z, and letting $ds \to 0$, this implies that a small increase $d\tau_s$ in the marginal savings tax rate induces the same earnings change as a small increase $s'_{inc}(z)\, d\tau_s$ in the marginal earnings tax rate, which proves Lemma 1.

Note that this proof does not rely on heterogeneity being unidimensional, meaning that this result will equally apply when we consider multidimensional heterogeneity.

B.5 Proof of Theorem 2 (Optimal Smooth Tax Systems)

Since $z(\theta)$ is a strictly increasing function, it defines a bijection and we denote $\vartheta(z)$ its inverse, which indicates the type θ with earnings z. This allows us to define $\alpha(z) := \alpha(\vartheta(z))$, $c(z) := c(z; \vartheta(z))$, and $s(z) := s(z; \vartheta(z))$.

With these notations, the problem of the government is to maximize the following Lagrangian

$$\mathcal{L} = \int_z \Big[\alpha(z)\, U\Big(c(z), s(z), z; \vartheta(z)\Big) + \lambda\Big(\mathcal{T}(s(z), z) - E\Big)\Big]\, dH_z(z). \tag{93}$$

where λ is the social marginal value of public funds, and the tax function implicitly enters agents' utility through $c(z) = z - s(z) - \mathcal{T}(s(z), z)$.

B.5.1 Optimality condition for marginal earnings tax rates

Reform. We consider a small reform at earnings level $z^0 = z(\theta^0)$ that consists in a small increase $d\tau_z$ of the marginal tax rate on earnings in a small bandwidth dz. Formally,

$$d\mathcal{T}(s,z) = \begin{cases} 0 & if\ z \leq z^0 \\ d\tau_z(z-z^0) & if\ z \in [z^0, z^0+dz] \\ d\tau_z dz & if\ z \geq z^0+dz \end{cases}$$

We characterize the impact of this reform on the government's objective function $\mathcal{L}$ as $dz \to 0$. Normalizing all effects by $1/\lambda$, the reform induces

- *mechanical effects*:

$$\int_{z^0}^{\bar{z}} \left(1 - \frac{\alpha(z)}{\lambda} U_c'\left(c(z), s(z), z; \vartheta(z)\right)\right) d\tau_z dz\, dH_z(z)$$

- *behavioral effects from earnings changes*:[19]

$$-\mathcal{T}_z'\left(s\left(z^0\right), z^0\right) \frac{z^0}{1-\mathcal{T}_z'\left(s\left(z^0\right), z^0\right)} \zeta_z^c(z^0) d\tau_z\, dz h_z(z^0) \\ - \int_{z^0}^{\bar{z}} \mathcal{T}_z'\left(s\left(z\right), z\right) \frac{\eta_z(z)}{1-\mathcal{T}_z'\left(s\left(z\right), z\right)} d\tau_z dz\, dH_z(z)$$

- *behavioral effects from savings changes*:

$$-\mathcal{T}_s'\left(s(z^0), z^0\right) s_{inc}'(z^0) \left[\frac{z^0}{1-\mathcal{T}_z'\left(s\left(z^0\right), z^0\right)} \zeta_z^c(z^0) d\tau_z\right] dz h_z(z^0) \\ - \int_{z^0}^{\bar{z}} \mathcal{T}_s'\left(s(z), z\right) \left[\frac{\eta_{s|z}(z)}{1+\mathcal{T}_s'\left(s(z), z\right)} + s_{inc}'(z) \frac{\eta_z(z)}{1-\mathcal{T}_z'\left(s\left(z\right), z\right)}\right] d\tau_z dz\, dH_z(z)$$

Summing over these different effects yields the total impact of the reform

$$\frac{1}{\lambda}\frac{d\mathcal{L}}{dz} = \int_{z^0}^{\bar{z}} \left(1 - \hat{g}(z)\right) d\tau_z dH_z(z) \\ - \left(\mathcal{T}_z'\left(s\left(z^0\right), z^0\right) + s_{inc}'(z^0)\mathcal{T}_s'\left(s(z^0), z^0\right)\right) \frac{z^0}{1-\mathcal{T}_z'\left(s\left(z^0\right), z^0\right)} \zeta_z^c(z^0) d\tau_z\, h_z(z^0) \tag{94}$$

[19] Note that by definition elasticity concepts include all circularities and adjustments induced by tax reforms such that savings and earnings changes are simply given by

$$\begin{cases} dz = -\frac{z}{1-\mathcal{T}_z'} \zeta_z^c(z) d\mathcal{T}_z'(s,z) - \frac{\eta_z(z)}{1-\mathcal{T}_z'} d\mathcal{T}(s,z) \\ ds = -\frac{\eta_{s|z}(z)}{1+\mathcal{T}_s'} d\mathcal{T}(s,z) + s_{inc}'(z) dz \end{cases}$$

where we have plugged in $\hat{g}(z)$ the social marginal welfare weights augmented with income effects

$$\hat{g}(z) = \frac{\alpha(z)}{\lambda} U_c'\left(c(z), s(z), z; \vartheta(z)\right) + \mathcal{T}_z'\left(s\left(z\right), z\right) \frac{\eta_z(z)}{1 - \mathcal{T}_z'\left(s\left(z\right), z\right)} + \mathcal{T}_s'\left(s(z), z\right) \left(\frac{\eta_{s|z}(z)}{1 + \mathcal{T}_s'\left(s(z), z\right)} + s'_{inc}(z) \frac{\eta_z(z)}{1 - \mathcal{T}_z'\left(s\left(z\right), z\right)} \right).$$

Optimality. A direct implication of this result is a sufficient statistics characterization of the optimal schedule of marginal earnings tax rate $\mathcal{T}_z'$. Indeed, at the optimum the reform should have a zero impact on the government objective, $d\mathcal{L} = 0$, meaning that at each earnings z^0 the optimal marginal earnings tax rate satisfies

$$\frac{\mathcal{T}_z'\left(s\left(z^0\right), z^0\right)}{1 - \mathcal{T}_z'\left(s\left(z^0\right), z^0\right)} = \frac{1}{\zeta_z^c(z^0)} \frac{1}{z^0 h_z(z^0)} \int_{z=z^0}^{\bar{z}} \left(1 - \hat{g}(z)\right) dH_z(z) - s'_{inc}(z^0) \frac{\mathcal{T}_s'\left(s(z^0), z^0\right)}{1 - \mathcal{T}_z'\left(s\left(z^0\right), z^0\right)} \tag{95}$$

which is the optimality condition (15) presented in Theorem (2).

B.5.2 Optimality condition for marginal savings tax rates

Reform. We consider a small reform of the savings tax at savings level $s^0 = s(\theta^0)$ that consists in a small increase $d\tau_s$ of the marginal tax rate on savings in a small bandwidth ds. Formally,

$$d\mathcal{T}(s, z) = \begin{cases} 0 & if\ s \leq s^0 \\ d\tau_s (s - s^0) & if\ s \in [s^0, s^0 + ds] \\ d\tau_s ds & if\ s \geq s^0 + ds \end{cases}$$

Assuming agents' preferences are smooth across the type distribution, there exists an increasing mapping between earnings z and savings s. Denote z^0 the earnings level such that $s^0 = s(z^0)$. We characterize the impact of this reform on the government objective function $\mathcal{L}$ as $ds \to 0$. Normalizing all effects by $1/\lambda$, the reform induces

- *mechanical effects:*

$$\int_{z \geq z^0} \left(1 - \frac{\alpha(z)}{\lambda} U_c'\left(c(z), s(z), z; \vartheta(z)\right)\right) d\tau_s ds\, dH_z(z)$$

- *behavioral effects from earnings changes*:[20]

$$-\mathcal{T}_z'(s^0,z^0)\left[\frac{z^0}{1-\mathcal{T}_z'(s^0,z^0)}\zeta_z^c(z^0)\,s_{inc}'(z)d\tau_s\right]ds\,\frac{h_z(z^0)}{s'(z^0)}-\int_{z^0}^{\bar{z}}\mathcal{T}_z'\left(s,z\right)\frac{\eta_z(z)}{1-\mathcal{T}_z'\left(s,z\right)}d\tau_s ds\,dH_z(z)$$

- *behavioral effects from savings changes*:

$$-\mathcal{T}_s'\left(s^0,z^0\right)\left[\frac{s^0}{1+\mathcal{T}_s'\left(s^0,z^0\right)}\zeta_{s|z}^c(z^0)d\tau_s+s_{inc}'(z^0)\frac{z^0}{1-\mathcal{T}_z'\left(s^0,z^0\right)}\zeta_z^c(z^0)\,s_{inc}'(z^0)d\tau_s\right]ds\,\frac{h_z(z^0)}{s'(z^0)}$$
$$-\int_{z^0}^{\bar{z}}\mathcal{T}_s'\left(s,z\right)\left[\frac{\eta_{s|z}(z)}{1+\mathcal{T}_s'\left(s,z\right)}+s_{inc}'(z)\frac{\eta_z(z)}{1-\mathcal{T}_z'\left(s,z\right)}\right]d\tau_s ds\,dH_z(z)$$

Summing over these different effects yields the total impact of the reform

$$\frac{1}{\lambda}\frac{d\mathcal{L}}{ds}=s'\left(z^0\right)\int_{z=z^0}^{\bar{z}}\left(1-\hat{g}(z)\right)d\tau_s\,dH_z(z)-\Big\{\mathcal{T}_s'\left(s^0,z^0\right)\frac{s^0}{1+\mathcal{T}_s'\left(s^0,z^0\right)}\zeta_{s|z}^c(z^0)$$
$$+\left[\mathcal{T}_z'(s^0,z^0)+s_{inc}'(z^0)\mathcal{T}_s'(s^0,z^0)\right]\frac{z^0}{1-\mathcal{T}_z'\left(s^0,z^0\right)}\zeta_z^c(z^0)s_{inc}'(z^0)\Big\}d\tau_s\,h_z(z^0) \tag{97}$$

Optimality. A direct implication of this result is a sufficient statistics characterization of the optimal schedule of marginal savings tax rates $\mathcal{T}_s'$. Indeed, at the optimum the reform should have a zero impact on the government objective, $d\mathcal{L}=0$, which implies that at each savings $s^0=s\left(z^0\right)$ the optimal marginal earnings tax rate satisfies

$$\frac{\mathcal{T}_s'(s^0,z^0)}{1+\mathcal{T}_s'(s^0,z^0)}s^0\zeta_{s|z}^c(z^0)h_z(z^0)$$
$$=s'\left(z^0\right)\int_{z^0}^{\bar{z}}(1-\hat{g}(z))\,dH_z(z)-s_{inc}'(z^0)\frac{\mathcal{T}_z'(s^0,z^0)+s_{inc}'(z^0)\mathcal{T}_s'(s^0,z^0)}{1-\mathcal{T}_z'(s^0,z^0)}z^0\zeta_z^c(z^0)h_z(z^0) \tag{98}$$

Note that if we use the formula for the optimal schedule of marginal earnings tax rates (95) to replace the last term, this formula can be rearranged as

$$\frac{\mathcal{T}_s'(s^0,z^0)}{1+\mathcal{T}_s'(s^0,z^0)}=s_{pref}'(z^0)\frac{1}{s^0\zeta_{s|z}^c(z^0)}\frac{1}{h_z(z^0)}\int_{z^0}^{\bar{z}}(1-\hat{g}(z))\,dH_z(z)$$

which is the optimality condition (16) presented in Theorem (2).

This optimal tax formula can also be obtained as a direct result through a joint reform of

[20] Applying Lemma 1, earnings and savings changes are here given by

$$\begin{cases}dz=-\frac{z}{1-T_z'}\zeta_z^c(z)\delta T_z^{\theta\prime}-\frac{\eta_z(z)}{1-T_z'}\delta T_z^{\theta}\\ ds=-\frac{s(z)}{1+T_s'}\zeta_{s|z}^c(z)\delta T_s'-\frac{\eta_{s|z}(z)}{1+T_s'}\delta T_s+s_{inc}'(z)dz\end{cases} \tag{96}$$

where the equivalent reform T_z^{θ} is a small increase in the marginal earnings tax by $s_{inc}'(z^*)d\tau_s$ for types $\theta\in[\theta^*,\theta^*+d\theta]$, and a lump sum increase in tax liability by $d\tau_s ds$ for types above. Moreover, the mass of individuals in the bandwitdh is $\delta s\;h_s(s(z^*))=\delta s\,\frac{h_z(z^*)}{s'(z^*)}$.

earnings and savings taxes at income level z^0. Starting from existing tax schedules, consider a *decrease* in marginal earnings tax rates by $d\tau_z$ in the bandwitdh $[z^0, z^0 + dz]$ combined with an *increase* in marginal savings tax rates by $d\tau_s$ in the corresponding bandwidth $[s(z^0), s(z^0 + dz)]$.

Imposing $d\tau_z = s'_{inc}(z^0)d\tau_s$, this joint reform is such that the behavioral effects from earnings changes exactly cancel out (Lemma 1). Moreover, lump-sum changes induced by the two reforms are given by $d\tau_z dz = d\tau_s s'_{inc}(z^0)dz$ and by $d\tau_s ds = d\tau_s s'(z^0)dz$, meaning that the net lump-sum change is proportional to $s'_{pref}(z^0) = s'(z^0) - s'_{inc}(z^0)$.

The total impact of the reform is thus

$$\frac{1}{\lambda}\frac{d\mathcal{L}}{dz} = s'_{pref}(z^0)\int_{z^0}^{\bar{z}}(1-\hat{g}(z))\,d\tau_s\,dH_z(z) - \mathcal{T}'_s\left(s^0,z^0\right)\frac{s^0}{1+\mathcal{T}'_s\left(s^0,z^0\right))}\zeta^c_{s|z}(z^0)d\tau_s\,h_z(z^0) \quad (99)$$

and characterizing the optimum as a situation in which this type of joint reform does not affect the government objective function, $d\mathcal{L} = 0$, yields the result.

B.5.3 Pareto-efficiency condition

We can combine formulas for optimal marginal earnings and savings tax rates (95) and (98) to obtain a characterization of Pareto-efficiency. Indeed, leveraging the previous optimal formula for marginal savings tax rates written in terms of $s'_{pref}(z^*)$, and replacing the integral term by its value given from the optimal formula for marginal earnings tax rates (98) yields

$$\frac{\mathcal{T}'_s(s^0,z^0)}{1+T'_s(s^0,z^0)}s^0\zeta^c_{s|z}(z^0) = s'_{pref}(z^0)\frac{\mathcal{T}'_z(z^0)+s'_{inc}(z^0)\mathcal{T}'_s(s^0,z^0)}{1-\mathcal{T}'_z(s^0,z^0)}z^0\zeta^c_z(z^0)$$

which is the Pareto-efficiency condition (17) presented in Theorem (2).

This Pareto-efficiency condition can also be obtained as a direct result through a joint reform of income and savings tax at income level z^0. Starting from existing tax schedules, consider an *increase* in marginal earnings tax rates by $d\tau_z$ in the bandwitdh $[z^0, z^0 + dz]$ and at the same time a *decrease* in marginal savings tax rates by $\delta\tau_s$ in the corresponding bandwidth $[s(z^0), s(z^0 + dz)]$.

Let construct these reforms such that they leave individuals above $z^0 + dz$ unaffected by setting the two lump-sum changes equal, meaning $d\tau_z dz = d\tau_s ds$ i.e. $d\tau_z = d\tau_s \frac{ds}{dz} = d\tau_s\, s'(z^0)$. Applying Lemma 1, the total impact of the reform on the government's tax revenue as $dz \to 0$ is

$$\begin{aligned}\frac{1}{\lambda}\frac{d\mathcal{L}}{dz} &= \mathcal{T}'_s\left(s^0,z^0\right)\frac{s^0}{1+\mathcal{T}'_s\left(s^0,z^0\right)}\zeta^c_{s|z}(z^0)d\tau_s\,h_z(z^0)\\ &\quad-\Big(\mathcal{T}'_z(s^0,z^0)+s'_{inc}(z^0)\mathcal{T}'_s(s^0,z^0)\Big)\frac{z^0}{1-\mathcal{T}'_z(s^0,z^0)}\zeta^c_z(z^0)d\tau_s\Big(s'(z^0)-s'_{inc}(z^0)\Big)\,h_z(z^0) \quad (100)\end{aligned}$$

while agents well-being remains unchanged. When the impact of this reform is non-zero, the type of joint reform we consider delivers a Pareto-improvement over the existing tax system. Characterizing a Pareto-efficient tax system as one that cannot be reformed in a Pareto-improving way yields the above sufficient statistics characterization.

B.6 Proof of Propositions 2, A.3, and A.4 (Optimal Simple Tax Systems)

For smooth tax system, the sufficient statistic characterization for optimal tax schedules and Pareto-efficiency (Theorem 2) are obtained from local reforms of marginal earnings tax rates and marginal savings tax rates.

Because all types of simple tax systems we consider (SL, SN, and LED) feature a nonlinear earnings tax schedule, such local reforms of marginal earnings tax rates are available. As a result, the derivation of optimal earnings tax formulas for simple tax systems exactly parallels that of general smooth tax systems and the optimal formula for marginal earnings tax rates formula (15) continues to hold. This proves Proposition A.4.

Moreover, SN tax systems also feature a nonlinear savings tax schedule meaning that local reforms of marginal savings tax rates are available as well. As a result, the derivation of the optimal formula for marginal savings tax ratesin SN tax systems exactly parallels that of general smooth tax systems, and the optimal formula (16) continues to hold. This implies that the Pareto-efficiency condition (17) also continues to hold, proving all sufficient statistics characterizations for SN tax systems.

In contrast, local reforms of marginal savings tax rates are not available under SL and LED tax systems which assume a linear savings tax rate. For SL and LED tax systems, we thus derive optimal savings tax formulas to prove Proposition A.3 and Pareto-efficiency formulas to prove Proposition 2.

B.6.1 SL tax system

Impact of a SL tax reform. When the government uses a linear savings tax such that $\mathcal{T}(s,z) = \tau_s\, s + T_z(z)$, we consider a small reform of the linear savings tax rate τ_s that consists in a small increase $d\tau_s$. Formally, this corresponds to a reform $dT_s(s) = d\tau_s\, s$ of the savings tax function $T_s(s) = \tau_s\, s$.

We characterize the impact of this reform on the government objective function. Normalizing all effects by $1/\lambda$, the reform induces

- *mechanical effects:*

$$\int_z \left(1 - \frac{\alpha(z)}{\lambda} U_c'\left(c(z), s(z), z; \vartheta(z)\right)\right) d\tau_s\, s(z)\, dH_z(z) \tag{101}$$

- *behavioral effects from earnings changes*:[21]

$$-\int_z T_z'(z) \left[\frac{z\zeta_z^c(z)}{1 - T_z'(z)}\, d\tau_s\, s'_{inc}(z) + \frac{\eta_z(z)}{1 - T_z'(z)}\, d\tau_s\, s(z)\right] dH_z(z) \tag{103}$$

[21] Applying Lemma 1, earnings and savings changes are here given by

$$\begin{cases} dz = -\frac{z\zeta_z^c(z)}{1-T_z'(z)}\, d\tau_s\, s'_{inc}(z) - \frac{\eta_z(z)}{1-T_z'(z)}\, d\tau_s\, s(z) \\ ds = -\frac{s(z)\zeta_{s|z}^c(z)}{1+\tau_s}\, d\tau_s - \frac{\eta_{s|z}(z)}{1+\tau_s} d\tau_s\, s(z) + s'_{inc}(z) dz \end{cases} \tag{102}$$

- *behavioral effects from savings changes*:

$$-\int_z \tau_s \left[\frac{s(z)\zeta^c_{s|z}(z)}{1+\tau_s} d\tau_s + \frac{\eta_{s|z}(z)}{1+\tau_s} d\tau_s \, s\,(z)\right] dH_z(z) \\ - \int_z \tau_s s'_{inc}(z) \left[\frac{z\zeta^c_z(z)}{1-T'_z} \, d\tau_s \, s'_{inc}(z) + \frac{\eta_z(z)}{1-T'_z} \, d\tau_s \, s\,(z)\right] dH_z(z) \quad (104)$$

Summing over these different effects yields the total impact of the reform

$$\frac{d\mathcal{L}}{\lambda} = \int_z \left\{ (1-\hat{g}(z)) \, s\,(z) - \frac{T'_z(z) + s'_{inc}(z)\tau_s}{1-T'_z\,(z)} z\zeta^c_z(z) \, s'_{inc}(z) - \frac{\tau_s}{1+\tau_s} s(z)\zeta^c_{s|z}(z) \right\} d\tau_s \, dH_z(z), \quad (105)$$

with social marginal welfare weights augmented with the fiscal impact of income effects given by

$$\hat{g}(z) = \frac{\alpha(z)}{\lambda} U'_c\left(c(z), s(z), z; \theta(z)\right) + \frac{T'_z(z)}{1-T'_z\,(z)} \eta_z(z) + \tau_s \left[\frac{\eta_{s|z}(z)}{1+\tau_s} + s'_{inc}(z) \frac{\eta_z(z)}{1-T'_z}\right].$$

Optimal savings tax formula. A direct implication of this result is a sufficient statistics characterization of the optimal linear savings tax schedule τ_s, given an existing and potentially suboptimal nonlinear tax $T_z\,(z)$. Indeed, at the optimum the reform should have a zero impact on the government objective meaning that the optimal savings tax schedule verifies

$$\frac{\tau_s}{1+\tau_s} \int_z s(z)\zeta^c_{s|z}(z) dH_z(z) = \int_z \left\{ (1-\hat{g}(z)) \; s\,(z) - \frac{T'_z(z) + \tau_s \, s'_{inc}(z)}{1-T'_z\,(z)} z\zeta^c_z(z) \, s'_{inc}(z) \right\} dH_z(z), \quad (106)$$

which is Equation (48) in Proposition A.3, and holds for any (potentially suboptimal) earnings tax schedule.

Now, let assume that the earnings tax schedule is optimal and make use of the optimal formula for marginal earnings tax rates. It states that at each earnings z,

$$\frac{T'_z(z) + s'_{inc}(z)\tau_s}{1-T'_z(z)} = \frac{1}{\zeta^c_z(z)} \frac{1}{z h_z(z)} \int_{x \geq z} (1-\hat{g}(x)) \, dH_z(x) \quad (107)$$

such that plugging in this expression to replace the last term, we obtain

$$\frac{\tau_s}{1+\tau_s} \int_z s(z)\zeta^c_{s|z}(z) dH_z(z) = \int_z \left\{ s\,(z) \left[1-\hat{g}(z)\right] \right\} dH_z(z) - \int_z \left\{ s'_{inc}(z) \int_{x \geq z} (1-\hat{g}(x)) \, h_z(x) dx \right\} dz \quad (108)$$

Defining $s_{inc}(z) \equiv \int_{x=0}^{z} s'_{inc}\,(x)\, dx$, we can now integrate by part to rewrite this last term as[22]

$$\int_z \left\{ s'_{inc}(z) \int_{x \geq z} (1-\hat{g}(x)) \, h_z(x) dx \right\} dz = \int_z \left\{ s_{inc}(z)(1-\hat{g}(z)) \, h_z(z) \right\} dz \quad (109)$$

[22]Define $\phi\,(z) = \int_{x=0}^{z} s'_{inc}\,(x)\, dx$ such that $\phi'\,(z) = s'_{inc}\,(z)$ and $\psi\,(z) = \int_{x=z}^{\bar{z}} (1-\hat{g}(x)) \, h_z(x) dx$ such that $\psi'\,(z) = -(1-\hat{g}(z)) \, h_z(z)$ and apply

$$\int_{x=z}^{\bar{z}} \left[\phi'\,(x)\, \psi\,(x)\right] dx = \left[\phi\,(\bar{z})\, \psi\,(\bar{z}) - \phi\,(z)\, \psi\,(z)\right] - \int_{x=z}^{\bar{z}} \left[\phi\,(x)\, \psi'\,(x)\right] dx$$

such that

$$\frac{\tau_s}{1+\tau_s}\int_z s(z)\zeta^c_{s|z}(z)dH_z(z) = \int_z \left\{ \left[s\left(z\right) - s_{inc}(z) \right] \left(1-\hat{g}(z)\right) \right\} dH_z(z). \tag{110}$$

Note that here $\int_z \frac{s(z)}{1+\tau_s}\zeta^c_{s|z}(z)dH_z(z)$ is the aggregate population responses to a change in τ_s. Defining $\overline{\zeta^c_{s|z}}$ as the aggregate elasticity of $\overline{s} := \int_z s(z)dH_z(z)$, we can rewrite this term as $\frac{\overline{s}}{1+\tau_s}\overline{\zeta^c_{s|z}}$ to obtain

$$\frac{\tau_s}{1+\tau_s} = \frac{1}{\overline{s}\overline{\zeta^c_{s|z}}}\int_z s_{pref}\left(z\right)\left(1-\hat{g}(z)\right)dH_z(z). \tag{111}$$

which is Equation (20) in Proposition 2. Integrating by part the right-hand side, this formula is also equivalent to

$$\frac{\tau_s}{1+\tau_s} = \frac{1}{\overline{s}\overline{\zeta^c_{s|z}}}\int_z \left[s'_{pref}\left(z\right) \int_{x\geq z} \left(1-\hat{g}(x)\right) dH_z(x) \right] dz. \tag{112}$$

Pareto-efficiency formula. To derive Pareto-efficiency, we combine savings tax and earnings tax reforms in a way that anihilates all lump-sum changes in tax liability in order to offset all utility changes.

We start with a small reform of the linear savings tax rate τ_s that consists in small increase $d\tau_s$. At the bottom of the earnings distribution ($z = \underline{z}$), the mechanical effect of the savings tax reform is an increase in tax liability by $s\left(\underline{z}\right)d\tau_s$. We thus adjust the earnings tax liability through a downward shift by $s\left(\underline{z}\right)d\tau_s$ at all earnings levels. This joint reform has the following impact on the government objective

$$\frac{d\mathcal{L}}{\lambda} = \int_{z=\underline{z}}^{\overline{z}} \left\{ \left[1-\hat{g}(z)\right]\left[s\left(z\right)-s\left(\underline{z}\right)\right] - \frac{T'_z(z)+\tau_s\, s'_{inc}(z)}{1-T'_z} z\zeta^c_z(z)\, s'_{inc}(z) - \tau_s \frac{s(z)}{1+T'_s}\zeta^c_{s|z}(z) \right\} d\tau_s\, dH_z(z) \tag{113}$$

meaning that the lump-sum change in tax liability is nil at earnings $z = \underline{z}$ but not at earnings above $z \geq \underline{z}$.

To cancel out lump-sum changes in tax liability at all earnings levels, we are going to construct a sequence of earnings tax reforms. We discretize the range of earnings $[\underline{z}, \overline{z}]$ into N bins and consider reforms in the small earnings bandwidths $dz = \frac{\Delta z}{N}$ where $\Delta z = \overline{z} - \underline{z}$. We proceed by induction to pin down a general formula for the tax reforms we consider:

- First, consider a decrease in the marginal earnings tax rate by $d\tau_z = s'\left(\underline{z}\right)d\tau_s$ over the bandwidth $[\underline{z}, \underline{z}+dz]$. In this bandwidth, this reform (i) cancels out lump-sum changes in tax liability to a first-order approximation since $\left[s\left(\underline{z}+dz\right)-s\left(\underline{z}\right)\right]d\tau_s \approx s'\left(\underline{z}\right)dzd\tau_s$, and (ii)

to obtain

$$\begin{aligned}&\int_{x=z}^{\bar{z}} \left[s'_{inc}\left(x\right) \int_{t=x}^{\bar{z}} \left(1-\hat{g}(t)\right) h_z(t)dt \right] dx \\ &= -\left(\int_{x=0}^{z} s'_{inc}\left(x\right)dx \right)\left(\int_{x=z}^{\bar{z}} \left(1-\hat{g}(x)\right) h_z(x)dx \right) + \int_{x=z}^{\bar{z}} \left[\left(\int_{t=0}^{x} s'_{inc}\left(t\right)dt \right)\left(1-\hat{g}(x)\right) h_z(x) \right] dx\end{aligned}$$

where the second term is equal to zero when evaluated at $z = \underline{z}$.

induces earnings responses through the change in marginal tax rates. Moreover, it decreases the lump-sum tax liability on all individuals with earnings $z \geq \underline{z} + dz$. The total impact of this sequence of reforms is then

$$\begin{aligned}\frac{d\mathcal{L}}{\lambda} = &\int_{z=\underline{z}+dz}^{\bar{z}} \left\{ [1-\hat{g}(z)] \left[s(z) - s(\underline{z}) - s'(\underline{z})\, dz \right] \right\} d\tau_s\, dH_z(z) \\ &- \int_{z=\underline{z}}^{\bar{z}} \left\{ \frac{T_z'(z) + \tau_s\, s_{inc}'(z)}{1 - T_z'} z \zeta_z^c(z)\, s_{inc}'(z) - \tau_s \frac{s(z)}{1+\tau_s} \zeta_{s|z}^c(z) \right\} d\tau_s\, dH_z(z) \\ &+ \int_{z=\underline{z}}^{\underline{z}+dz} \frac{T_z'(z) + s_{inc}'(z)\tau_s}{1 - T_z'(z)} z \zeta_z^c(z) \left(s'(\underline{z})\, d\tau_s \right) dH_z(z) \end{aligned} \quad (114)$$

- Second, consider a decrease in the marginal earnings tax rate by $d\tau_z = s'(\underline{z} + dz)\, d\tau_s$ over the bandwidth $[\underline{z} + dz, \underline{z} + 2dz]$, which again cancels out lump-sum changes in this bandwidth up to a first-order approximatince since $[s(\underline{z} + 2dz) - s(\underline{z}) - s'(\underline{z})\, dz] \approx s'(\underline{z} + dz)\, dz$. The total impact of this sequence of reforms is then

$$\begin{aligned}\frac{d\mathcal{L}}{\lambda} = &\int_{z=\underline{z}+2dz}^{\bar{z}} \left\{ [1-\hat{g}(z)] \left[s(z) - s(\underline{z}) - s'(\underline{z})\, dz - s'(\underline{z}+dz)\, dz \right] \right\} d\tau_s\, dH_z(z) \\ &- \int_{z\geq\underline{z}} \left\{ \frac{T_z'(z) + \tau_s\, s_{inc}'(z)}{1 - T_z'} z \zeta_z^c(z)\, s_{inc}'(z) - \tau_s \frac{s(z)}{1+\tau_s} \zeta_{s|z}^c(z) \right\} d\tau_s\, dH_z(z) \\ &+ \int_{z=\underline{z}}^{\underline{z}+dz} \frac{T_z'(z) + s_{inc}'(z)\tau_s}{1 - T_z'(z)} z \zeta_z^c(z) \left(s'(\underline{z})\, d\tau_s \right) dH_z(z) \\ &+ \int_{z=\underline{z}+dz}^{\underline{z}+2dz} \frac{T_z'(z) + s_{inc}'(z)\tau_s}{1 - T_z'(z)} z \zeta_z^c(z) \left(s'(\underline{z}+dz)\, d\tau_s \right) dH_z(z) \end{aligned} \quad (115)$$

- Iterating over to step k, in which we consider a decrease in the marginal earnings tax rate by $d\tau_z = s'\left(\underline{z} + (k-1)\frac{\Delta z}{N}\right) d\tau_s$ over the bandwidth $[\underline{z} + (k-1)\frac{\Delta z}{N}, \underline{z} + k\frac{\Delta z}{N}]$. The total impact of this sequence of reforms is then

$$\begin{aligned}\frac{d\mathcal{L}}{\lambda} = &\int_{z=\underline{z}+k\frac{\Delta z}{N}}^{\bar{z}} \left\{ [1-\hat{g}(z)] \left[s(z) - s(\underline{z}) - \frac{\Delta z}{N} \left[\sum_{p=0}^{k-1} s'\left(\underline{z} + p\frac{\Delta z}{N} \right) \right] \right] \right\} d\tau_s\, dH_z(z) \\ &- \int_{z=\underline{z}}^{\bar{z}} \left\{ \frac{T_z'(z) + \tau_s\, s_{inc}'(z)}{1 - T_z'} z \zeta_z^c(z)\, s_{inc}'(z) - \tau_s \frac{s(z)}{1+\tau_s} \zeta_{s|z}^c(z) \right\} d\tau_s\, dH_z(z) \\ &+ \sum_{p=0}^{k-1} \int_{z=\underline{z}+p\frac{\Delta z}{N}}^{\underline{z}+(p+1)\frac{\Delta z}{N}} \frac{T_z'(z) + s_{inc}'(z)\tau_s}{1 - T_z'(z)} z \zeta_z^c(z) s'\left(\underline{z} + p\frac{\Delta z}{N} \right) d\tau_s\, dH_z(z) \end{aligned} \quad (116)$$

- Pushing the iteration forward until $k = N$, the first integral disappears (integration over an

empty set) such that the total impact of this sequence of reforms is given by

$$\frac{d\mathcal{L}}{\lambda} = -\int_{z=\underline{z}}^{\bar{z}} \left\{ \frac{T_z'(z) + \tau_s\, s_{inc}'(z)}{1 - T_z'} z\zeta_z^c(z)\, s_{inc}'(z) - \tau_s \frac{s(z)}{1+\tau_s} \zeta_{s|z}^c(z) \right\} d\tau_s\, dH_z(z) + \sum_{p=0}^{N-1} \int_{z=\underline{z}+p\frac{\Delta z}{N}}^{\underline{z}+(p+1)\frac{\Delta z}{N}} \frac{T_z'(z) + s_{inc}'(z)\tau_s}{1 - T_z'(z)} z\zeta_z^c(z) s'\left(\underline{z} + p\frac{\Delta z}{N}\right) d\tau_s\, dH_z(z) \quad (117)$$

Let's now compute the last term at the limit $N \to \infty$. We have

$$\sum_{p=0}^{N-1} \int_{z=\underline{z}+p\frac{\Delta z}{N}}^{\underline{z}+(p+1)\frac{\Delta z}{N}} \frac{T_z'(z) + s_{inc}'(z)\tau_s}{1 - T_z'(z)} z\zeta_z^c(z) s'\left(\underline{z} + p\frac{\Delta z}{N}\right) d\tau_s\, dH_z(z) \quad (118)$$

$$\approx \sum_{p=0}^{N-1} \frac{T_z'(\underline{z} + p\frac{\Delta z}{N}) + s_{inc}'(\underline{z} + p\frac{\Delta z}{N})\tau_s}{1 - T_z'\left(\underline{z} + p\frac{\Delta z}{N}\right)} \left(\underline{z} + p\frac{\Delta z}{N}\right) \zeta_z^c\left(\underline{z} + p\frac{\Delta z}{N}\right) s'\left(\underline{z} + p\frac{\Delta z}{N}\right) d\tau_s h_z\left(\underline{z} + p\frac{\Delta z}{N}\right) \frac{\Delta z}{N} \quad (119)$$

$$\underset{N\to\infty}{\longrightarrow} \int_{z=\underline{z}}^{\bar{z}} \frac{T_z'(z) + s_{inc}'(z)\tau_s}{1 - T_z'(z)} z\zeta_z^c(z)\, s'(z)\, d\tau_s\, h_z(z)\, dz \quad (120)$$

where the last line follows from the Riemann definition of the integral in terms of Riemann sums. Hence, the total impact of this sequence of reforms is at the limit given by

$$\frac{d\mathcal{L}}{\lambda} = -\int_{z=\underline{z}}^{\bar{z}} \left\{ \frac{T_z'(z) + \tau_s\, s_{inc}'(z)}{1 - T_z'} z\zeta_z^c(z)\, s_{inc}'(z) - \tau_s \frac{s(z)}{1+\tau_s} \zeta_{s|z}^c(z) \right\} d\tau_s\, h_z(z) dz + \int_{z=\underline{z}}^{\bar{z}} \frac{T_z'(z) + s_{inc}'(z)\tau_s}{1 - T_z'(z)} z\zeta_z^c(z)\, s'(z)\, d\tau_s\, h_z(z)\, dz \quad (121)$$

By construction, the sequence of reforms we have constructed does not affect agents' utility, and only affects tax revenue through the expression above. When the impact of this reform is non-zero, the type of sequence of reforms we consider delivers a Pareto-improvement over the existing tax system. Characterizing a Pareto-efficient tax system as one that cannot be reformed in a Pareto-improving way yields the following sufficient statistics characterization:

$$\frac{\tau_s}{1+\tau_s} \int_{z=\underline{z}}^{\bar{z}} s(z) \zeta_{s|z}^c(z) h_z(z) dz = \int_{z=\underline{z}}^{\bar{z}} \left[s'(z) - s_{inc}'(z)\right] \frac{T_z'(z) + s_{inc}'(z)\tau_s}{1 - T_z'(z)} z\zeta_z^c(z)\, h_z(z)\, dz, \quad (122)$$

which is Equation (24) in Proposition 2.

B.6.2 Linear earnings-dependent savings tax

Earnings responses to LED savings tax reform. We begin by translating Lemma 1 for LED tax systems, where we consider particular savings tax reforms.

For a LED tax system, we consider savings tax reform $dT_s(s|z)$ that consists in a small increase $d\tau_s$ of the linear savings tax rate phased-in over the earnings bandwidth $[z^0, z^0 + dz]$. Formally,

this corresponds to the following reform $dT_s(s|z)$ of the savings tax function $T_s(s|z) = \tau_s(z)\, s$:

$$dT_s(s|z) = \begin{cases} 0 & if\ z \leq z^0 \\ d\tau_s\left(z - z^0\right) s & if\ z \in [z^0, z^0 + dz] \\ d\tau_s dz\, s & if\ z \geq z^0 + dz \end{cases} \tag{123}$$

Now, let

$$V\left(T_s(.|z), y, z; \theta\right) = \max_s\ U\left(y - s - T_s\left(s|z\right), s, z; \theta\right) \tag{124}$$

be agent θ indirect utility function at earnings z and (pre savings tax) disposable income y.

The idea is to construct for each type θ a perturbation of the earnings tax $dT_z^\theta(z)$ that induces the same earnings response as the initial perturbation $dT_s(s|z)$. Suppose we define this perturbation for each type θ such that at all earnings z,

$$V(T_s(.|z) + dT_s(.|z), z - T_z(z), z; \theta) = V(T_s(.|z), z - T_z(z) - dT_z^\theta(z), z; \theta) \tag{125}$$

then, by construction, dT_z^θ induces the same earnings response as the initial perturbation dT_s. Moreover, dT_z^θ induces the same utility change as the initial perturbation dT_s.

Applying the envelope theorems to agents in the earnings bandwidth $[z^0, z^0 + dz]$ yields

$$\begin{aligned} & U_c' \cdot d\tau_s\left(z - z^0\right) s\left(z; \theta\right) = U_c' \cdot dT_z^\theta(z) \\ \iff & dT_z^\theta(z) = d\tau_s\left(z - z^0\right) s\left(z; \theta\right) \end{aligned}$$

As can thus be seen differentiating both sides by earnings z, and letting $dz \to 0$, a small increase by $d\tau_s$ in the linear savings tax rate at earnings z^0 leads to the same earnings responses as a small increase by $s\left(z; \theta\right) d\tau_s$ in the marginal earnings tax rate.

Applying the envelope theorems to agents above the earnings bandwidth $[z^0, z^0 + dz]$ yields

$$dT_z^\theta(z) = d\tau_s dz\, s\left(z; \theta\right)$$

such that the equivalent earnings tax reform combines a $d\tau_s dz\, s\left(z\right)$ increase in tax liability and a $d\tau_s dz\, s'_{inc}\left(z\right)$ increase in the marginal earnings tax rate.

Impact of a LED savings tax reform. We consider savings tax reform $dT_s(s|z)$ that consists in a small increase $d\tau_s$ of the linear savings tax rate phased-in over the earnings bandwidth $[z^0, z^0 + dz]$. We characterize the impact of this reform on the government objective function as $dz \to 0$. Normalizing all effects by $1/\lambda$, the reform induces

- *mechanical effects:*

$$\int_{z \geq z^0} \left(1 - \frac{\alpha(z)}{\lambda} U_c'\left(c(z), s(z), z; \theta(z)\right)\right) d\tau_s dz\, s\left(z\right)\, dH_z(z) \tag{126}$$

- *behavioral effects from earnings changes*:[23]

$$
\begin{aligned}
&- \mathcal{T}_z'\left(s\left(z^0\right), z^0\right)\left[\frac{z^0\zeta_z^c(z^0)}{1-\mathcal{T}_z'\left(s\left(z^0\right), z^0\right)}\, d\tau_s s(z^0)\right] h_z(z^0)dz \\
&- \int_{z\geq z^0} \mathcal{T}_z'\left(s\left(z\right), z\right)\left[\frac{z\zeta_z^c(z)}{1-\mathcal{T}_z'\left(s\left(z\right), z\right)}\, s'_{inc}(z) + \frac{\eta_z(z)}{1-\mathcal{T}_z'\left(s\left(z\right), z\right)}\, s(z)\right] d\tau_s dz\, dH_z(z) \qquad (127)
\end{aligned}
$$

- *behavioral effects from savings changes*:

$$
\begin{aligned}
&- \tau_s(z^0)s'_{inc}(z^0)\left[\frac{z^0\zeta_z^c(z^0)}{1-\mathcal{T}_z'\left(s\left(z^0\right), z^0\right)}\, d\tau_s\, s(z^0)\right] h_z(z^0)dz \\
&- \int_{z\geq z^0} \tau_s(z)\left[\frac{\zeta^c_{s|z}(z)+\eta_{s|z}(z)}{1+\tau_s(z)}\, s(z) + s'_{inc}(z)\left(\frac{z\zeta_z^c(z)}{1-\mathcal{T}_z'\left(s\left(z\right), z\right)} s'_{inc}(z) + \frac{\eta_z(z)}{1-\mathcal{T}_z'\left(s\left(z\right), z\right)}\, s(z)\right)\right] d\tau_s dz\, dH_z(z \qquad (128)
\end{aligned}
$$

Summing over these different effects yields the total impact of the reform

$$
\begin{aligned}
\frac{1}{\lambda}\frac{d\mathcal{L}}{dz} = &\int_{z\geq z^0}\left\{(1-\hat{g}(z))\, s\left(z\right) - \frac{\mathcal{T}_z'\left(s\left(z\right), z\right) + s'_{inc}(z)\tau_s(z)}{1-\mathcal{T}_z'\left(s\left(z\right), z\right)} z\zeta_z^c(z)s'_{inc}(z) - \frac{\tau_s(z)}{1+\tau_s(z)} s(z)\zeta^c_{s|z}(z)\right\} d\tau_s\, dH_z(z) \\
&- \frac{\mathcal{T}_z'\left(z^0, s\left(z^0\right)\right) + s'_{inc}(z^0)\tau_s(z^0)}{1-\mathcal{T}_z'\left(s\left(z^0\right), z^0\right)} z^0\zeta_z^c(z^0)\, s(z^0) d\tau_s\, h_z(z^0) \qquad (129)
\end{aligned}
$$

with

$$
\hat{g}(z) = \frac{\alpha(z)}{\lambda} U_c'\left(c(z), s(z), z; \theta(z)\right) - \frac{\mathcal{T}_z'\left(s\left(z\right), z\right) + s'_{inc}(z)\tau_s(z)}{1-\mathcal{T}_z'\left(s\left(z\right), z\right)}\eta_z(z) - \frac{\tau_s(z)}{1+\tau_s(z)}\eta_{s|z}(z).
$$

Optimal LED savings tax formula. A direct implication of this result is a sufficient statistics characterization of the optimal linear earnings-dependent savings tax $\tau_s(z)$, given an existing and potentially suboptimal nonlinear earnings tax $T_z\left(z\right)$. Indeed, at the optimum the reform should have a zero impact on the government objective, meaning that the optimal linear earnings-dependent savings tax schedule verifies at each earnings z^0

$$
\begin{aligned}
&\frac{\mathcal{T}_z'\left(s\left(z^0\right), z^0\right) + s'_{inc}(z^0)\tau_s(z^0)}{1-\mathcal{T}_z'\left(s\left(z^0\right), z^0\right)} z^0\zeta_z^c(z^0)\, s(z^0)\, h_z(z^0) \qquad (130) \\
&= \int_{z\geq z^0}\left\{(1-\hat{g}(z))\, s\left(z\right) - \frac{\mathcal{T}_z'\left(s\left(z\right), z\right) + s'_{inc}(z)\tau_s(z)}{1-\mathcal{T}_z'\left(s\left(z\right), z\right)} z\zeta_z^c(z)s'_{inc}(z) - \frac{\tau_s(z)}{1+\tau_s(z)} s(z)\zeta^c_{s|z}(z)\right\} dH_z(z),
\end{aligned}
$$

which is Equation (47) in Proposition A.3, and holds for any (potentially suboptimal) earnings tax schedule.

[23] Applying Lemma 1, earnings and savings changes at earnings z^0 and at earnings above z^0 are respectively given by

$$
\begin{cases} dz = -\frac{z^0\zeta_z^c(z^0)}{1-\mathcal{T}_z'} d\tau_s\, s(z^0) \\ ds = s'_{inc}(z^0)dz \end{cases} \quad \text{and} \quad \begin{cases} dz = -\frac{z\zeta_z^c(z)}{1-\mathcal{T}_z'} d\tau_s dz\, s'_{inc}(z) - \frac{\eta_z(z)}{1-\mathcal{T}_z'} d\tau_s dz\, s\left(z\right) \\ ds = -\frac{s(z)\zeta^c_{s|z}(z)}{1+\tau_s(z)} d\tau_s dz - \frac{\eta_s(z)}{1+\tau_s(z)} d\tau_s dz\ s\left(z\right) + s'_{inc}(z)dz \end{cases}
$$

Now, let assume that the earnings tax schedule is optimal and make use of the optimal formula for marginal earnings tax rates. It states that at each earnings z,

$$\frac{T_z'\left(z^0\right)+\tau_s'\left(z^0\right)s\left(z^0\right)+s_{inc}'(z^0)\tau_s\left(z^0\right)}{1-T_z'\left(z^0\right)-\tau_s'\left(z^0\right)s\left(z^0\right)}\zeta_z^c(z^0)z^0h_z(z^0)=\int_{z\geq z^0}(1-\hat{g}(z))\,dH_z(z) \tag{131}$$

such that replacing the first term in the first equation we get

$$\begin{aligned}&s(z^0)\int_{z\geq z^0}(1-\hat{g}(z))\,dH_z(z)+\int_{z\geq z^0}\frac{\tau_s(z)}{1+\tau_s(z)}s(z)\zeta_{s|z}^c(z)\,dH_z(z) \\ &=\int_{z\geq z^0}\left\{(1-\hat{g}(z))\,s\left(z\right)-\frac{T_z'\left(z\right)+\tau_s'\left(z\right)s\left(z\right)+s_{inc}'(z)\tau_s(z)}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s\left(z\right)}z\zeta_z^c(z)s_{inc}'(z)\right\}dH_z(z)\end{aligned} \tag{132}$$

Differentiating with respect to z^0 yields

$$\begin{aligned}&s'(z^0)\int_{z\geq z^0}(1-\hat{g}(z))\,dH_z(z)-s(z^0)(1-\hat{g}(z^0))h_z(z^0)-\frac{\tau_s(z^0)}{1+\tau_s(z^0)}s(z^0)\zeta_{s|z}^c(z^0)h_z(z^0) \\ &=-\left(1-\hat{g}(z^0)\right)s\left(z^0\right)h_z(z^0)+s_{inc}'(z^0)\frac{T_z'\left(z^0\right)+\tau_s'\left(z^0\right)s\left(z^0\right)+s_{inc}'(z^0)\tau_s\left(z^0\right)}{1-T_z'\left(z^0\right)-\tau_s'\left(z^0\right)s\left(z^0\right)}\zeta_z^c(z^0)z^0h_z(z^0)\end{aligned}$$

now both $s(z^0)(1-\hat{g}(z^0))h_z(z^0)$ cancel out such that replacing the last term we obtain

$$\frac{\tau_s(z^0)}{1+\tau_s(z^0)}s(z^0)\zeta_{s|z}^c(z^0)h_z(z^0)=s_{pref}'(z^0)\int_{z\geq z^0}(1-\hat{g}(z))\,dH_z(z) \tag{133}$$

which is Equation (23) in Proposition 2.

Pareto-efficiency formula for LED tax systems. To derive Pareto-efficiency, we combine savings tax and earnings tax reforms in a way that anihilates all lump-sum changes in tax liability in order to offset all utility changes. To do so, consider the following sequence of reforms.

First, consider an *increase* $d\tau_s$ of the income-contingent savings tax $\tau_s(z)$ by phased-in over the earnings bandwidth $[z^0, z^0+dz]$, together with a *decrease* $d\tau_z$ of the marginal tax rate on earnings in the same bandwidth. Imposing $d\tau_z = s(z^0)d\tau_s$, this joint reform cancels out substitution effects at z^0 and, as $dz\to 0$, it has the following impact on the government objective:

$$\begin{aligned}\frac{1}{\lambda}\frac{d\mathcal{L}}{dz}=&\int_{z\geq z^0}\left\{(1-\hat{g}(z))\,s\left(z\right)-\frac{\mathcal{T}_z'\left(s\left(z\right),z\right)+s_{inc}'(z)\tau_s(z)}{1-\mathcal{T}_z'\left(s\left(z\right),z\right)}z\zeta_z^c(z)s_{inc}'(z)-\frac{\tau_s(z)}{1+\tau_s(z)}s(z)\zeta_{s|z}^c(z)\right\}d\tau_s\,dH_z(z) \\ &-\int_{z\geq z^0}(1-\hat{g}(z))\;s(z^0)d\tau_s\,dH_z(z)\end{aligned} \tag{134}$$

Second, consider a *decrease* by $d\tau_s$ of the income-contingent savings tax $\tau_s(z)$ phased-in over the earnings bandwidth $[z^0+dz, z^0+2dz]$, together with an *increase* $d\tau_z$ of the marginal tax rate on earnings in the same bandwidth. Imposing $d\tau_z = s(z^0+dz)d\tau_s$, this joint reform cancels out

substitution effects at z^0 and, as $dz \to 0$, it has the following impact on the government objective:

$$\frac{1}{\lambda}\frac{d\mathcal{L}}{dz} = -\int_{z\geq z^0+dz}\left\{(1-\hat{g}(z))\,s\,(z) - \frac{\mathcal{T}_z'\,(s\,(z)\,,z) + s'_{inc}(z)\tau_s(z)}{1-\mathcal{T}_z'\,(s\,(z)\,,z)} z\zeta_z^c(z)s'_{inc}(z) - \frac{\tau_s(z)}{1+\tau_s(z)}s(z)\zeta^c_{s|z}(z)\right\} d\tau_s\, dH_z(z)$$
$$+ \int_{z\geq z^0+dz}(1-\hat{g}(z))\ s(z^0+dz)d\tau_s\, dH_z(z) \tag{135}$$

The total impact of these two joint reforms on the objective function of the government is thus

$$\frac{1}{\lambda}\frac{d\mathcal{L}}{dz} = \int_{z^0}^{z^0+dz}\left\{(1-\hat{g}(z))\,s\,(z) - \frac{\mathcal{T}_z'\,(s\,(z)\,,z) + s'_{inc}(z)\tau_s(z)}{1-\mathcal{T}_z'\,(s\,(z)\,,z)} z\zeta_z^c(z)s'_{inc}(z) - \frac{\tau_s(z)}{1+\tau_s(z)}s(z)\zeta^c_{s|z}(z)\right\} d\tau_s dH_z(z)$$
$$+ \left[s(z^0+dz)-s(z^0)\right]\int_{z^0}^{\bar{z}}(1-\hat{g}(z))\,d\tau_s dH_z(z) - s(z^0)\int_{z^0}^{z^0+dz}(1-\hat{g}(z))\,d\tau_s dH_z(z) \tag{136}$$

where, as $dz \to 0$, the term $(1-\hat{g}(z))\,s\,(z)$ in the integral on the first line excatly cancels out with the last term on the second line.

Third, we consider a *decrease* by $d\tau_z$ of the marginal tax rate on earnings in the bandwidth $[z^0, z^0+dz]$. Imposing $d\tau_z = \left[s(z^0+dz)-s(z^0)\right]d\tau_s$, this reform cancels out the other integral term on the second line and generates earnings responses at the margin. As a result, the total impact of this sequence of reforms is at the limit $dz \to 0$ given by

$$\frac{1}{\lambda}\frac{d\mathcal{L}}{dz} = \left\{-s'_{inc}(z^0)\frac{\mathcal{T}_z'\left(s\left(z^0\right),z^0\right) + s'_{inc}(z^0)\tau_s(z^0)}{1-\mathcal{T}_z'\left(s\left(z^0\right),z^0\right)} z^0\zeta_z^c(z^0) - \frac{\tau_s(z^0)}{1+\tau_s(z^0)}s(z^0)\zeta^c_{s|z}(z^0)\right\} d\tau_s\, h_z(z^0)dz$$
$$+ s'(z^0)\frac{\mathcal{T}_z'\left(s\left(z^0\right),z^0\right) + s'_{inc}(z^0)\tau_s(z^0)}{1-\mathcal{T}_z'\left(s\left(z^0\right),z^0\right)} z^0\zeta_z^c(z^0)d\tau_s\, h_z(z^0)dz \tag{137}$$

where the $s'(z^*)$ in the last line comes from the fact that at the limit $dz \to 0$, we have $\left[s(z^0+dz)-s(z^0)\right]d\tau_s \to s'(z^0)\,dz\,d\tau_s$. This expression is the impact of this sequence of reforms on the government's tax revenue, while agents well-being remains unchanged. When this impact is non-zero, the type of reforms we consider delivers a Pareto-improvement over the existing tax system. Characterizing a Pareto-efficient tax system as one that cannot be reformed in a Pareto-improving way yields the following sufficient statistics characterization:

$$\frac{\tau_s(z^0)}{1+\tau_s(z^0)}s(z^0)\zeta^c_{s|z}(z^0) = \left[s'(z^0) - s'_{inc}(z^0)\right]\frac{\mathcal{T}_z'\left(s\left(z^0\right),z^0\right) + s'_{inc}(z^0)\tau_s(z^0)}{1-\mathcal{T}_z'\left(s\left(z^0\right),z^0\right)} z^0\zeta_z^c(z^0), \tag{138}$$

which is Equation (26) in Proposition 2.

B.7 Proof of Proposition 3 (Simple Tax systems and Multidimensional Heterogeneity)

We characterize in Proposition 3 optimal savings tax formulas for each type of simple tax system in the presence of multidimensional heterogeneity. These formulas take the actual earnings tax

schedule as given, be they optimally set or not, and extend the results derived in the unidimensional case. Crucially, we are able to provide similar characterizations because Lemma 1 still holds in the presence of multidimensional heterogeneity.

B.7.1 Separable linear (SL) tax system

Consider a reform that consists in a $d\tau_s$ increase in the linear savings tax rate τ_s. For all agents, this triggers an increase in tax liability by $s\,d\tau_s$ and an increase in the marginal savings tax rate by $d\tau_s$ – which by Lemma 1 produces earnings responses equivalent to an increase in the marginal earnings tax rate by $s'_{inc}d\tau_s$.

This reform has the following effects on the government objective:

- mechanical effects

$$\int_z \int_s \Big[(1-g(s,z))\, s\, d\tau_s\Big] h(s,z)\, ds dz \tag{139}$$

$$= \int_z \mathbb{E}\left[(1-g(s,z))\, s|z\right]\, d\tau_s h_z(z)\, dz \tag{140}$$

- earnings responses[24]

$$\int_z T'_z(z) \left\{ \int_s \left(-\frac{z}{1-T'_z(z)} \zeta^c_z(s,z)\, s'_{inc}(s,z)\, d\tau_s - \frac{\eta_z(s,z)}{1-T'_z(z)} s\, d\tau_s \right) h(s,z)\, ds \right\} dz \tag{141}$$

$$= -\int_z \frac{T'_z(z)}{1-T'_z(z)} \left\{ \mathbb{E}\left[z\zeta^c_z(s,z) s'_{inc}(s,z) + \eta_z(s,z) s|z \right] \right\} d\tau_s h_z(z)\, dz \tag{142}$$

- savings responses[25]

$$\tau_s \int_z \int_s \Bigg\{ -\frac{s}{1+\tau_s} \zeta^c_{s|z}(s,z)\, d\tau_s - \frac{\eta_{s|z}(s,z)}{1+\tau_s}\, s\, d\tau_s \tag{143}$$

$$+ s'_{inc}(s,z) \left(-\frac{z}{1-T'_z(z)} \zeta^c_z(s,z)\, s'_{inc}(s,z)\, d\tau_s - \frac{\eta_z(s,z)}{1-T'_z(z)} s\, d\tau_s \right) \Bigg\} h(s,z)\, ds dz$$

$$= -\tau_s \int_z \Bigg\{ \frac{1}{1+\tau_s} \mathbb{E}\left[s\zeta^c_{s|z}(s,z) + \eta_{s|z}(s,z)\, s|z \right] \tag{144}$$

$$+ \frac{1}{1-T'_z(z)} \left(\mathbb{E}\left[z\zeta^c_z(s,z) \left(s'_{inc}(s,z)\right)^2 + \eta_z(s,z)\, s\, s'_{inc}(s,z)\, |z \right] \right) \Bigg\} d\tau_s h_z(z)\, dz$$

[24] We have

$$dz = \frac{\partial z}{\partial T'_z}\delta T'_z + \frac{\partial z}{\partial T_z}\delta T_z = -\frac{z}{1-T'_z}\zeta^c_z\, \delta T'_z - \frac{\eta_z}{1-T'_z}\delta T_z$$

[25] We have

$$ds = \frac{\partial s}{\partial T'_s}\delta T'_s + \frac{\partial s}{\partial T_s}\delta T_s + \frac{ds(z;\theta)}{dz}dz = -\frac{s}{1+T'_s}\zeta^c_{s|z}\, \delta T'_s - \frac{\eta_{s|z}}{1+T'_s}\delta T_s + s'_{inc}dz$$

Such that the total impact of the reform on the government objective is

$$\frac{d\mathcal{L}}{d\tau_s} = \int_z \mathbb{E}\left[\left(1-g\left(s,z\right)\right)s|z\right]\, h_z\left(z\right)dz \tag{145}$$

$$- \int_z \frac{T_z'\left(z\right)}{1-T_z'\left(z\right)}\left\{\mathbb{E}\left[z\zeta_z^c(s,z)s_{inc}'\left(s,z\right)+\eta_z(s,z)\,s|z\right]\right\}h_z\left(z\right)dz \tag{146}$$

$$- \tau_s\int_z\left\{\frac{1}{1+\tau_s}\mathbb{E}\left[s\zeta_{s|z}^c(s,z)+\eta_{s|z}(s,z)\,s|z\right]\right.$$

$$\left.+\frac{1}{1-T_z'\left(z\right)}\left(\mathbb{E}\left[z\zeta_z^c(s,z)\,\left(s_{inc}'\left(s,z\right)\right)^2+\eta_z(s,z)\,s\,s_{inc}'\left(s,z\right)|z\right]\right)\right\}h_z\left(z\right)dz \tag{147}$$

Redefining augmented social marginal welfare weights as

$$\hat{g}(s,z) = g(s,z) + \frac{T_z'(z)}{1-T_z'\left(z\right)}\eta_z(s,z) + \frac{\tau_s}{1+\tau_s}\eta_{s|z}(s,z) + \frac{\tau_s}{1-T_z'\left(z\right)}\eta_z(s,z)s_{inc}'(s,z) \tag{148}$$

we finally get

$$\frac{d\mathcal{L}}{d\tau_s} = \int_z \mathbb{E}\left[\left(1-\hat{g}\left(s,z\right)\right)s|z\right]\, h_z\left(z\right)dz - \int_z \frac{T_z'\left(z\right)}{1-T_z'\left(z\right)}\left\{\mathbb{E}\left[z\zeta_z^c(s,z)s_{inc}'\left(s,z\right)|z\right]\right\}h_z\left(z\right)dz \tag{149}$$

$$- \tau_s\int_z\left\{\frac{1}{1+\tau_s}\mathbb{E}\left[s\zeta_{s|z}^c(s,z)|z\right]+\frac{1}{1-T_z'\left(z\right)}\left(\mathbb{E}\left[z\zeta_z^c(s,z)\,\left(s_{inc}'\left(s,z\right)\right)^2|z\right]\right)\right\}h_z\left(z\right)dz$$

The optimal savings linear tax rate τ_s thus satisfies

$$\frac{\tau_s}{1+\tau_s}\int_z\left\{\mathbb{E}\left[s\zeta_{s|z}^c(s,z)|z\right]\right\}dH_z\left(z\right) \tag{150}$$

$$= \int_z\left\{\mathbb{E}\left[\left(1-\hat{g}\left(s,z\right)\right)s|z\right] - \mathbb{E}\left[\frac{T_z'\left(z\right)+s_{inc}'\left(s,z\right)\tau_s}{1-T_z'\left(z\right)}z\zeta_z^c(s,z)s_{inc}'\left(s,z\right)|z\right]\right\}dH_z\left(z\right)$$

B.7.2 Separable nonlinear (SN) tax system

Consider a reform that consists in a small $\delta\tau_s$ increase in the marginal savings tax rate across the savings bandwidth $\left[s^0, s^0+\delta s\right]$.[26] For all agents with savings above s^0, this triggers a $\delta s\,\delta\tau_s$ increase in tax liability. For agents at s^0, this triggers a $\delta\tau_s$ increase in the marginal savings tax rate – which by Lemma 1 produces earnings responses equivalent to a $s_{inc}'\delta\tau_s$ increase in the marginal earnings tax rate.

This reform has the following effects on the government objective:

- mechanical effects

$$\int_{s\geq s^0}\int_z\left\{\left(1-g\left(s,z\right)\right)\delta s\delta\tau_s\right\}h\left(s,z\right)dsdz \tag{151}$$

$$= \int_z\left\{\mathbb{E}\left[1-g\left(s,z\right)|z,s\geq s^0\right]\right\}\delta s\delta\tau_s h_z\left(z\right)dz$$

[26] To avoid any ambiguity in this multidimensional context, we here use d for integration and δ for attributes of the reform we consider.

- earnings responses[27]

$$
\begin{aligned}
&-\int_z T_z'(z)\left\{\frac{z}{1-T_z'}\zeta_z^c\left(s^0,z\right)s_{inc}'\left(s^0,z\right)\delta\tau_s\right\}\delta s\,h\left(s^0,z\right)dz \qquad (152)\\
&-\int_{s\geq s^0}\int_z T_z'(z)\left\{\frac{\eta_z(s,z)}{1-T_z'(z)}\,\delta\tau_s\delta s\right\}h(s,z)\,dsdz\\
&=-\int_z\frac{T_z'(z)}{1-T_z'}\left\{z\zeta_z^c\left(s^0,z\right)s_{inc}'\left(s^0,z\right)+\mathbb{E}\left[\eta_z(s,z)s|z,s\geq s^0\right]\right\}\delta\tau_s\delta s\,h_z(z)\,dz
\end{aligned}
$$

- savings responses[28]

$$
\begin{aligned}
&-T_s'\left(s^0\right)\int_z\left\{\frac{s^0}{1+T_s'\left(s^0\right)}\zeta_{s|z}^c(s^0,z)\,\delta\tau_s+s_{inc}'\left(s^0,z\right)\frac{z}{1-T_z'(z)}\zeta_z^c(s^0,z)\,s_{inc}'\left(s^0,z\right)\delta\tau_s\right\}\delta s\,h\left(s^0,z\right)dz \qquad (153)\\
&-\int_{s\geq s^0}\int_z\left\{T_s'(s)\left(\frac{\eta_{s|z}(s,z)}{1+T_s'(s)}\delta s\delta\tau_s+s_{inc}'(s,z)\,\frac{\eta_z(s,z)}{1-T_z'(z)}\delta s\delta\tau_s\right)\right\}h(s,z)\,dsdz\\
&=-\int_z\left\{\frac{T_s'\left(s^0\right)}{1+T_s'\left(s^0\right)}s^0\zeta_{s|z}^c(s^0,z)+\frac{T_s'\left(s^0\right)}{1-T_z'(z)}s_{inc}'\left(s^0,z\right)^2z\zeta_z^c(s^0,z)\right\}\delta\tau_s\delta s\,h_z(z)\,dz \qquad (154)\\
&-\int_z\left\{\mathbb{E}\left[\frac{T_s'(s)}{1+T_s'(s)}\eta_{s|z}(s,z)\middle|z,s\geq s^0\right]+\mathbb{E}\left[s_{inc}'(s,z)\,\frac{T_s'(s)}{1-T_z'(z)}\eta_z(s,z)\middle|z,s\geq s^0\right]\right\}\delta s\delta\tau_s h_z(z)\,dz
\end{aligned}
$$

Such that the total impact of the reform on the government objective is

$$
\begin{aligned}
\frac{d\mathcal{L}}{\delta s\delta\tau_s}&=\int_z\left\{\mathbb{E}\left[1-g(s,z)\,|z,s\geq s^0\right]\right\}dH_z(z) \qquad (155)\\
&-\int_z\frac{T_z'(z)}{1-T_z'}\left\{z\zeta_z^c\left(s^0,z\right)s_{inc}'\left(s^0,z\right)+\mathbb{E}\left[\eta_z(s,z)|z,s\geq s^0\right]\right\}dH_z(z)\\
&-\int_z\left\{\frac{T_s'\left(s^0\right)}{1+T_s'\left(s^0\right)}s^0\zeta_{s|z}^c(s^0,z)+\frac{T_s'\left(s^0\right)}{1-T_z'(z)}s_{inc}'\left(s^0,z\right)^2z\zeta_z^c(s^0,z)\right\}dH_z(z)\\
&-\int_z\left\{\mathbb{E}\left[\frac{T_s'(s)}{1+T_s'(s)}\eta_{s|z}(s,z)\middle|z,s\geq s^0\right]+\mathbb{E}\left[s_{inc}'(s,z)\,\frac{T_s'(s)}{1-T_z'(z)}\eta_z(s,z)\middle|z,s\geq s^0\right]\right\}dH_z(z)
\end{aligned}
$$

Redefining augmented social marginal welfare weights as

$$
\hat{g}(s,z)=g(s,z)+\frac{T_z'(z)}{1-T_z'(z)}\eta_z(s,z)+\frac{T_s'(s)}{1+T_s'(s)}\eta_{s|z}(s,z)+s_{inc}'(s,z)\frac{T_s'(s)}{1-T_z'(z)}\eta_z(s,z) \qquad (156)
$$

[27] We have

$$
dz=\frac{\partial z}{\partial T_z'}\delta T_z'+\frac{\partial z}{\partial T_z}\delta T_z=-\frac{z}{1-T_z'}\zeta_z^c(z)\,\delta T_z'-\frac{\eta_z(z)}{1-T_z'}\delta T_z
$$

[28] We have

$$
ds=\frac{\partial s}{\partial T_s'}\delta T_s'+\frac{\partial s}{\partial T_s}\delta T_s+\frac{ds(z;\theta)}{dz}dz=-\frac{s}{1+T_s'}\zeta_{s|z}^c\,\delta T_s'-\frac{\eta_{s|z}(z)}{1+T_s'}\delta T_s+s_{inc}'(z)dz
$$

we finally get

$$\frac{d\mathcal{L}}{\delta s \delta \tau_s} = \int_z \left\{ \mathbb{E}\left[1 - \hat{g}\left(s,z\right)|z, s \geq s^0\right] \right\} dH_z\left(z\right) - \int_z \left\{ \frac{T_z'\left(z\right)}{1-T_z'} z \zeta_z^c\left(s^0,z\right) s_{inc}'\left(s^0,z\right) \right\} dH_z\left(z\right) \quad (157)$$
$$- \int_z \left\{ \frac{T_s'\left(s^0\right)}{1+T_s'\left(s^0\right)} s^0 \zeta_{s|z}^c(s^0,z) + \frac{T_s'\left(s^0\right)}{1-T_z'\left(z\right)} s_{inc}'\left(s^0,z\right)^2 z \zeta_z^c(s^0,z) \right\} dH_z\left(z\right)$$

The optimal marginal savings tax rate $T_s'\left(.\right)$ thus satisfies at each savings s^0,

$$\frac{T_s'\left(s^0\right)}{1+T_s'\left(s^0\right)} \int_z \left\{ s^0 \zeta_{s|z}^c(s^0,z) \right\} dH_z\left(z\right) \quad (158)$$
$$= \int_z \left\{ \mathbb{E}\left[1 - \hat{g}\left(s,z\right)|z, s \geq s^0\right] \right\} dH_z\left(z\right) - \int_z \left\{ \frac{T_z'\left(z\right) + s_{inc}'\left(s^0,z\right) T_s'\left(s^0\right)}{1-T_z'\left(z\right)} z \zeta_z^c\left(s^0,z\right) s_{inc}'\left(s^0,z\right) \right\} dH_z\left(z\right)$$

which is equation (52) in Proposition (A.4).

B.7.3 Linear earnings-dependent (LED) tax system

Consider a reform that consists in a $\delta\tau_s \delta z$ increase in the linear savings tax rate phased-in across the earnings bandwidth $\left[z^0, z^0 + \delta z\right]$.[29]

For all agents with earnings above z^0, this triggers an increase in the linear savings tax rate by $\delta\tau_s\delta z$ meaning that the marginal savings tax rate increases by the same magnitude – which triggers by Lemma 1 earnings responses equivalent to those induced by a $s_{inc}'\delta\tau_s\delta z$ increase in the marginal earnings tax rate – and that agents' tax liability increases by $s\,\delta\tau_s\delta z$.

For agents at z^0, the only direct effect of the reform is to induce earnings responses which by Lemma 1 are equivalent to an increase in the marginal earnings tax rate given by $s\,\delta\tau_s$ (as in the unidimensional case).

This reform has the following effects on the government objective:

- mechanical effects[30]

$$\int_{z \geq z^0} \int_s \left\{ \left(1 - g\left(s,z\right)\right) \delta z \delta \tau_s s \right\} h\left(s,z\right) ds dz \quad (159)$$
$$= \int_{z \geq z^0} \left\{ E_s\left[\left(1 - g\left(s,z\right)\right) s \middle| z \right] \delta z \delta \tau_s \right\} h_z\left(z\right) dz$$

[29] To avoid any ambiguity, we here use d for integration and δ for attributes of the reform we consider.

[30] We use

$$h\left(s,z\right) = h\left(s|z\right) h_z\left(z\right)$$

- earnings responses[31]

$$-\int_s \left(T_z'\left(z^0\right)+\tau_s'\left(z^0\right)s\right)\left\{\frac{z^0}{1-T_z'\left(z^0\right)-\tau_s'\left(z^0\right)s}\zeta_z^c(s,z^0)\,s\,\delta\tau_s\right\}\delta z\,h\left(s,z^0\right)ds \tag{160}$$

$$-\int_{z\geq z^0}\int_s \left(T_z'\left(z\right)+\tau_s'\left(z\right)s\right)\left\{\frac{z}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s}\zeta_z^c(s,z)s'_{inc}\left(s,z\right)\delta z\delta\tau_s+\frac{\eta_z(s,z)}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s}\,s\,\delta z\delta\tau_s\right\}h\left(s,\right.$$

$$=-\mathbb{E}\left[\frac{T_z'\left(z\right)+\tau_s'\left(z\right)s}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s}z\zeta_z^c(s,z)s\middle|z=z^0\right]\delta z\delta\tau_s\,h_z\left(z^0\right) \tag{161}$$

$$-\int_{z\geq z^0}\left\{\mathbb{E}\left[\frac{T_z'\left(z\right)+\tau_s'\left(z\right)s}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s}\Big(z\zeta_z^c(s,z)s'_{inc}\left(s,z\right)+s\,\eta_z(s,z)\Big)\middle|z\right]\right\}\delta z\delta\tau_s h_z\left(z\right)dz$$

- savings responses[32]

$$-\tau_s\left(z^0\right)\int_s s'_{inc}\left(s,z^0\right)\left\{\frac{z^0}{1-T_z'\left(z^0\right)-\tau_s'\left(z^0\right)s}\zeta_z^c(s,z^0)\,s\,\delta\tau_s\right\}\delta z\,h\left(s,z^0\right)ds \tag{162}$$

$$-\int_{z\geq z^0}\int_s \tau_s\left(z\right)\left\{\frac{s}{1+\tau_s\left(z\right)}\zeta_{s|z}^c(s,z)\,\delta z\delta\tau_s+s'_{inc}\left(s,z\right)\frac{z}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s}\zeta_z^c(s,z)\,s'_{inc}\left(s,z\right)\delta z\delta\tau_s\right\}h\left(s,z\right)dz$$

$$-\int_{z\geq z^0}\int_s\left\{\tau_s\left(z\right)\left(\frac{\eta_{s|z}(s,z)}{1+\tau_s\left(z\right)}\,s\,\delta z\delta\tau_s+s'_{inc}\left(s,z\right)\frac{\eta_z(s,z)}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s}\,s\,\delta z\delta\tau_s\right)\right\}h\left(s,z\right)dsdz$$

$$=-\tau_s\left(z^0\right)\mathbb{E}\left[s'_{inc}\left(s,z\right)\frac{z\zeta_z^c(s,z)s}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s}\middle|z=z^0\right]\delta z\delta\tau_s\,h_z\left(z^0\right) \tag{163}$$

$$-\int_{z\geq z^0}\left\{\mathbb{E}\left[\frac{\tau_s\left(z\right)}{1+\tau_s\left(z\right)}s\zeta_{s|z}^c(s,z)\middle|z\right]+\mathbb{E}\left[\frac{\tau_s\left(z\right)}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s}z\zeta_z^c(s,z)s'_{inc}\left(s,z\right)^2\right]\right\}\delta z\delta\tau_s\,h_z\left(z\right)dz$$

$$-\int_{z\geq z^0}\left\{\mathbb{E}\left[\frac{\tau_s\left(z\right)}{1+\tau_s\left(z\right)}s\eta_{s|z}(s,z)\middle|z\right]+\mathbb{E}\left[\frac{\tau_s\left(z\right)}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s}s\eta_z(s,z)s'_{inc}\left(s,z\right)\middle|z\right]\right\}\delta z\delta\tau_s\,h_z\left(z\right)dz$$

Such that the total impact of the reform on the government objective is

[31] We have

$$dz=\frac{\partial z}{\partial\mathcal{T}_z'}\delta\mathcal{T}_z'+\frac{\partial z}{\partial\mathcal{T}_z}\delta\mathcal{T}_z=-\frac{z}{1-\mathcal{T}_z'}\zeta_z^c\,\delta\mathcal{T}_z'-\frac{\eta_z}{1-\mathcal{T}_z'}\delta\mathcal{T}_z$$

[32] We have

$$ds=\frac{\partial s}{\partial T_s'}\delta T_s'+\frac{\partial s}{\partial T_s}\delta T_s+\frac{ds(z;\theta)}{dz}dz=-\frac{s}{1+T_s'}\zeta_{s|z}^c\,\delta T_s'-\frac{\eta_{s|z}}{1+T_s'}\delta T_s+s'_{inc}(z)dz$$

$$\frac{d\mathcal{L}}{\delta s \delta \tau_s} = \int_{z\geq z^0} \Big\{ E_s\Big[(1-g\left(s,z\right))\, s \Big| z\Big] \Big\} h_z\left(z\right) dz \tag{164}$$
$$- \mathbb{E}\left[\frac{T_z'\left(z\right)+\tau_s'\left(z\right)s}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s} z\zeta_z^c(s,z)s \Big| z=z^0 \right] h_z\left(z^0\right)$$
$$- \int_{z\geq z^0} \left\{ \mathbb{E}\left[\frac{T_z'\left(z\right)+\tau_s'\left(z\right)s}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s} \Big(z\zeta_z^c(s,z)s_{inc}'\left(s,z\right) + s\,\eta_z(s,z) \Big) \Big| z \right] \right\} h_z\left(z\right) dz$$
$$- \mathbb{E}\left[\frac{s_{inc}'\left(s,z\right)\tau_s\left(z\right)}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s} z\zeta_z^c(s,z)s \Big| z=z^0 \right] \delta z \delta\tau_s\, h_z\left(z^0\right)$$
$$- \int_{z\geq z^0} \left\{ \mathbb{E}\left[\frac{\tau_s\left(z\right)}{1+\tau_s\left(z\right)} s\zeta_{s|z}^c(s,z) \Big| z \right] + \mathbb{E}\left[\frac{s_{inc}'\left(s,z\right)\tau_s\left(z\right)}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s} z\zeta_z^c(s,z)s_{inc}'\left(s,z\right) \Big| z \right] \right\} h_z\left(z\right) dz \tag{165}$$
$$- \int_{z\geq z^0} \left\{ \mathbb{E}\left[\frac{\tau_s\left(z\right)}{1+\tau_s\left(z\right)} s\eta_{s|z}(s,z) \Big| z \right] + \mathbb{E}\left[\frac{s_{inc}'\left(s,z\right)\tau_s\left(z\right)}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s} s\eta_z(s,z) \Big| z \right] \right\} h_z\left(z\right) dz$$

Redefining augmented social marginal welfare weights as

$$\hat{g}(s,z) = g(s,z) + \frac{T_z'\left(z\right)+\tau_s'\left(z\right)s + s_{inc}'(s,z)\tau_s\left(z\right)}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s}\eta_z(s,z) + \frac{\tau_s\left(z\right)}{1+\tau_s\left(z\right)}\eta_{s|z}(s,z) \tag{166}$$

we finally get

$$\frac{d\mathcal{L}}{\delta s \delta \tau_s} = \int_{z\geq z^0} \mathbb{E}\Big[(1-\hat{g}\left(s,z\right))\, s \Big| z\Big] h_z\left(z\right) dz \tag{167}$$
$$- \mathbb{E}\left[\frac{T_z'\left(z\right)+\tau_s'\left(z\right)s + s_{inc}'\left(s,z\right)\tau_s\left(z\right)}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s} z\zeta_z^c(s,z)s \Big| z=z^0 \right] h_z\left(z^0\right)$$
$$- \int_{z\geq z^0} \mathbb{E}\left[\frac{T_z'\left(z\right)+\tau_s'\left(z\right)s + s_{inc}'\left(s,z\right)\tau_s\left(z\right)}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s} z\zeta_z^c(s,z)s_{inc}'\left(s,z\right) \Big) \Big| z \right] h_z\left(z\right) dz$$
$$- \int_{z\geq z^0} \mathbb{E}\left[\frac{\tau_s\left(z\right)}{1+\tau_s\left(z\right)} s\zeta_{s|z}^c(s,z) \Big| z \right] h_z\left(z\right) dz$$

The optimal linear earnings-dependent savings tax rate $\tau_s\left(.\right)$ satisfies at each earnings z^0,

$$\mathbb{E}\left[\frac{T_z'\left(z\right)+\tau_s'\left(z\right)s + s_{inc}'\left(s,z\right)\tau_s\left(z\right)}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s} z\zeta_z^c(s,z)s \Big| z=z^0 \right] h_z\left(z^0\right) + \int_{z\geq z^0} \mathbb{E}\left[\frac{\tau_s\left(z\right)}{1+\tau_s\left(z\right)} s\zeta_{s|z}^c(s,z) \Big| z \right] h_z\left(z\right) dz$$
$$= \int_{z\geq z^0} \mathbb{E}\Big[(1-\hat{g}\left(s,z\right))\, s \Big| z\Big] h_z\left(z\right) dz - \int_{z\geq z^0} \mathbb{E}\left[\frac{T_z'\left(z\right)+\tau_s'\left(z\right)s + s_{inc}'\left(s,z\right)\tau_s\left(z\right)}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s} z\zeta_z^c(s,z)s_{inc}'\left(s,z\right) \Big| z \right] h_z\left(z\right) dz \tag{168}$$

which is equation (53) in Proposition (A.4).

B.8 Proof of Proposition 4 (Many Goods)

B.8.1 Setting and definitions

The problem of the government is to maximize the following Lagrangian

$$\mathcal{L} = \int_z \left\{ \alpha(z)\, U\Big(z - \mathcal{T}(\boldsymbol{s}(z), z) - \sum_{i=1}^{n} s_i(z), \boldsymbol{s}(z), z; \vartheta\,(z)\Big) + \lambda \mathcal{T}(\boldsymbol{s}(z), z) - E \right\} dH_z(z). \tag{169}$$

where we use the fact that $z\,(\theta)$ is a bijective mapping to denote $\vartheta\,(z)$ its inverse to denote Pareto-weights $\alpha(z) := \alpha(\vartheta(z))$ and the vector of n consumption goods as $\boldsymbol{s}\,(z) := \boldsymbol{s}\,(z; \vartheta\,(z))$.

Note that when characterizing optimal marginal tax rates on good i, we use the existence of a mapping between s_i consumption and earnings z, such that we implicitly assume that consumption patterns are strictly monotonic. That is, for each good i, we assume that $s_i'(z) > 0$ for normal goods, and that $s_i'(z) < 0$ for inferior goods.

In this setting, we are going to express optimal tax formulas of the following elasticity concepts that measure consumption responses of s_i and s_j to changes in $\mathcal{T}'_{s_i}$

$$\zeta^c_{s_i|z}\,(z(\theta)) := -\frac{1 + \mathcal{T}'_{s_i}\,(\boldsymbol{s}\,(z;\theta)\,, z)}{s_i\,(z;\theta)} \frac{\partial s_i\,(z;\theta)}{\partial \mathcal{T}'_{s_i}\,(\boldsymbol{s}\,(z;\theta)\,, z)}\Big|_{z=z(\theta)} \tag{170}$$

$$\xi^c_{s_{j,i}|z}\,(z(\theta)) := \frac{\mathcal{T}'_{s_i}\,(\boldsymbol{s}\,(z;\theta)\,, z)}{s_j\,(z;\theta)} \frac{\partial s_j\,(z;\theta)}{\partial \mathcal{T}'_{s_i}\,(\boldsymbol{s}\,(z;\theta)\,, z)}\Big|_{z=z(\theta)} \tag{171}$$

and in terms of the following statistics

$$s'_{i,inc}(z(\theta)) := \frac{\partial s_i(z;\theta)}{\partial z}\Big|_{z=z(\theta)} \tag{172}$$

$$\hat{g}(z(\theta)) := \left[\alpha(z) \frac{U_c'(z)}{\lambda} - \left(\mathcal{T}'_z\,(\boldsymbol{s}(z), z) + \sum_{i=1}^{n} s'_{i,inc}(z) \mathcal{T}'_{s_i}\,(\boldsymbol{s}(z), z) \right) \frac{\partial z(.)}{\partial \mathcal{T}} - \sum_{i=1}^{n} \mathcal{T}'_{s_i}\,(\boldsymbol{s}(z), z)\, \frac{\partial s_i(.)}{\partial \mathcal{T}} \right]\Big|_{z=z(\theta)}. \tag{173}$$

B.8.2 Optimal marginal tax rates on earnings z

We consider a small reform at earnings level z^0 that consists in a small increase $d\tau_z$ of the marginal earnings tax rate $\mathcal{T}'_z$ in a small bandwidth dz. The impact of this reform on the Lagrangian as $dz \to 0$ is

$$\begin{aligned} \frac{1}{\lambda} \frac{d\mathcal{L}}{d\tau_z dz} = & \int_{x=z^0}^{\bar{z}} \Big(1 - \alpha(x) \frac{U_c'(x)}{\lambda}\Big) dH_z(x) \\ & + \mathcal{T}'_z\,\big(\boldsymbol{s}(z^0), z^0\big)\, \frac{\partial z(.)}{\partial \mathcal{T}'_z}\Big|_{z=z^0} h_z(z^0) + \int_{x=z_0}^{\bar{z}} \mathcal{T}'_z\,(\boldsymbol{s}(x), x)\, \frac{\partial z(.)}{\partial \mathcal{T}}\Big|_{z=x} dH_z(x) \\ & + \sum_{i=1}^{n} \mathcal{T}'_{s_i}\,\big(\boldsymbol{s}(z^0), z^0\big)\, s'_{i,inc}(z^0) \frac{\partial z(.)}{\partial \mathcal{T}'_z}\Big|_{z=z^0} h_z(z^0) + \int_{x=z_0}^{\bar{z}} \sum_{i=1}^{n} \mathcal{T}'_{s_i}\,(\boldsymbol{s}(x), x) \left[\frac{\partial s_i(.)}{\partial \mathcal{T}}\Big|_{z=x} + s'_{i,inc}(x) \frac{\partial z(.)}{\partial \mathcal{T}}\Big|_{z=x} \right] dH \end{aligned} \tag{174}$$

We characterize optimal taxes through $d\mathcal{L} = 0$. Plugging in social marginal welfare weights augmented with the fiscal impacts of income effects $\hat{g}(z)$, we obtain

$$-\left[\mathcal{T}_z'\left(\boldsymbol{s}(z^0), z^0\right) + \sum_{i=1}^{n} \mathcal{T}_{s_i}'\left(\boldsymbol{s}(z^0), z^0\right) s_{i,inc}'(z^0)\right] \frac{\partial z(.)}{\partial \mathcal{T}_z'}\Big|_{z=z^0} h_z(z_0) = \int_{x=z_0}^{\bar{z}} \Big(1 - \hat{g}(x)\Big) dH_z(x) \quad (175)$$

B.8.3 Optimal marginal tax rates on good i

Intuitively, an increase in the marginal tax rate on good i induces (a) substitution effects away from s_i, (b) labor supply distortions on earnings z, and, new to this setting, (c) cross-effects on the consumption of goods s_{-i}.

We consider a small reform at $s_i^0 = s_i(z^0)$ that consists in a small increase $d\tau_s$ of the marginal tax rate $\mathcal{T}_{s_i}'$ in a small bandwidth ds. The impact of this reform on the Lagrangian as $ds \to 0$ is

$$\begin{aligned}
\frac{1}{\lambda}\frac{d\mathcal{L}}{d\tau_z dz} = &\int_{x=z^0}^{\bar{z}} \Big(1 - \alpha(x)\frac{U_c'(x)}{\lambda}\Big) dH_z(x) \\
&+ \mathcal{T}_z'\left(\boldsymbol{s}(z^0), z^0\right) \frac{\partial z(.)}{\partial \mathcal{T}_{s_i}'}\Big|_{z=z^0} \frac{h_z(z^0)}{s_i'(z^0)} + \int_{x=z^0}^{\bar{z}} \mathcal{T}_z'\left(\boldsymbol{s}(x), x\right) \frac{\partial z(.)}{\partial \mathcal{T}}\Big|_{z=x} dH_z(z) \\
&+ \sum_{j=1}^{n} \mathcal{T}_{s_j}'\left(\boldsymbol{s}(z^0), z^0\right) \left[\frac{\partial s_j(.)}{\partial \mathcal{T}_{s_i}'}\Big|_{z=z^0} + s_{j,inc}'(z^0) \frac{\partial z(.)}{\partial \mathcal{T}_{s_i}'}\Big|_{z=z^0}\right] \frac{h_z(z^0)}{s_i'(z^0)} \\
&+ \int_{x=z^0}^{\bar{z}} \sum_{j=1}^{n} \mathcal{T}_{s_j}'\left(\boldsymbol{s}(x), x\right) \left[\frac{\partial s_j(.)}{\partial \mathcal{T}}\Big|_{z=x} + s_{j,inc}'(x) \frac{\partial z(.)}{\partial \mathcal{T}}\Big|_{z=x}\right] dH_z(z)
\end{aligned} \quad (176)$$

where $\frac{\partial s_j(.)}{\partial \mathcal{T}_{s_i}'}$ capture cross-effects for all $j \neq i$.

We characterize optimal taxes through $d\mathcal{L} = 0$. Plugging in social marginal welfare weights augmented with the fiscal impacts of income effects $\hat{g}(x)$, and using the fact that by Lemma 1 (which still holds in this setting) $\frac{\partial z(.)}{\partial \mathcal{T}_{s_i}'} = s_{i,inc}' \frac{\partial z(.)}{\partial \mathcal{T}_z'}$, we obtain

$$\begin{aligned}
&-\left[\left(\mathcal{T}_z'\left(\boldsymbol{s}(z^0), z^0\right) + \sum_{j=1}^{n} s_{j,inc}'(z^0) \mathcal{T}_{s_j}'\left(\boldsymbol{s}(z^0), z^0\right)\right) s_{i,inc}'(z^0) \frac{\partial z(.)}{\partial \mathcal{T}_z'}\Big|_{z=z^0} + \sum_{j=1}^{n} \mathcal{T}_{s_j}'\left(\boldsymbol{s}(z^0), z^0\right) \frac{\partial s_j(.)}{\partial \mathcal{T}_{s_i}'}\Big|_{z=z^0}\right] h_z(z^0) \\
&= s_i'(z^0) \int_{x=z^0}^{\bar{z}} \Big(1 - \hat{g}(x)\Big) dH_z(x)
\end{aligned} \quad (177)$$

B.8.4 Deriving Proposition 4

For any good i, we combine the optimality condition for marginal tax rates on earnings z with the one for i to obtain

$$-\sum_{j=1}^{n} \mathcal{T}_{s_j}'\left(z^0\right) \frac{\partial s_j(.)}{\partial \mathcal{T}_{s_i}'}\Big|_{z=z^0} = \frac{1}{h_z(z^0)} \Big(s_i'(z^0) - s_{i,inc}'(z^0)\Big) \int_{x=z^0}^{\bar{z}} \Big(1 - \hat{g}(x)\Big) dH_z(x). \quad (178)$$

Isolating the term relative to $\mathcal{T}'_{s_i}\left(z^0\right)$ on the left-hand side yields the following optimal tax formula in terms of $s'_{i,pref}$

$$-\mathcal{T}'_{s_i}\left(z^0\right)\frac{\partial s_i(.)}{\partial \mathcal{T}'_{s_i}}\Big|_{z=z^0}=\frac{1}{h_z(z^0)}s'_{i,pref}(z^0)\int_{x=z^0}^{\bar{z}}\Big(1-\hat{g}(x)\Big)dH_z(x)+\sum_{j\neq i}\mathcal{T}'_{s_j}\left(z^0\right)\frac{\partial s_j(.)}{\partial \mathcal{T}'_{s_i}}\Big|_{z=z^0} \quad (179)$$

where $\frac{\partial s_j(.)}{\partial \mathcal{T}'_{s_i}}$ capture cross-effects for all $j\neq i$.

We can rewrite this optimality condition in terms of the compensated elasticity $\zeta^c_{s_i|z}$ and the cross elasticity $\xi^c_{s_{j,i}|z}$ to finally obtain

$$\begin{aligned}\frac{\mathcal{T}'_{s_i}(z^0)}{1+\mathcal{T}'_{s_i}(z^0)}=&\frac{1}{s_i(z^0)\zeta^c_{s_i|z}(z^0)}\frac{1}{h_z(z^0)}s'_{i,pref}(z^0)\int_{z=z^0}^{\bar{z}}\Big(1-\hat{g}(z)\Big)dH_z(z)\\&+\sum_{j\neq i}\frac{\mathcal{T}'_{s_j}(z^0)}{\mathcal{T}'_{s_i}(z^0)}\frac{s_j(z^0)\xi^c_{s_{j,i}|z}(z^0)}{s_i(z^0)\zeta^c_{s_i|z}(z^0)}\end{aligned} \quad (180)$$

which is the first condition stated in Proposition 4.

To derive the second condition stated in Proposition 4, we again combine the optimality condition for marginal tax rates on earnings z with the one for i eliminating this time the integral terms. This yields

$$\sum_{j=1}^{n}\mathcal{T}'_{s_j}\left(z^0\right)\frac{\partial s_j(.)}{\partial \mathcal{T}'_{s_i}}\Big|_{z=z^0}=\Big(s'_i(z^0)-s'_{i,inc}(z^0)\Big)\left(\mathcal{T}'_z(z^0)+\sum_{j=1}^{n}s'_{j,inc}(z^0)\mathcal{T}'_{s_j}(z^0)\right)\frac{\partial z(.)}{\partial \mathcal{T}'_z}\Big|_{z=z^0}. \quad (181)$$

Isolating the term relative to $\mathcal{T}'_{s_i}\left(z^0\right)$ on the left-hand side yields the following Pareto-efficiency condition in terms of $s'_{i,pref}$

$$\mathcal{T}'_{s_i}\left(z^0\right)\frac{\partial s_i(.)}{\partial \mathcal{T}'_{s_i}}\Big|_{z=z^0}=s'_{i,pref}(z^0)\left(\mathcal{T}'_z(z^0)+\sum_{j=1}^{n}s'_{j,inc}(z^0)\mathcal{T}'_{s_j}(z^0)\right)\frac{\partial z(.)}{\partial \mathcal{T}'_z}-\sum_{j\neq i}\mathcal{T}'_{s_j}\left(z^0\right)\frac{\partial s_j(.)}{\partial \mathcal{T}'_{s_i}}\Big|_{z=z^0} \quad (182)$$

where $\frac{\partial s_j(.)}{\partial \mathcal{T}'_{s_i}}$ capture cross-effects for all $j\neq i$.

We can rewrite this optimality condition in terms of the compensated elasticity $\zeta^c_{s_i|z}$ and the cross elasticity $\xi^c_{s_{j,i}|z}$ to finally obtain

$$\begin{aligned}\frac{\mathcal{T}'_{s_i}(z^0)}{1+\mathcal{T}'_{s_i}(z^0)}=&s'_{i,pref}(z^0)\frac{\mathcal{T}'_z(z^0)+\sum_{j=1}^{n}\mathcal{T}'_{s_j}(z^0)s'_{j,inc}(z^0)}{1-\mathcal{T}'_z(z^0)}\frac{z^0\zeta^c_z(z^0)}{s_i(z^0)\zeta^c_{s_i|z}(z^0)}\\&+\sum_{j\neq i}\frac{\mathcal{T}'_{s_j}(z^0)}{\mathcal{T}'_{s_i}(z^0)}\frac{s_j(z^0)\xi^c_{s_{j,i}|z}(z^0)}{s_i(z^0)\zeta^c_{s_i|z}(z^0)}\end{aligned} \quad (183)$$

This completes the proof of Proposition 4.

B.9 Proof of Proposition 5 (Bequest Taxation and Behavioral Agents)

We here provide a sufficient statistics characterization of a smooth tax system $\mathcal{T}(s,z)$ under the following additively separable representation of agents' preferences

$$U\big(c,s,z;\theta\big) = u\left(c;\theta\right) - k\left(z;\theta\right) + \beta\left(\theta\right) v\left(s;\theta\right),$$

and for a utilitarian government that maximizes

$$\int_{\theta}\Big[U\left(c\left(\theta\right),s\left(\theta\right),z\left(\theta\right);\theta\right) + \nu\left(\theta\right) v\left(s\left(\theta\right);\theta\right)\Big] dF\left(\theta\right), \tag{184}$$

where $\nu(\theta)$ parametrizes the degree of misalignment on the valuation of s.

Using the mapping between types θ and earnings z, the Lagrangian of the problem writes

$$\mathcal{L} = \int_{z}\Big[U\left(c\left(z\right),s\left(z\right),z;\vartheta\left(z\right)\right) + \nu\left(z\right) v\left(s\left(z\right);\vartheta\left(z\right)\right) + \lambda\left(\mathcal{T}\left(s,z\right) - E\right)\Big] dH_{z}\left(z\right), \tag{185}$$

As before, we derive optimal tax formulas by considering reforms of marginal tax rates on z and s. Thanks to the additively separable representation of preferences, there are no income effects on labor supply choices. As a result, the only substantial change is that savings changes now lead to changes in social welfare proportional to the degree of misalignment.

Optimal marginal tax rates on z. A small reform at earnings z^0 that consists in a small increase $d\tau_z$ of the marginal earnings tax rate in a small bandwidth dz has the following effect as $dz \to 0$,

$$\begin{aligned}\frac{1}{\lambda}\frac{d\mathcal{L}}{d\tau_z dz} = &\int_{z^0}^{\bar{z}} \left(1-\hat{g}(z)\right) dH_z(z) \\ &- \left(\mathcal{T}_z'\left(s\left(z^0\right),z^0\right) + s'_{inc}(z^0)\Big(\mathcal{T}_s'\left(s(z^0),z^0\right) + \nu\left(z^0\right)\frac{v'\left(s\left(z^0\right)\right)}{\lambda}\Big)\right)\frac{z^0}{1-\mathcal{T}_z'\left(s\left(z^0\right),z^0\right)}\zeta_z^c(z^0)\, h_z(z^0).\end{aligned}$$

In this context, social marginal welfare weights augmented with income effects $\hat{g}(z)$ are equal to

$$\hat{g}(z) = \frac{u'\left(c\left(z\right)\right)}{\lambda} + \Big(\mathcal{T}_s'\left(s(z),z\right) + \nu\left(z\right)\frac{v'\left(s\left(z\right)\right)}{\lambda}\Big)\frac{\eta_{s|z}(z)}{1+\mathcal{T}_s'\left(s(z),z\right)}$$

and we can use agents' first-order condition for s, $\left(1+\mathcal{T}_s'\right)u'\left(c\right) = \beta v'\left(s\right)$, to express the misalignment wedge in terms of the social marginal welfare weights $g\left(z\right) := \frac{u'(c(z))}{\lambda}$ as

$$\nu\left(z\right)\frac{v'\left(s\left(z\right)\right)}{\lambda} = \frac{\nu\left(z\right)}{\beta\left(z\right)} g(z)\left(1+\mathcal{T}_s'\right).$$

The optimal schedule of marginal earnings tax rates is thus characterized by

$$\frac{\mathcal{T}_z'\left(s\left(z^0\right),z^0\right)}{1-\mathcal{T}_z'\left(s\left(z^0\right),z^0\right)}=\frac{1}{\zeta_z^c(z^0)}\frac{1}{z^0h_z(z^0)}\int_{z=z^0}^{\bar{z}}(1-\hat{g}(z))\,dH_z(z)$$
$$-s'_{inc}(z^0)\frac{\mathcal{T}_s'\left(s(z^0),z^0\right)}{1-\mathcal{T}_z'\left(s\left(z^0\right),z^0\right)}-s'_{inc}(z^0)\frac{\nu\left(z^0\right)}{\beta\left(z^0\right)}g(z^0)\frac{1+\mathcal{T}_s'\left(s(z^0),z^0\right)}{1-\mathcal{T}_z'\left(s\left(z^0\right),z^0\right)}$$

Optimal marginal tax rates on s. A small reform at $s^0=s(z^0)$ that consists in a small increase $d\tau_s$ of the marginal tax rate on s in a small bandwidth ds has the following effect as $ds\to 0$,

$$\frac{1}{\lambda}\frac{d\mathcal{L}}{d\tau_s ds}=s'\left(z^0\right)\int_{z^0}^{\bar{z}}(1-\hat{g}(z))\,dH_z(z)$$
$$-\left\{\left(\mathcal{T}_s'\left(s(z^0),z^0\right)+\nu\left(z^0\right)\frac{v'\left(s\left(z^0\right)\right)}{\lambda}\right)\frac{s(z^0)}{1+\mathcal{T}_s'\left(s\left(z^0\right),z^0\right)}\zeta_{s|z}^c(z^0)+\right.$$
$$\left.\left[\mathcal{T}_z'\left(s\left(z^0\right),z^0\right)+s'_{inc}(z^0)\left(\mathcal{T}_s'\left(s(z^0),z^0\right)+\nu\left(z^0\right)\frac{v'\left(s\left(z^0\right)\right)}{\lambda}\right)\right]\frac{z^0}{1-\mathcal{T}_z'\left(s\left(z^0\right),z^0\right)}\zeta_z^c(z^0)s'_{inc}(z^0)\right\}h_z(z^0)$$

Replacing the misalignment wedge by its expression in terms of social marginal welfare weights $g\left(z\right)$, we obtain that the optimal schedule of marginal tax rates on s is characterized by

$$\begin{aligned}&\frac{\mathcal{T}_s'\left(s(z^0),z^0\right)}{1+\mathcal{T}_s'\left(s\left(z^0\right),z^0\right)}\zeta_{s|z}^c(z^0)s(z^0)h_z(z^0)\\&=s'\left(z^0\right)\int_{z^0}^{\bar{z}}(1-\hat{g}(z))\,dH_z(z)-s'_{inc}(z^0)\frac{\mathcal{T}_z'\left(s\left(z^0\right),z^0\right)+s'_{inc}(z^0)\mathcal{T}_s'\left(s(z^0),z^0\right)}{1-\mathcal{T}_z'\left(s\left(z^0\right),z^0\right)}\zeta_z^c(z^0)z^0h_z(z^0)\\&-\frac{\nu\left(z^0\right)}{\beta\left(z^0\right)}g(z^0)\left[\zeta_{s|z}^c(z^0)s(z^0)+\left(s'_{inc}(z^0)\right)^2\frac{1+\mathcal{T}_s'\left(s(z^0),z^0\right)}{1-\mathcal{T}_z'\left(s\left(z^0\right),z^0\right)}\zeta_z^c(z^0)z^0\right]h_z(z^0)\end{aligned}\tag{186}$$

Plugging in the formula for optimal marginal earnings tax rates, we finally obtain

$$\frac{\mathcal{T}_s'\left(s(z^0),z^0\right)}{1+\mathcal{T}_s'\left(s\left(z^0\right),z^0\right)}=s'_{pref}\left(z^0\right)\frac{1}{\zeta_{s|z}^c(z^0)}\frac{1}{s(z^0)h_z(z^0)}\int_{z=z^0}^{\bar{z}}(1-\hat{g}(z))\,dH_z(z)-\frac{\nu\left(z^0\right)}{\beta\left(z^0\right)}g(z^0)$$

which is the first optimality condition in Proposition (5).

Pareto-efficiency condition. Combining optimal formulas for earnings and savings (or bequests) taxes to eliminate the integral term yields the following Pareto-efficiency condition

$$\begin{aligned}&\frac{\mathcal{T}_s'\left(s(z^0),z^0\right)}{1+\mathcal{T}_s'\left(s\left(z^0\right),z^0\right)}+\frac{\nu\left(z\right)}{\beta\left(z\right)}g(z)\\&=s'_{pref}\left(z^0\right)\frac{\zeta_z^c(z^0)z^0}{\zeta_{s|z}^c(z^0)s(z^0)}\left[\frac{\mathcal{T}_z'\left(s\left(z^0\right),z^0\right)+s'_{inc}(z^0)\mathcal{T}_s'\left(s(z^0),z^0\right)}{1-\mathcal{T}_z'\left(s\left(z^0\right),z^0\right)}+s'_{inc}(z^0)\frac{\nu\left(z^0\right)}{\beta\left(z^0\right)}g(z^0)\right]\end{aligned}$$

which is the second optimality condition in Proposition (5).

B.10 Proof of Proposition 6 (Multi-Dimensional Tax Range with Heterogeneous Prices)

B.10.1 Setting

We consider heterogeneous marginal rates of transformation or "prices" $p(z,\theta)$ between c and s, and a two-part tax structure, where a person must pay a tax $T_1(z)$ in units of c and a tax $T_2(s,z)$ in units of s.

In this setting, we can write agent θ problem as

$$\max_{c,s,z} U(c,s,z;\theta) \text{ s.t. } c + p(z,\theta)s \leq z - T_1(z) - p(z,\theta)T_2(s,z) \tag{187}$$

$$\iff \max_{z} \left\{ \max_{s} U\Big(z - T_1(z) - p(z,\theta)\left(s + T_2(s,z)\right), s, z; \theta\Big)\right\}. \tag{188}$$

where the inner problem leads to consumption choices $c(z;\theta)$ and $s(z;\theta)$, while the outter problem leads to an earnings choice $z(\theta)$. Maintaining the assumption that $z(\theta)$ is a bijective mapping, we again denote $\vartheta(z)$ its inverse, which indicates the type θ with earnings z. This allows us to define $s(z) := s(z;\vartheta(z))$, $p(z) := p(z(\vartheta(z));\vartheta(z))$ and to formulate the problem in terms of observable earnings z.

Let λ_1 and λ_2 be the marginal values of public funds associated with the following resource constraints:

$$\int_z T_1(z) dH_z(z) \geq E_1 \tag{189}$$

$$\int_z T_2(s(z),z) dH_z(z) \geq E_2 \tag{190}$$

The problem of the government is to maximize the following Lagrangian

$$\mathcal{L} = \int_z \Bigg\{ \alpha(z)\, U\Big(z - T_1(z) - p(z)\big(s(z) + T_2(s(z),z)\big), s(z), z; \vartheta(z)\Big) + \lambda_1 T_1(z) + \lambda_2 T_2(s(z),z) - E_1 - E_2 \Bigg\} dH_z(z) \tag{191}$$

B.10.2 Adapting Lemma 1

Lemma B.4. *For an agent $\theta = \vartheta(z)$, we have that:*
(1a) a small increase $d\tau_z$ in the marginal tax rate $\frac{\partial T_2}{\partial z}$ generates the same earnings change as a small increase $p(z)d\tau_z$ in the marginal tax rate $\frac{\partial T_1}{\partial z}$.
(1b) a small increase $d\tau_s$ in the marginal tax rate $\frac{\partial T_2}{\partial s}$ generates the same earnings change as a small increase $p(z)s'_{inc}(z)d\tau_s$ in the marginal tax rate $\frac{\partial T_1}{\partial z}$.
(2) a small increase dT in the T_2 tax liability faced by agent $\theta = \vartheta(z)$ generates the same earnings change as a small increase $p(z)dT$ in the T_1 tax liability.

Proof. We first derive an abstract characterization, that we then apply to different tax reforms.

Let agent θ indirect utility function at earnings z be

$$V\left(T_1(z), T_2(.,z), z; \theta\right) := \max_s \; U\Big(z - T_1(z) - p(z,\theta)\left(s + T_2(s,z)\right), s, z; \theta\Big). \tag{192}$$

Consider a small reform $dT_2(s,z)$ of T_2, and let construct for each type θ a perturbation $dT_1^\theta(z)$ of T_1 that induces the same earnings response as the initial perturbation. Suppose we define this perturbation for each type θ such that at all earnings z,

$$V(T_1(z) + dT_1^\theta(z), T_2(.,z), z; \theta) = V(T_1(z), T_2(.,z) + dT_2(.,z), z; \theta) \tag{193}$$

then, by construction, the perturbation $dT_1^\theta(z)$ induces the same earnings response dz as the initial perturbation $dT_2(.,z)$. Moreover, both tax reforms must induce the same utility change for agent θ. Applying the envelope theorem yields

$$-U_c'(z;\theta) \cdot dT_1^\theta(z) = -U_c'(z;\theta)\, p(z,\theta) \cdot dT_2\left(s(z;\theta), z\right) \tag{194}$$

such that finally, the perturbation $dT_1^\theta(z)$ is

$$dT_1^\theta(z) = p(z,\theta) \cdot dT_2\left(s(z;\theta), z\right). \tag{195}$$

and we can now apply this abstract characterization to different tax reforms.

(1a) Consider a small increase $d\tau_z$ in the marginal tax rate $\frac{\partial T_2}{\partial z}$ over a small bandwidth of income $\left[z^0, z^0 + dz\right]$. Then, for any agent θ such that $z(\theta) \in \left[z^0, z^0 + dz\right]$, we have $dT_2\left(s(z;\theta), z\right) = d\tau_z\left(z - z^0\right)$ such that $dT_1^\theta(z) = p(z,\theta) d\tau_z\left(z - z^0\right)$ and differentiating with respect to z we get

$$\Big(dT_1^\theta(z)\Big)' = \frac{\partial p(z,\theta)}{\partial z} d\tau_z\left(z - z^0\right) + p(z,\theta) d\tau_z \tag{196}$$

At the limit $dz \to 0$ such that $z \to z^0$, a small increase $d\tau_z$ in the marginal tax rate $\frac{\partial T_2}{\partial z}$ generates the same earnings change as a small increase $p(z)d\tau_z$ in the marginal tax rate $T_1'(z)$.

(1b) Consider a small increase $d\tau_s$ in the marginal tax rate $\frac{\partial T_2}{\partial s}$ over a small bandwidth of savings $\left[s^0, s^0 + ds\right]$. Then, for any agent θ such that $s(\theta) \in \left[s^0, s^0 + ds\right]$, we have $dT_2\left(s(z;\theta), z\right) = d\tau_s\left(s(z;\theta) - s^0\right)$ such that $dT_1^\theta(z) = p(z,\theta) d\tau_z\left(s(z;\theta) - s^0\right)$ and differentiating with respect to z we get

$$\Big(dT_1^\theta(z)\Big)' = \frac{\partial p(z,\theta)}{\partial z} d\tau_z\left(s(z;\theta) - s^0\right) + p(z,\theta) d\tau_z s_{inc}'(z) \tag{197}$$

At the limit $ds \to 0$ such that $s \to s^0$, a small increase $d\tau_s$ in the marginal tax rate $\frac{\partial T_2}{\partial s}$ generates the same earnings change as a small increase $p(z)s_{inc}'(z)d\tau_z$ in the marginal tax rate $T_1'(z)$.

(2) Consider a small lump-sum increase dT in the T_2 tax liability for an agent θ who earns z, we then have $dT_1^\theta(z) = p(z,\theta)dT$ such that the equivalent reform is no longer a lump-sum increase. Hence, a small increase dT in the T_2 tax liability faced by an agent $\vartheta(z)$ generates the same earnings change as a small increase $p(z)dT$ in the T_1 tax liability. □

B.10.3 Marginal values of public funds

An important prerequisite to derive optimality conditions is to pin down the marginal values of public funds λ_1 and λ_2. At the optimum, λ_1 and λ_2 are pinned down by optimally setting the tax level T_1 and T_2. Characterizing the impact of lump-sum changes in tax liabilities yields the following two equations that can be solved for λ_1 and λ_2:

$$\int_{x=\underline{z}}^{\overline{z}} \left\{ -\alpha(x)U_c'(x) + \lambda_1 + \left(\lambda_1 T_1'(x) + \lambda_2 \frac{\partial T_2}{\partial z} + s_{inc}' \lambda_2 \frac{\partial T_2}{\partial s} \right) \frac{\partial z(.)}{\partial T_1} + \lambda_2 \frac{\partial T_2}{\partial s} \frac{\partial s(.)}{\partial T_1} \right\} dH_z(x) = 0 \tag{198}$$

$$\int_{x=\underline{z}}^{\overline{z}} \left\{ -\alpha(x)p(x)U_c'(x) + \lambda_2 + \left(\lambda_1 T_1'(x) + \lambda_2 \frac{\partial T_2}{\partial z} + s_{inc}' \lambda_2 \frac{\partial T_2}{\partial s} \right) \frac{\partial z(.)}{\partial T_2} + \lambda_2 \frac{\partial T_2}{\partial s} \frac{\partial s(.)}{\partial T_2} \right\} dH_z(x) = 0 \tag{199}$$

where $z(.)$ and $s(.)$ denote with a slight abuse of notations the earnings and savings choices and all partial derivatives are evaluated at earnings x.

Renormalizing these equations by λ_1, we can use the fact that by Lemma 2, $\frac{\partial z(.)}{\partial T_2} = \frac{\partial z(.)}{\partial T_1} p(z) + \frac{\partial z(.)}{\partial T_1'} \frac{\partial p}{\partial z}$ and that $\frac{\partial s(.)}{\partial T_2} = \frac{\partial s(.)}{\partial T_1} p(z)$ to obtain

$$\int_{x=\underline{z}}^{\overline{z}} \left\{ 1 - \left[\alpha(x) \frac{U_c'(x)}{\lambda_1} - \left(T_1'(x) + \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial z} + s_{inc}' \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial s} \right) \frac{\partial z(.)}{\partial T_1} - \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial s} \frac{\partial s(.)}{\partial T_1} \right] \right\} dH_z(z) = 0 \tag{200}$$

$$\begin{aligned} \int_{x=\underline{z}}^{\overline{z}} \Bigg\{ \frac{\lambda_2}{\lambda_1} - p(x) \left[\alpha(x) \frac{U_c'(x)}{\lambda_1} - \left(T_1'(x) + \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial z} + s_{inc}' \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial s} \right) \frac{\partial z(.)}{\partial T_1} - \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial s} \frac{\partial s(.)}{\partial T_1} \right] \\ + \left(T_1'(z) + \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial z} + s_{inc}' \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial s} \right) \frac{\partial z(.)}{\partial T_1'} \frac{\partial p}{\partial z} \Bigg\} dH_z(x) = 0 \end{aligned} \tag{201}$$

At any given earnings x, defining social marginal welfare weights augmented with the fiscal impact of income effects $\hat{g}(x)$, and the fiscal impacts of the novel substitution effects $\varphi(x)$ as respectively

$$\hat{g}(x) := \alpha(x) \frac{U_c'(x)}{\lambda_1} - \left(T_1'(x) + \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial z} + s_{inc}'(x) \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial s} \right) \frac{\partial z(.)}{\partial T_1} - \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial s} \frac{\partial s(.)}{\partial T_1} \tag{202}$$

$$\varphi(x) := \left(T_1'(x) + \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial z} + s_{inc}'(x) \frac{\lambda_2}{\lambda_1} \frac{\partial T_2}{\partial s} \right) \frac{\partial z(.)}{\partial T_1'} \frac{\partial p}{\partial z} \tag{203}$$

where all partial derivatives are evaluated at x, we finally obtain

$$\overline{\hat{g}(x)} := \int_{x=\underline{z}}^{\overline{z}} \hat{g}(x)\, dH_z(x) = 1 \tag{204}$$

$$\overline{\hat{g}(x)p(x) - \varphi(x)} := \int_{x=\underline{z}}^{\overline{z}} \Big(\hat{g}(x)p(x) - \varphi(x) \Big)\, dH_z(x) = \frac{\lambda_2}{\lambda_1} \tag{205}$$

B.10.4 Optimal marginal tax rates $T_1'(z)$

We consider a small reform at earnings level z^0 that consists in a small increase $d\tau_z$ of the marginal earnings tax rate $T_1'(z)$ in a small bandwidth dz. The impact on the Lagrangian is as $dz \to 0$,

$$\begin{aligned}\frac{d\mathcal{L}}{d\tau_z dz} &= \int_{x=z^0}^{\bar{z}} \Big(\lambda_1 - \alpha(x)U_c'(x)\Big) dH_z(x) \\ &+ \left[\lambda_1 T_1'\left(z^0\right) + \lambda_2 \frac{\partial T_2}{\partial z}\Big|_{z=z^0}\right] \frac{\partial z(.)}{\partial T_1'\left(z^0\right)} h_z(z^0) \\ &+ \int_{x=z^0}^{\bar{z}} \left[\lambda_1 T_1'\left(x\right) + \lambda_2 \frac{\partial T_2}{\partial z}\right] \frac{\partial z(.)}{\partial T_1} dH_z(x) \\ &+ \lambda_2 \frac{\partial T_2}{\partial s}\Big|_{z=z^0} s_{inc}'(z^0) \frac{\partial z(.)}{\partial T_1'\left(z^0\right)} h_z(z^0) \\ &+ \int_{x=z^0}^{\bar{z}} \lambda_2 \frac{\partial T_2}{\partial s} \left[\frac{\partial s(.)}{\partial T_1} + s_{inc}'(x) \frac{\partial z(.)}{\partial T_1}\right] dH_z(x).\end{aligned} \tag{206}$$

We characterize optimal taxes through $d\mathcal{L} = 0$. Renormalizing everything by λ_1, plugging in social marginal welfare weights augmented with income effects $\hat{g}(x)$, we obtain the following optimality condition for marginal earnings tax rates at each earnings z^0

$$-\left[T_1'(z^0) + \frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial z}\Big|_{z^0} + s_{inc}'(z^0)\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\Big|_{z^0}\right] \frac{\partial z(.)}{\partial T_1'\left(z^0\right)} = \frac{1}{h_z(z^0)} \int_{x=z^0}^{\bar{z}} \Big[1 - \hat{g}(x)\Big] dH_z(x) \tag{207}$$

B.10.5 Optimal marginal tax rates $\frac{\partial T_2}{\partial z}$

We consider a small reform at earnings level z^0 that consists in a small increase $d\tau_z$ of the marginal earnings tax rate $\frac{\partial T_2}{\partial z}$ in a small bandwidth dz. Using Lemma 2, we characterize the impact of the reform on the Lagrangian as $dz \to 0$,

$$\begin{aligned}\frac{d\mathcal{L}}{d\tau_z dz} &= \int_{x=z^0}^{\bar{z}} \Big(\lambda_2 - \alpha(x)p(x)U_c'(x)\Big) dH_z(x) \\ &+ \left[\lambda_1 T_1'(z^0) + \lambda_2 \frac{\partial T_2}{\partial z}\Big|_{z=z^0}\right] \frac{\partial z(.)}{\partial T_1'\left(z^0\right)} p(z^0) h_z(z^0) \\ &+ \int_{x=z^0}^{\bar{z}} \left[\lambda_1 T_1'(x) + \lambda_2 \frac{\partial T_2}{\partial z}\right] \left(\frac{\partial z(.)}{\partial T_1} p(x) + \frac{\partial z(.)}{\partial T_1'\left(x\right)} \frac{\partial p}{\partial z}\right) dH_z(x) \\ &+ \lambda_2 \frac{\partial T_2}{\partial s}\Big|_{z=z^0} \left[s_{inc}'(z^0) \frac{\partial z(.)}{\partial T_1'(z^0)} p(z^0)\right] h_z(z^0) \\ &+ \int_{x=z^0}^{\bar{z}} \lambda_2 \frac{\partial T_2}{\partial s} \left[\frac{\partial s(.)}{\partial T_2} + s_{inc}'(x) \left(\frac{\partial z(.)}{\partial T_1} p(x) + \frac{\partial z(.)}{\partial T_1'\left(x\right)} \frac{\partial p}{\partial z}\right)\right] dH_z(x)\end{aligned} \tag{208}$$

We characterize optimal taxes through $d\mathcal{L}=0$. Renormalizing by λ_1 and using $\frac{\partial s(.)}{\partial T_2}=\frac{\partial s(.)}{\partial T_1}p(x)$, we can plug in $\hat{g}(x)$ and $\varphi(x)$ to obtain the following optimality condition at each earnings z^0

$$-\left[T_1'(z^0)+\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial z}\Big|_{z^0}+s'_{inc}(z^0)\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\Big|_{z^0}\right]\frac{\partial z(.)}{\partial T_1'(z^0)}p(z^0)=\frac{1}{h_z(z^0)}\int_{x=z^0}^{\bar{z}}\left\{\frac{\lambda_2}{\lambda_1}-\hat{g}(x)p(x)+\varphi(x)\right\}dH_z(x) \tag{209}$$

B.10.6 Optimal marginal tax rates $\frac{\partial T_2}{\partial s}$

We consider a small reform at $s^0=s(z^0)$ that consists in a small increase $d\tau_s$ of $\frac{\partial T_2}{\partial s}$, the marginal tax rate on s, in a small bandwidth ds. Using Lemma 2, we characterize the impact of the reform on the Lagrangian as $ds\to 0$

$$\begin{aligned}\frac{d\mathcal{L}}{d\tau_s ds}=&\int_{x=z^0}^{\bar{z}}\Big(\lambda_2-\alpha(x)p(x)U_c'(x)\Big)\,dH_z(x)\\&+\left[\lambda_1T_1'(z^0)+\lambda_2\frac{\partial T_2}{\partial z}\Big|_{z=z^0}\right]\frac{\partial z(.)}{\partial T_1'(z^0)}s'_{inc}(z^0)p(z^0)\,\frac{h_z(z^0)}{s'(z^0)}\\&+\int_{x=z^0}^{\bar{z}}\left[\lambda_1T_1'(x)+\lambda_2\frac{\partial T_2}{\partial z}\right]\left(\frac{\partial z(.)}{\partial T_1}p(x)+\frac{\partial z(.)}{\partial T_1'(x)}\frac{\partial p}{\partial z}\right)dH_z(x)\\&+\lambda_2\frac{\partial T_2}{\partial s}\Big|_{z=z^0}\left[\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\Big|_{z=z^0}\right)}+s'_{inc}(z^0)\frac{\partial z(.)}{\partial T_1'(z^0)}s'_{inc}(z^0)p(z^0)\right]\frac{h_z(z^0)}{s'(z^0)}\\&+\int_{x=z^0}^{\bar{z}}\lambda_2\frac{\partial T_2}{\partial s}\left[\frac{\partial s(.)}{\partial T_2}+s'_{inc}(x)\left(\frac{\partial z(.)}{\partial T_1}p(x)+\frac{\partial z(.)}{\partial T_1'(x)}\frac{\partial p}{\partial z}\right)\right]dH_z(x)\end{aligned} \tag{210}$$

We characterize optimal taxes through $d\mathcal{L}=0$. Renormalizing by λ_1 and using $\frac{\partial s(.)}{\partial T_2}=\frac{\partial s(.)}{\partial T_1}p(x)$, we can plug in $\hat{g}(x)$ and $\varphi(x)$ to obtain the following optimality condition for marginal tax rates on s at each savings $s^0=s(z^0)$

$$\begin{aligned}&-\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\Big|_{z^0}\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\Big|_{z^0}\right)}h_z(z^*)=s'(z^0)\int_{x=z^0}^{\bar{z}}\left\{\frac{\lambda_2}{\lambda_1}-\hat{g}(x)p(x)+\varphi(x)\right\}dH_z(x)\\&+\left[T_1'(z^0)+\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial z}\Big|_{z^0}+s'_{inc}(z^0)\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\Big|_{z^0}\right]\frac{\partial z(.)}{\partial T_1'(z^0)}s'_{inc}(z^0)p(z^0)h_z(z^0)\end{aligned} \tag{211}$$

B.10.7 Deriving Proposition 6

At this stage, let us highlight that a general two-part tax system $\{T_1(z),T_2(s,z)\}$ satisfies three optimality conditions: the optimality condition (207) for $T_1'(z)$, the optimality condition (209) for $\frac{\partial T_2}{\partial z}$, and the optimality condition (211) for $\frac{\partial T_2}{\partial s}$.

Combining the optimality condition (207) for $T_1'(z)$ and the optimality condition (209) for $\frac{\partial T_2}{\partial z}$, we find that at the optimum the distortions in marginal earnings tax rates induced by T_1 and T_2

are such that an optimal general two-part tax system satisfies

$$\int_{x=z^0}^{\bar{z}} \Big[1-\hat{g}(x)\Big]\, dH_z(x) = \frac{1}{p(z^0)} \int_{x=z^0}^{\bar{z}} \left\{ \frac{\lambda_2}{\lambda_1} - \hat{g}(x)p(x) + \varphi(x) \right\} dH_z(x) \tag{212}$$
$$\iff p(z^0) \int_{x=z^0}^{\bar{z}} \Big[\overline{\hat{g}(x)} - \hat{g}(x)\Big]\, dH_z(x) = \int_{x=z^0}^{\bar{z}} \left\{ \overline{\hat{g}(x)p(x) - \varphi(x)} - \Big(\hat{g}(x)p(x) - \varphi(x)\Big) \right\} dH_z(z)$$

where the equivalence follows from $\overline{\hat{g}(x)p(x) - \varphi(x)} = \frac{\lambda_2}{\lambda_1}$ and $\overline{\hat{g}(x)} = 1$.

In contrast, a restricted two-part tax system $\{T_1(z), T_2(s)\}$ satisfies only two optimality conditions: the optimality condition (207) for $T_1'(z)$ and the optimality condition (211) for $\frac{\partial T_2}{\partial s}$. As a result, the previous relationship does not hold in an optimal restricted two-part tax system.

Combining the optimality condition (207) for $T_1'(z)$ and the optimality condition (211) for $\frac{\partial T_2}{\partial s}$, we get that optimal general and restricted tax systems both satisfy at each earnings z^0

$$-\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\Big|_{z^0} \frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\Big|_{z^0}\right)} = \frac{s'(z^0)}{h_z(z^0)} \int_{x=z^0}^{\bar{z}} \left\{ \frac{\lambda_2}{\lambda_1} - \hat{g}(x)p(x) + \varphi(x) \right\} dH_z(x) - p(z^0)\frac{s'_{inc}(z^0)}{h_z(z^0)} \int_{x=z^0}^{\bar{z}} \Big[1-\hat{g}(x)\Big]\, dH_z(x) \tag{213}$$

Adding and subtracting $p(z^0)\frac{s'(z^0)}{h_z(z^0)} \int_{x=z^0}^{\bar{z}} \Big[1-\hat{g}(x)\Big]\, dH_z(x)$ yields

$$\begin{aligned} -\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\Big|_{z^0} \frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\Big|_{z^0}\right)} &= p(z^0)\frac{s'(z^0) - s'_{inc}(z^0)}{h_z(z^0)} \int_{x=z^0}^{\bar{z}} \Big[1-\hat{g}(x)\Big]\, dH_z(x) \\ &+ \frac{s'(z^0)}{h_z(z^0)} \int_{x=z^0}^{\bar{z}} \left\{ \frac{\lambda_2}{\lambda_1} - \hat{g}(x)p(x) + \varphi(x) \right\} dH_z(x) - p(z^0)\frac{s'(z^0)}{h_z(z^0)} \int_{x=z^0}^{\bar{z}} \Big[1-\hat{g}(x)\Big]\, dH_z(x) \end{aligned} \tag{214}$$

Hence, for a general tax system in which $T_2(s,z)$ depends on both s and z, distortions in marginal earnings tax rates induced by T_1 and T_2 are such that the last two terms cancel out. Plugging in $\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\Big|_{z^0}\right)} = -\frac{s(z^0)}{1+p(z^0)\frac{\partial T_2}{\partial s}\Big|_{z^0}}\zeta^c_{s|z}(z^0)$ as well as $\frac{\lambda_2}{\lambda_1} = \overline{\hat{g}(x)p(x) - \varphi(x)}$, we then obtain

$$\frac{\overline{\hat{g}(x)p(x) - \varphi(x)}\frac{\partial T_2}{\partial s}\Big|_{z^0}}{1 + p(z^0)\frac{\partial T_2}{\partial s}\Big|_{z^0}} = \frac{1}{s(z^0)\zeta^c_{s|z}(z^0)}\frac{1}{h_z(z^0)} \left\{ s'_{pref}(z^0)p(z^0) \int_{x=z^0}^{\bar{z}} \Big[1-\hat{g}(x)\Big]\, dH_z(x) \right\} \tag{215}$$

which is the first condition stated in Proposition 6.

For a restricted tax system in which $T_2(s)$ depends only on s, the last two terms do not cancel out. Plugging in $\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\Big|_{z^0}\right)} = -\frac{s(z^0)}{1+p(z^0)\frac{\partial T_2}{\partial s}\Big|_{z^0}}\zeta^c_{s|z}(z^0)$ as well as $\frac{\lambda_2}{\lambda_1} = \overline{\hat{g}(x)p(x) - \varphi(x)}$, we then obtain

$$\begin{aligned} &\frac{\overline{\hat{g}(x)p(x) - \varphi(x)}\frac{\partial T_2}{\partial s}\Big|_{z^0}}{1 + p(z^0)\frac{\partial T_2}{\partial s}\Big|_{z^0}} \\ &= \frac{1}{s(z^0)\zeta^c_{s|z}(z^0)}\frac{1}{h_z(z^0)} \left\{ s'_{pref}(z^0)p(z^0) \int_{x=z^0}^{\bar{z}} \Big[1-\hat{g}(x)\Big]\, dH_z(x) + s'(z^0)\Big[\Psi(z^0) + \Phi(z^0)\Big] \right\} \end{aligned} \tag{216}$$

where we again use $\overline{\hat{g}(x)p(x)-\varphi(x)}=\frac{\lambda_2}{\lambda_1}$ and $\overline{\hat{g}(x)}=1$ to obtain the additional terms

$$\begin{aligned}\Psi(z^0) &:= \int_{z=z^0}^{\bar{z}}\Big[\overline{\hat{g}(z)p(z)}-\hat{g}(z)p(z)\Big]dH_z(z)-p(z^0)\int_{z=z^0}^{\bar{z}}\Big[\overline{\hat{g}(z)}-\hat{g}(z)\Big]dH_z(z) \qquad (217)\\
&=\int_{z=z^0}^{\bar{z}}\left[\int_{x=\underline{z}}^{\bar{z}}\hat{g}(x)p(x)dH_z(x)-\hat{g}(z)p(z)\right]dH_z(z)-p(z^0)\int_{z=z^0}^{\bar{z}}\Big[\int_{x=\underline{z}}^{\bar{z}}\hat{g}(x)dH_z(x)-\hat{g}(z)\Big]dH_z(z)\\
&=\big(1-H_z(z^0)\big)\int_{x=\underline{z}}^{\bar{z}}\hat{g}(x)p(x)dH_z(x)-\int_{z=z^0}^{\bar{z}}\hat{g}(z)p(z)dH_z(z)\\
&\qquad -p(z^0)\big(1-H_z(z^0)\big)\int_{x=\underline{z}}^{\bar{z}}\hat{g}(x)dH_z(x)-p(z^0)\int_{z=z^0}^{\bar{z}}\hat{g}(z)dH_z(z)\\
&=\big(1-H_z(z^0)\big)\int_{x=\underline{z}}^{\bar{z}}\hat{g}(x)\Big(p(x)-p(z^0)\Big)dH_z(x)-\int_{x=z^0}^{\bar{z}}\hat{g}(x)\Big(p(x)-p(z^0)\Big)dH_z(x)\\
&=\Big(1-H_z(z^0)\Big)\int_{x=\underline{z}}^{z^0}\hat{g}(x)\Big(p(x)-p(z^0)\Big)dH_z(z)+H_z(z^0)\int_{x=z^0}^{\bar{z}}\hat{g}(x)\Big(p(z^0)-p(x)\Big)dH_z(x) \qquad (218)\end{aligned}$$

$$\Phi(z^0) := \int_{x=z^0}^{\bar{z}}\Big[\varphi(x)-\overline{\varphi(x)}\Big]dH_z(x) \qquad (219)$$

This proves the second part of Propoition 6.

Now, combining the optimality condition (209) for $\frac{\partial T_2}{\partial z}$ and the optimality condition (211) for $\frac{\partial T_2}{\partial s}$, we obtain that in a general tax system

$$\begin{aligned}&\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\Big|_{z^0}\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\Big|_{z^0}\right)} \qquad (220)\\
&=\Big(s'(z^0)-s'_{inc}(z^0)\Big)p(z^0)\left[T_1'(z^0)+\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial z}\Big|_{z^0}+s'_{inc}(z^0)\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\Big|_{z^0}\right]\frac{\partial z(.)}{\partial T_1'(z^0)}\end{aligned}$$

Plugging in $\frac{\lambda_2}{\lambda_1}=\overline{\hat{g}(x)p(x)-\varphi(x)}$ as well as $\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\Big|_{z^0}\right)}=-\frac{s(z^0)}{1+p(z^0)\frac{\partial T_2}{\partial s}\Big|_{z^0}}\zeta^c_{s|z}(z^0)$ and $\frac{\partial z(.)}{\partial T_1'(z^0)}=-\frac{1}{T_1'(z^0)+p(z^0)\frac{\partial T_2}{\partial z}\Big|_{z^0}+\frac{\partial p}{\partial z}\Big|_{z^0}T_2(s(z^0),z^0)}z^0\zeta^c_z(z^0)$ yields the following Pareto-efficiency condition for a general two-part tax system

$$\begin{aligned}&\frac{\overline{\hat{g}(x)p(x)-\varphi(x)}\frac{\partial T_2}{\partial s}\Big|_{z^0}}{1+p(z^0)\frac{\partial T_2}{\partial s}\Big|_{z^0}} \qquad (221)\\
&=s'_{pref}(z^0)p(z^0)\frac{T_1'(z^0)+\overline{\hat{g}(x)p(x)-\varphi(x)}\frac{\partial T_2}{\partial z}\Big|_{z^0}+s'_{inc}(z^0)\overline{\hat{g}(x)p(x)-\varphi(x)}\frac{\partial T_2}{\partial s}\Big|_{z^0}}{T_1'(z^0)+p(z^0)\frac{\partial T_2}{\partial z}\Big|_{z^0}+\frac{\partial p}{\partial z}\Big|_{z^0}T_2(s(z^0),z^0)}\frac{z^0\zeta^c_z(z^0)}{s(z^0)\zeta^c_{s|z}(z^0)}\end{aligned}$$

C Details on the Empirical Application

This appendix describes the details underlying the numerical results presented in Section 6. Those results report the savings tax rates that would be Pareto efficient — holding fixed the status quo income tax — within each class simple tax systems described in Table 1: SL, SN, and LED.

Throughout this section, we assume that the nonlinear income tax is levied during each individual's working life, denoted $T_1(z)$ to clarify that the tax is measured in units of "period 1" (i.e., working life) dollars. We assume that the savings tax is collected at the time of retirement, denoting this component $T_2(s, z)$. Consistent with the notation in Appendix A.5, we use $s_g = s + T_2(s, g)$ to denote gross-of-tax savings, and we write $T_2^g(s_g, z)$ as the equivalent savings tax system, written as a function of gross rather than net savings. Appendix A.5 demonstrates that the simplicity structure of a tax system (SL, SN, and LED) is preserved when translating between $T_2(s, z)$ and $T_2^g(s_g, z)$.

C.1 Details of the Calibration of Income, Savings, and Elasticities

Here we describe the details of our calibration of a two-period version of the U.S. economy, with one period corresponding to working life and another to retirement. Throughout this section, we assume these periods are separated by 20 years, with a risk-free annual rate of return of 3.8% per year between period 1 and period 2.

C.1.1 Distribution of earnings $h(z)$, gross savings $s_g(z)$, and the nonlinear income tax schedule $T(z)$

The distribution of earnings and savings is calibrated for the U.S. from the Distributional National Accounts micro-files of Piketty et al. (2018), henceforth PSZ. We use individual measures of pretax labor income (*plinc*) and net personal wealth (*hweal*) as well as the age category (20 to 44 years old, 45 to 64, and above 65) and household information. We discretize the income distribution into percentiles by age group. Our measure of annualized earnings during work life z at the n-th percentile is constructed by averaging earnings at the n-th percentile across those aged 20 to 44 and those aged 45 to 64. For married households, we use the average earnings of the couple and assign both members of the couple to the same percentile of income. This can be important for households in which only one member of the household is working. If one member of the household is above 65 years old, we only keep the younger spouse in the sample. We drop the bottom 2% of observations with non-positive labor income; these individuals have positive average income from other sources, suggesting they are not representative of the zero ability types which would correspond to $z = 0$ in our model. We construct a smoothed savings distribution by fitting the log of our discretized probability mass function to log income.[33]

Our measure of gross retirement savings per year worked, which we denote s_g in the notation of Appendix A.5, at the n-th percentile of income is constructed by projecting forward to age 65 the average wealth we observe in the 45 to 65 age category at each percentile, assuming that

[33]This fitted schedule, and the savings and post-tax income fitted schedules below, are constructed using the smoothing spline fit in Matlab. The density uses a smoothing parameter of 0.99, while the savings and post-tax income schedules use a parameter of 0.9. Measures of savings are noisy at low incomes, which also have outlier values of $\ln(z)$ after the logarithmic transformation used for our savings fit. To avoid having those percentiles generate a strong pull on the fit, we fit the log of savings to $\ln(z + k)$, where a larger k reduces the extent to which the low incomes are outliers. Our baseline uses $k = \$20,000$.

agents in this category experience another 10 years of returns on average before retirement.[34] For married households, we take household wealth to be the average wealth of its members. This can be important for couples if assets are legally held by one household member. We then normalize the total wealth at retirement by the number of working years ($65 - 25 = 40$) so that z and s_g are in comparable units measured per working year. This yields a monotonic distribution of earnings z and gross savings $s_g(z)$, and pins down the cross-sectional variation in savings $s_g'(z)$.

We construct the status quo income tax function using the PSZ measure *diinc* ("extended disposable income") as a measure of post-tax income $z - T_1(z)$, using the median value within each pre-tax income percentile.

C.1.2 Measures of s'_{inc}

A key input for our sufficient statistics is the marginal propensity to save out of earned income, $s'_{inc}(z) := \frac{\partial s(z)}{\partial z}\big|_{\theta=\theta(z)}$, which relates changes in the amount of pre-tax (or gross) savings at the time of retirement, denoted s, to changes in the amount of pre-tax earnings z. We here show how we draw from two sources of empirical data to calibrate our marginal propensities to consume (or save), translated into measures of $s'_{inc}(z)$. These results are plotted in Figure 2.

Norwegian estimates from Fagereng et al. (2019). Fagereng et al. (2019) estimate marginal propensities to consume (MPC) across the earnings distribution using information on lottery prizes linked with administrative data in Norway. They find that agents' consumption peaks during the winning year and gradually reverts to their previous value afterwards. Over a 5-year horizon, they estimate agents consume close to 90% of the tax-exempt lottery prize, which translates into an MPC of 0.9, and thus a marginal propensity to save of 0.1. Under the assumption that preferences are weakly separable with respect to the disutility of labor supply, this is also the marginal propensity to save out of net earned income from labor supply. (See Proposition 1.)

They find little evidence of variation in MPCs across income levels which implies

$$\frac{\partial c(z)}{\partial (z - T_1(z))} = 0.9$$

and thus

$$\frac{\partial s_1(z)}{\partial (z - T_1(z))} = 1 - \frac{\partial c(z)}{\partial (z - T(z))} = 0.1.$$

[34] To account for the remaining 10 years (on average) of saving among 45- to 65-year-olds before retirement, we first impute the amount of net-of-tax income saved by middle-aged workers at each income percentile, which are compounded until the time of retirement. We do this imputation by constructing a representative working agent in each income percentile in each age category: a "young" agent of age 35 (the 20 to 44 age category, for which we assume work begins at age 25), and a "middle-aged" agent of age 55 (the 45 to 64 age category). We assume wealth at middle-aged consists of the sum of 20 years worth of savings while young, with returns compounded for an average of $55 - 35 = 20$ years, and 10 years of saving while middle-aged, compounded for an average of 5 years. Assuming that agents within each percentile save a the same share of post-tax income while young and middle-aged, this calculation pins down annual savings among middle-aged workers, which are assumed to continue until retirement at age 65.

This implies

$$\frac{\partial s_g(z)}{\partial(z - T_1(z))} = (1+r)0.1.$$

Therefore the local causal effect on *net* savings satisfies $\frac{\partial s(z)}{\partial(z-T_1(z))} = \frac{\partial s_g(z)}{\partial(z-T_1(z))}\frac{1}{1+T_s'}$. And again using $\partial(z - T_1(z)) = (1 - T_1'(z))\partial z$, we thus have

$$s'_{inc}(z) = (1+r) \times 0.1 \times \frac{1 - T_1'(z)}{1 + T_s'(s(z))}.$$

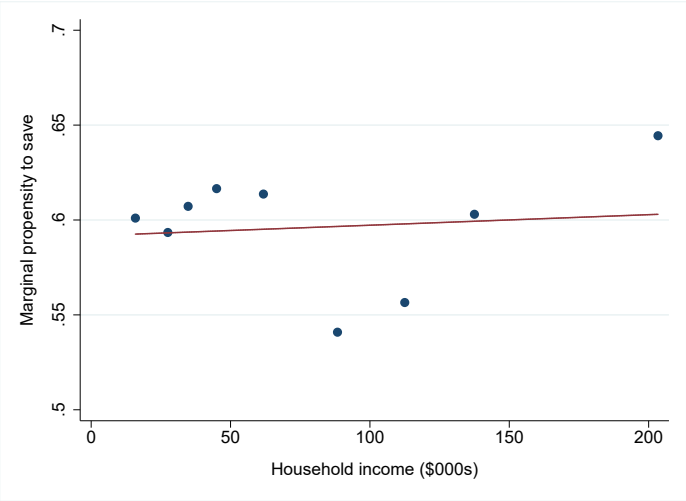
We can then use our calibrated US tax schedule to obtain a profile of $s'_{inc}(z)$, under the key assumption that US households have similar MPCs as Norwegian households. This profile is plotted in Figure 2.

U.S. estimates from a new Amerispeak survey In a survey experiment ran on a representative sample of US households, we ask the following question:

> Imagine that you or someone else in your household gets a raise such that over the next five years, your household's income is $1,000 higher each year than what you expected. How much of this would your household spend, and how much would your household save over each of the next five years? (For purposes of this question, consider paying off debt, such as reducing your mortgage, a form of saving.) If no one in your household is going to be employed for most of the next five years, please write "N/A."
>
> Spend an extra $ _____ per year
>
> Save an extra $ _____ per year

Answers to this question provide information about agents' reported marginal propensity to consume (MPC) and marginal propensity to save (MPS) out of earned income. Since we also collect information on household income in the survey, we can observe marginal propensity to save across earnings levels, plotted in Figure A1. Agents report on average a marginal propensity to save of 0.6 in the year of the raise. Moreover, marginal propensities to save seem strikingly constant across income levels which is consistent with the results of Fagereng et al. (2019).

Figure A1: Marginal propensity to save across household income (own survey)



Notes: Marginal propensities to save are computed from the answers to our survey question. They are computed as the ratio between the amount agents report they would save and the amount of the raise.

Since we ask about consumption and spending within each year, we interpret these estimates as short-run responses. Yet, Fagereng et al. (2019) show that positive income shocks are followed by consumption responses that can last up to 5 years. We thus rely on the evidence that they provide to convert these 1-year MPS into 5-year MPS that we interpret as a total effect on savings before returns. To do so, we use the fact that they report 1-year MPC of of 0.52 and a 5-year MPC of 0.90 and accordingly compute our long run marginal propensity to save as

$$MPS_{5y} = MPS_{1y}\frac{1-0.90}{1-0.52}\frac{1}{1+T_s'(s(z))} = \frac{0.125(1+r)}{1+T_s'(s(z))}.$$

We then multiply by the rate of return and adjust for savings taxes to reach our measure of $s'_{inc}(z)$. This estimate is slightly larger than that obtained by Fagereng et al. (2019) for Norway but remains of the same order of magnitude.

C.2 Comparison to Golosov et al. (2013)

In their baseline calibration, Golosov et al. (2013) assume agents preferences are CRRA

$$U = \frac{\alpha\left(w^i\right)}{1+\alpha\left(w^i\right)}\ln c_1^i + \frac{1}{1+\alpha\left(w^i\right)}\ln c_2^i - \frac{1}{\sigma}\left(l^i\right)^\sigma$$

where c_1 and c_2 represent first period and second period consumption levels and l is labor supply. The risk aversion parameter is here set to $\gamma = 1$, the isoelastic disulity from labor effort is such that $\sigma = 3$, and the taste parameter is given by

$$\alpha\left(w^i\right) = 1.0526\left(w^i\right)^{-0.0036}$$

meaning that it varies from 1.0433 for individuals in the bottom quintile of the earnings distribution (mean hourly wage of $12.35, in 1992 dollars) to 1.0406 for individuals in the top quintile of the earnings distribution (mean hourly wage of $25.39, in 1992 dollars). In other words, this parameter is almost constant with income around an average of $\overline{\alpha} = 1.042$.

To translate this structural calibration into our sufficient statistics, note that the earnings elasticity is given by $\zeta_z^c = \frac{1}{\sigma} = 0.33$.

Given the separability assumptions, we can recover the causal income effect on savings s'_{inc} through $s'_{inc} = \frac{1-\mathcal{T}'_z}{1+\mathcal{T}'_s}\eta_{s|z} = \frac{1-\mathcal{T}'_z}{1+\alpha(1+\mathcal{T}'_s)}$. Assuming a constant marginal earnings tax rate of 30%, and a constant marginal savings tax rate of 2% yields $s'_{inc} = \frac{1-0.3}{1+1.042*(1+0.02)} = 0.34$.

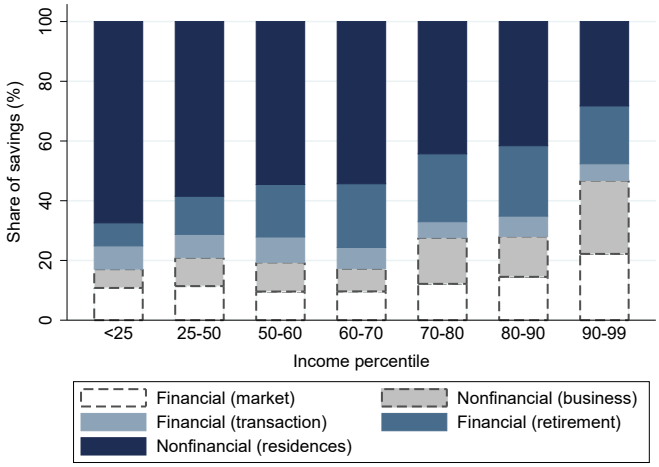
C.3 Estimates of Savings Tax Rates in the United States

Bricker et al. (2019) provide a decomposition of saving types by asset ownership percentile; we summarize the analogous decomposition by income percentile in Figure A2 below. We use the share of savings of each type in Figure A2(a) to estimate the savings tax rates in Figure A2(b). We estimate that Financial (market) and Nonfinancial (business) savings are subject to a 15% tax, while the other types of savings are not subject to any additional savings taxes.[35] Thus, we estimate the savings tax rate for each income group as 15%×(share of Financial (market) savings + share of Nonfinancial (business) savings).

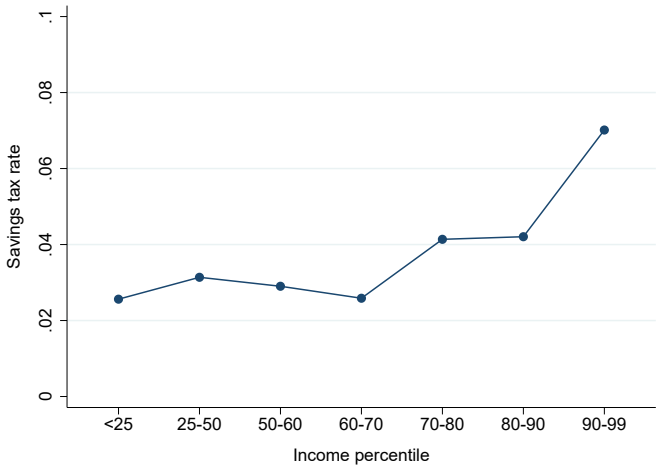
[35]The 15% estimated tax on Financial (market) savings and Nonfinancial (business) savings represent capital gains taxes. We assume that Financial (transaction) savings—which include checkings and savings accounts among other transaction accounts—and Financial (retirement) savings are tax exempt. We view property taxes on "Nonfinancial (residences)" savings as a tax that is incident on renters, and thus a component of imputed rent, which is paid regardless of whether the asset is owned by the user, so we also assume the tax rate on these savings is 0%.

Figure A2: Calibration of Savings Tax Rates Across Incomes in the U.S.

(a) Decomposition of Savings Types: Bricker et al. (2019)



(b) Calibrated Savings Tax Rates in the United States, by Income Percentile



Notes: This figure illustrates the calibration of savings tax rates in the U.S. across the income distribution. Panel (a) plots the composition of asset types in individuals' portfolios across the income distribution, reported by Bricker et al. (2019). Panel (b) plots the implied weighted average savings tax rate in each bin. See Appendix C.3 for details.

C.4 Inverse Optimum Approach

The inverse optimum approach aims at inferring the social marginal welfare weights (SMWW) consistent with actual tax policy (Bourguignon and Spadaro, 2012; Lockwood and Weinzierl, 2016). Indeed, assuming the actual policy is optimal, one can invert optimality conditions to obtain the social marginal welfare weights.

In the general case where we consider arbitrary smooth tax systems, we derive two optimal-

ity conditions gathered in Proposition 2. Proposition A.4 characterizes the optimal schedule of marginal earnings tax rates for any given (and potentially suboptimal) schedule of marginal savings tax rates, whereas Proposition 3 characterizes the optimal schedule of marginal savings tax rates for any given (and potentially suboptimal) schedule of marginal earnings tax rates.

There are thus two methodological choices to be made here. The first choice is whether we assume that the actual schedule of marginal earnings tax rates is optimal, or whether we assume that the actual schedule of marginal savings tax rates is optimal—if both are optimal then the actual tax system should verify Pareto-efficiency, which does not appear to be the case for our baseline specification in Figure 3. The second choice is how to compute the actual tax schedules on earnings and savings, and in particular whether we impose a particular functional form on the actual tax system.

If we assume that the actual schedule of marginal earnings tax rates is optimal, that it only depends on earnings, and that the savings tax is of the separable nonlinear (SN) type, we can infer social marginal welfare weights at each earnings z^* through

$$\frac{T_z'(z^*)}{1-T_z'(z^*)} = \frac{1}{\zeta_z^c(z^*)}\frac{1}{z^* h_z(z^*)}\int_{z^*}^{\bar{z}}(1-\hat{g}(x))\,dH_z(x) - s_{inc}'(z^*)\frac{T_s'(s(z^*))}{1-T_z'(z^*)} \tag{222}$$

$$\iff \int_{z^*}^{\bar{z}}(1-\hat{g}(x))\,dH_z(x) = \zeta_z^c(z^*)z^* h_z(z^*)\frac{T_z'(z^*)+s_{inc}'(z^*)T_s'(s(z^*))}{1-T_z'(z^*)} \tag{223}$$

where the right-hand side term can be identified from the data.

Differentiating with respect to z^* yields

$$\begin{aligned} -(1-\hat{g}(z^*))\,h_z(z^*) = {} & \frac{d\zeta_z^c(z^*)}{dz^*}z^* h_z(z^*)\frac{T_z'(z^*)+s_{inc}'(z^*)T_s'(s(z^*))}{1-T_z'(z^*)} \\ & + \zeta_z^c(z^*)\left[h_z(z^*)+z^* h_z'(z^*)\right]\frac{T_z'(z^*)+s_{inc}'(z^*)T_s'(s(z^*))}{1-T_z'(z^*)} \\ & + \zeta_z^c(z^*)z^* h_z(z^*)\frac{d}{dz^*}\left[\frac{T_z'(z^*)+s_{inc}'(z^*)T_s'(s(z^*))}{1-T_z'(z^*)}\right] \end{aligned} \tag{224}$$

and further assuming that the compensated elasticity of earnings is approximately constant across incomes ($\frac{d\zeta_z^c(z^*)}{dz^*}\approx 0$) yields

$$\begin{aligned} \hat{g}(z^*) = {} & 1+\zeta_z^c(z^*)\left[1+\frac{z^* h_z'(z^*)}{h_z(z^*)}\right]\frac{T_z'(z^*)+s_{inc}'(z^*)T_s'(s(z^*))}{1-T_z'(z^*)} \\ & + \zeta_z^c(z^*)z^*\left[\frac{[1+s_{inc}'(z^*)T_s'(s(z^*))]\,T_z''(z^*)}{[1-T_z'(z^*)]^2} + \frac{s_{inc}''(z^*)T_s'(s(z^*))+s_{inc}'(z^*)s'(z^*)T_s''(s(z^*))}{1-T_z'(z^*)}\right] \end{aligned} \tag{225}$$

which gives augmented social marginal welfare weights as a function of the local Pareto parameter of the income distribution $1+\frac{z^* h_z'(z^*)}{h_z(z^*)}$ and other observables.

Using the fact that augmented social marginal welfare weights are defined as

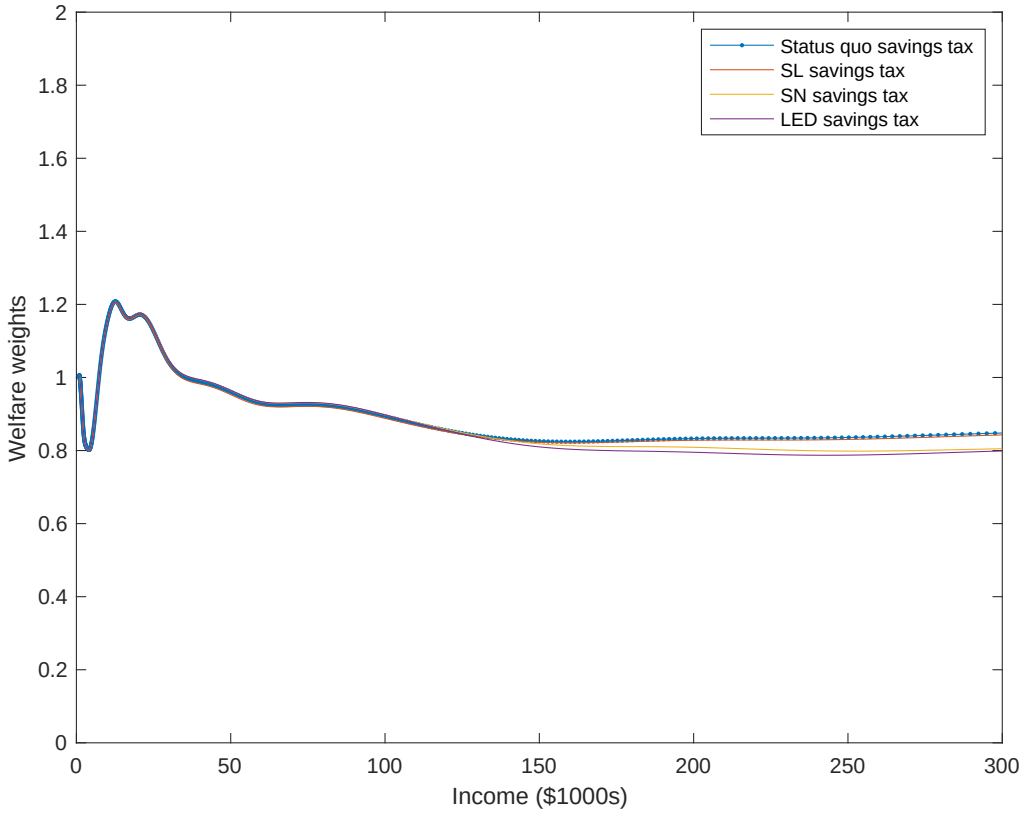
$$\hat{g}(z) := g(z) + T_z'(z)\frac{\eta_z(z)}{1-T_z'(z)} + T_s'(s(z))\left(\frac{\eta_{s|z}(z)}{1+T_s'(s(z))} + s_{inc}'(z)\frac{\eta_z(z)}{1-T_z'(z)}\right) \tag{226}$$

and assuming agents' preferences are weakly separable, such that by Proposition 1 we have $s'_{inc}(z) = \frac{1-T'_z(z)}{1+T'_s(s(z))}\eta_{s|z}(z)$, we finally get

$$\begin{aligned} g(z^*) = 1 &+ \zeta_z^c(z^*)\left[1 + \frac{z^* h'_z(z^*)}{h_z(z^*)}\right]\frac{T'_z(z^*) + s'_{inc}(z^*)T'_s(s(z^*))}{1 - T'_z(z^*)} \\ &+ \zeta_z^c(z^*)z^*\left[\frac{[1 + s'_{inc}(z^*)T'_s(s(z^*))]\,T''_z(z^*)}{[1 - T'_z(z^*)]^2} + \frac{s''_{inc}(z^*)T'_s(s(z^*)) + s'_{inc}(z^*)s'(z^*)T''_s(s(z^*))}{1 - T'_z(z^*)}\right] \\ &- T'_z(z)\frac{\eta_z(z)}{1 - T'_z(z)} - s'_{inc}(z)\,T'_s(s(z))\frac{1 + \eta_z(z)}{1 - T'_z(z)}. \end{aligned} \tag{227}$$

This expression depends on both the current schedule of marginal earnings tax rates, $T'_z(z)$, and savings tax rates, T'_s. We seek the schedule of weights that rationalizes the existing income tax schedule, and thus we set $T'_z(z)$ to reflect the observed schedule of income tax rates. Since the current schedule of savings tax rates is not Pareto efficient, there is a question of whether we compute weights assuming that the savings tax were counterfactually optimal, or whether we use the status quo (suboptimal) savings tax schedule. For consistency with the inverse "optimum" motivation, we use the former approach, but results look essentially identical if we instead use the suboptimal status quo schedule. Figure A3 plots our estimated profile of inverse optimum weights, both under the assumption that marginal savings tax rates are SN Pareto-efficient, and assuming the status quo schedule of savings tax rates.

Figure A3: Schedule of Inverse Optimum Social Welfare Weights in the U.S.



Notes: This figure plots the schedule of inverse optimum welfare weights that would rationalize the U.S. income tax schedule. Separate schedules are plotted using the status quo (Pareto-inefficient) savings tax, as well as under the assumption that the savings tax were (counterfactually) Pareto-efficient, with either a separable linear, separable nonlinear, or linear earnings-dependent structure.

C.5 Simulations of Optimal Savings Taxes with Multidimensional Heterogeneity

C.5.1 Overview and assumptions

Through numerical simulations, we assess how the presence of multidimensional heterogeneity affects optimal savings taxes in each of the simple tax system that we consider. To this end, we use optimal savings tax formulas derived in the presence of multidimensional heterogeneity (Proposition 3).

We assume that the existing earnings tax schedule is optimal to infer the government objective in an inverse optimum approach. This pins down social marginal welfare weights $g(z)$ that are consistent with actual tax policy under unidimensional heterogeneity. Further assuming that social marginal welfare weights are independent of savings s, we can then simulate optimal savings taxes under multidimensional heterogeneity keeping the government objective constant.

Doing so requires a number of assumptions regarding the structure of multidimensional heterogeneity. First, we assume that the compensated elasticity of earnings ζ_z^c and savings $\zeta_{s|z}^c$ are constant across earnings and savings, thereby disregarding any heterogeneity in these parameters.

Second, we assume that earnings income effects can be neglected (i.e. $\eta_z \approx 0$) meaning that changes in tax liabilities do not trigger labor supply responses – this simplifying assumption is by no means crucial for this simulation exercise because it is made for both unidimensional and multidimensional cases.

Third, we assume that agents' preferences are weakly separable, such that by Proposition 1 we have $s'_{inc}(z) = \frac{1-T'_z(z)}{1+T'_s(s(z))}\eta_{s|z}(z)$. Since we identify $s'_{inc}(z)$ from the data, this allows to pin down the income effect parameter on savings $\eta_{s|z}(z)$ in the unidimensional case. We then infer the value of $s'_{inc}(s,z)$ in the multidimensional case by assuming that the income elasticity of savings is constant within earnings and equal to its unidimensional counterpart. Formally, this means that at a given earnings z, for any savings s we have $\eta_{s|z}(s,z) = \frac{s}{\overline{s(z)}}\eta_{s|z}(z)$, where $\overline{s(z)} := \mathbb{E}\left[s\middle|z\right]$ denotes the average savings level at earnings z. And similarly, $s'_{inc}(s,z) = \frac{s}{\overline{s(z)}}s'_{inc}(z)$.

Under these assumptions, we derive for each type of simple tax system an optimal savings tax formula amenable to numerical simulations in the presence of multidimensional heterogeneity.

C.5.2 Separable linear (SL) tax system

The optimal savings tax formula with multidimensional heterogeneity (Proposition 3) is

$$\begin{aligned}&\frac{\tau_s}{1+\tau_s}\int_z\left\{\mathbb{E}\left[s\zeta_{s|z}^c(s,z)\middle|z\right]\right\}dH_z(z) \\ &=\int_z\left\{\mathbb{E}\left[(1-\hat{g}(s,z))\,s\middle|z\right]-\mathbb{E}\left[\frac{T'_z(z)+s'_{inc}(s,z)\,\tau_s}{1-T'_z(z)}z\zeta_z^c(s,z)\,s'_{inc}(s,z)\middle|z\right]\right\}dH_z(z)\end{aligned} \tag{228}$$

Under the aforementioned assumptions, expanding $\hat{g}(s,z)$, replacing $s'_{inc}(s,z)$ and $\eta_{s|z}(s,z)$ by their values, and assuming η_z is negligible gives

$$\begin{aligned}&\frac{\tau_s}{1+\tau_s}\int_z\left\{\overline{s(z)}\zeta_{s|z}^c\right\}dH_z(z) \\ &=\int_z\left\{\mathbb{E}\left[\left(1-g(z)-\tau_s\frac{\eta_{s|z}(z)}{1+\tau_s}\frac{s}{\overline{s(z)}}\right)s\middle|z\right]-\frac{z\zeta_z^c s'_{inc}(z)}{1-T'_z(z)}\mathbb{E}\left[T'_z(z)\frac{s}{\overline{s(z)}}+s'_{inc}(z)\,\tau_s\left(\frac{s}{\overline{s(z)}}\right)^2\middle|z\right]\right\}dH_z(z)\end{aligned} \tag{229}$$

which after rearranging yields

$$\begin{aligned}&\frac{\tau_s}{1+\tau_s}\zeta_{s|z}^c\int_z\overline{s(z)}dH_z(z) \\ &=\int_z\left\{(1-g(z))\,\overline{s(z)}-\frac{\tau_s}{1+\tau_s}\frac{\eta_{s|z}(z)}{\overline{s(z)}}\mathbb{E}\left[s^2\middle|z\right]-\frac{z\zeta_z^c s'_{inc}(z)}{1-T'_z(z)}\mathbb{E}\left[\frac{T'_z(z)}{\overline{s(z)}}s+\frac{s'_{inc}(z)\,\tau_s}{\overline{s(z)}^2}s^2\middle|z\right]\right\}dH_z(z).\end{aligned} \tag{230}$$

We can now use $\mathbb{E}\left[s^2\middle|z\right] = \mathbb{E}\left[\left(s-\overline{s\left(z\right)}\right)^2 + 2s\overline{s\left(z\right)} - \overline{s\left(z\right)}^2\middle|z\right] = \mathbb{V}\left(s|z\right) + \overline{s\left(z\right)}^2$ to obtain

$$\begin{aligned}&\frac{\tau_s}{1+\tau_s}\zeta^c_{s|z}\int_z \overline{s\left(z\right)}dH_z\left(z\right) \\ &= \int_z \left\{\left[1-g(z)-\frac{\tau_s\eta_{s|z}\left(z\right)}{1+\tau_s}\left(1+\frac{\mathbb{V}\left(s|z\right)}{\overline{s\left(z\right)}^2}\right)\right]\overline{s\left(z\right)} - \frac{z\zeta^c_z s'_{inc}\left(z\right)}{1-T'_z\left(z\right)}\left[T'_z\left(z\right)+s'_{inc}\left(z\right)\tau_s\left(1+\frac{\mathbb{V}\left(s|z\right)}{\overline{s\left(z\right)}^2}\right)\right]\right\}dH_z\left(z\right)\end{aligned} \tag{231}$$

which we can finally rewrite as

$$\begin{aligned}&\frac{\tau_s}{1+\tau_s}\int_z \overline{s\left(z\right)}\zeta^c_{s|z}dH_z\left(z\right) \\ &= \int_z \left\{\left(1-g(z)-\frac{\tau_s}{1+\tau_s}\eta_{s|z}\left(z\right)\right)\overline{s\left(z\right)} - \frac{T'_z\left(z\right)+s'_{inc}\left(z\right)\tau_s}{1-T'_z\left(z\right)}z\zeta^c_z s'_{inc}\left(z\right)\right\}dH_z\left(z\right) \\ &- \int_z \underbrace{\left\{\frac{\mathbb{V}\left(s|z\right)}{\overline{s\left(z\right)}^2}\right\}}_{\geq 0}\underbrace{\left(\frac{\tau_s}{1+\tau_s}\eta_{s|z}\left(z\right)\overline{s\left(z\right)} + \frac{s'_{inc}\left(z\right)\tau_s}{1-T'_z\left(z\right)}z\zeta^c_z s'_{inc}\left(z\right)\right)}_{\geq 0}\right\}dH_z\left(z\right).\end{aligned} \tag{232}$$

The first two lines correspond to the optimal savings tax formula under unidimensional heterogeneity (Proposition A.3) and the last line captures the effect of multidimensional heterogeneity through $\mathbb{V}\left(s|z\right)$. Multidimensional heterogeneity adds a corrective term which is unambiguously negative, it thus prescribes a lower linear savings tax rate.

C.5.3 Separable nonlinear (SN) tax system

The optimal savings tax formula with multidimensional heterogeneity (Proposition 3) is

$$\begin{aligned}&\frac{T'_s\left(s^*\right)}{1+T'_s\left(s^*\right)}\int_z \left\{s^*\zeta^c_{s|z}(s^*,z)\right\}h\left(s^*,z\right)dz \\ &= \int_z \left\{\mathbb{E}\Big[1-\hat{g}\left(s,z\right)\Big|z,s\geq s^*\Big]\right\}h_z\left(z\right)dz - \int_z \left\{\frac{T'_z\left(z\right)+s'_{inc}\left(s^*,z\right)T'_s\left(s^*\right)}{1-T'_z\left(z\right)}z\zeta^c_z(s^*,z)s'_{inc}\left(s^*,z\right)\right\}h\left(s^*,z\right)dz.\end{aligned} \tag{233}$$

Under the aforementioned assumptions, expanding $\hat{g}\left(s,z\right)$ and assuming η_z is negligible gives

$$\begin{aligned}&\frac{T'_s\left(s^*\right)}{1+T'_s\left(s^*\right)}s^*\zeta^c_{s|z}\int_z h\left(s^*,z\right)dz \\ &= \int_{s^*}^{\bar{s}}\left\{\int_z\left[1-g(z)-\frac{T'_s\left(s\right)}{1+T'_s\left(s\right)}s\frac{\eta_{s|z}\left(z\right)}{\overline{s\left(z\right)}}\right]h\left(s,z\right)dz\right\}ds - \int_z\left[\frac{T'_z\left(z\right)+s'_{inc}\left(s^*,z\right)T'_s\left(s^*\right)}{1-T'_z\left(z\right)}z\zeta^c_z s'_{inc}\left(s^*,z\right)\right]h\left(s^*,z\right)\end{aligned} \tag{234}$$

or equivalently, expressing this as a function of the savings density $h_s\left(s\right) = \int_z h\left(s,z\right)dz$,

$$\frac{T_s'(s^*)}{1+T_s'(s^*)}s^*\zeta^c_{s|z}h_s(s^*) \tag{235}$$
$$=\int_{s^*}^{\bar{s}}\left\{\mathbb{E}\left[1-g(z)-\frac{T_s'(s)}{1+T_s'(s)}\eta_{s|z}(s,z)\Big|s\right]\right\}h_s(s)\,ds-\mathbb{E}\left[\frac{T_z'(z)+s_{inc}'(s,z)\,T_s'(s)}{1-T_z'(z)}z\zeta^c_z s_{inc}'(s,z)\Big|s=s^*\right]h_s(s^*)$$

where expectations operator denote integration with respect to earnings conditional on savings.

For implementation, we assume that at each point in the income continuum, there are M different equal-sized saver-bins (e.g., bottom-, middle-, and top-third savers), indexed by $m = 1,\ldots,M$. Thus we can write $s_m(z)$ as the savings map for saver-bin m at each income, with $s_m'(z)$ the cross-sectional savings profile within each saver-bin. Then the income density in each saver-bin is $h_{z,m}(z) = h(z)/M$, since the bins are equally sized conditional on income. The savings density among saver-bin m is therefore $h_{s,m}(s) = h_{z,m}(z)/s_m'(z)$, and we have $H(s) = \sum_{m=1}^M \int_{s=0}^{\infty} h_{s,m}(s)ds$, and $h_s(s) = \sum_{m=1}^M h_{s,m}(s)$. And the savings-conditional average of some $x(s,z)$ is $\mathbb{E}[x(s,z)|s] = \frac{\sum_{m=1}^M x(s_m,z)h_{s,m}(s)}{h_s(s)}$.

To better picture the link with the unidimensional formula (46), let also rewrite the latter as a function of the savings density $h_s(s)$ – implicitly defining $z(s)$ as the earnings level of agents with savings s – this yields

$$\frac{T_s'(s^*)}{1+T_s'(s^*)}s^*\zeta^c_{s|z}h_s(s^*) \tag{236}$$
$$=\int_{s^*}^{\bar{s}}\left\{1-g(z(s))-\frac{T_s'(s)}{1+T_s'(s)}\eta_{s|z}(z(s))\right\}h_s(s)\,ds-\frac{T_z'(z(s^*))+s_{inc}'(z(s^*))T_s'(s^*)}{1-T_z'(z(s^*))}z(s^*)\,\zeta^c_z s_{inc}'(z(s^*))h_s(s^*).$$

While it is clear that the multidimensional formula extends the unidimensional formula, determining the impact of multidimensional heterogeneity on tax rates is analytically more difficult and we thus rely on numerical simulations.

C.5.4 Linear earnings dependent (LED) tax system

The optimal savings tax formula in the presence of multidimensional heterogeneity (Proposition 3) is

$$\int_{z\geq z^*}\mathbb{E}\left[\frac{\tau_s(z)}{1+\tau_s(z)}s\zeta^c_{s|z}(s,z)\Big|z\right]dH_z(z)+\mathbb{E}\left[\frac{T_z'(z)+\tau_s'(z)\,s+s_{inc}'(s,z)\,\tau_s(z)}{1-T_z'(z)-\tau_s'(z)\,s}z\zeta^c_z(s,z)s\Big|z^*\right]h_z(z^*)$$
$$=\int_{z\geq z^*}\left\{\mathbb{E}\Big[(1-\hat{g}(s,z))\,s\Big|z\Big]-\mathbb{E}\left[\frac{T_z'(z)+\tau_s'(z)\,s+s_{inc}'(s,z)\,\tau_s(z)}{1-T_z'(z)-\tau_s'(z)\,s}z\zeta^c_z(s,z)s_{inc}'(s,z)\Big)\Big|z\right]\right\}dH_z(z)$$

Under the aforementioned assumptions, expanding $\hat{g}(s,z)$, replacing $s_{inc}'(s,z)$ and $\eta_{s|z}(s,z)$ by

their values, and assuming η_z is negligible gives

$$\int_{z\geq z^*} \frac{\tau_s(z)}{1+\tau_s(z)} \overline{s(z)} \zeta^c_{s|z} dH_z(z) \tag{237}$$
$$+ z^* \zeta^c_z \mathbb{E}\left[\frac{T'_z(z)+\tau'_s(z)s}{1-T'_z(z)-\tau'_s(z)s} s + \frac{\tau_s(z)}{1-T'_z(z)-\tau'_s(z)s} \frac{s'_{inc}(z)}{\overline{s(z)}} s^2 \Big| z^* \right] h_z(z^*)$$
$$= \int_{z\geq z^*} \left\{ \mathbb{E}\left[\left(1 - g(z) - \frac{\tau_s(z)}{1+\tau_s(z)} \frac{\eta_{s|z}(z)}{\overline{s(z)}} s \right) s \Big| z \right] \right.$$
$$\left. - z\zeta^c_z \frac{s'_{inc}(z)}{\overline{s(z)}} \mathbb{E}\left[\frac{T'_z(z)+\tau'_s(z)s}{1-T'_z(z)-\tau'_s(z)s} s + \frac{\tau_s(z)}{1-T'_z(z)-\tau'_s(z)s} \frac{s'_{inc}(z)}{\overline{s(z)}} s^2 \Big) \Big| z \right] \right\} dH_z(z) \tag{238}$$

Since the marginal tax rate on earnings $T'_z(z)+\tau'_s(z)s$ features savings s, it is hard to further simplify this formula while retaining an exact characterization. To obtain exact results, we must rely on numerical simulations. We can neverthless derive simpler formulas, either relying on approximations or making additional assumptions.

Using approximations. We can nevertheless provide simple analytical formulas disregarding this dependence and setting $s = \overline{s(z)}$ in marginal earnings tax rates. We believe that these formulas are informative in that they converge to exact expressions as the linear earnings dependent savings tax rate tends to a simple linear savings tax rate – that is $\tau'_s(z) = 0$ for all z. Moreover, although these approximations are not unbiased in that they provide an upper bound on the linear-earnings dependent savings tax rate, these upper bounds are tight as the approximation only amounts to assuming $\tau'_s(z^*)\mathbb{V}(s|z^*)$ is negligible.[36]

Further using $\mathbb{E}\left[s^2 \Big| z^*\right] = \mathbb{V}(s|z^*) + \overline{s(z^*)}^2$, our approximation for the optimal savings tax

[36]Indeed, because functions $X \mapsto \frac{X}{1-X}$ and $X \mapsto \frac{1}{1-X}$ are positive and increasing in X over $[0,1]$ the interval where optimal earnings marginal tax rates lie, we have that

$$\mathbb{E}\left[\frac{T'_z(z)+\tau'_s(z)s}{1-T'_z(z)-\tau'_s(z)s} s \Big| z^* \right] \geq \frac{T'_z(z^*)+\tau'_s(z^*)\overline{s(z^*)}}{1-T'_z(z^*)-\tau'_s(z^*)\overline{s(z^*)}} \mathbb{E}\left[s \Big| z^* \right] \tag{239}$$
$$\mathbb{E}\left[\frac{\tau_s(z)}{1-T'_z(z)-\tau'_s(z)s} s^2 \Big| z^* \right] \geq \frac{\tau_s(z^*)}{1-T'_z(z^*)-\tau'_s(z^*)\overline{s(z^*)}} \mathbb{E}\left[s^2 \Big| z^* \right] \tag{240}$$

formula is

$$\begin{aligned}
&\int_{z\geq z^*}\frac{\tau_s(z)}{1+\tau_s(z)}\overline{s(z)}\zeta^c_{s|z}dH_z(z) \qquad (241)\\
&+z^*\zeta^c_z\overline{s(z)}\left[\frac{T'_z(z^*)+\tau'_s(z^*)\overline{s(z^*)}}{1-T'_z(z^*)-\tau'_s(z^*)\overline{s(z^*)}}+\frac{s'_{inc}(z)\tau_s(z^*)}{1-T'_z(z^*)-\tau'_s(z^*)\overline{s(z^*)}}\left(1+\frac{\mathbb{V}(s|z^*)}{\overline{s(z^*)}^2}\right)\right]h_z(z^*)\\
&=\int_{z\geq z^*}\left\{(1-g(z))\overline{s(z)}-\frac{\tau_s(z)}{1+\tau_s(z)}\eta_{s|z}(z)\overline{s(z)}\left(1+\frac{\mathbb{V}(s|z^*)}{\overline{s(z^*)}^2}\right)\right.\\
&\left.-z\zeta^c_z s'_{inc}(z)\left[\frac{T'_z(z^*)+\tau'_s(z^*)\overline{s(z^*)}}{1-T'_z(z^*)-\tau'_s(z^*)\overline{s(z^*)}}+\frac{s'_{inc}(z)\tau_s(z)}{1-T'_z(z)-\tau'_s(z)s}\left(1+\frac{\mathbb{V}(s|z^*)}{\overline{s(z^*)}^2}\right)\right]\right\}dH_z(z)
\end{aligned}$$

such that finally we obtain

$$\begin{aligned}
&\int_{z^*}^{\bar z}\left\{\frac{\tau_s(z)}{1+\tau_s(z)}\overline{s(z)}\zeta^c_{s|z}\right\}dH_z(z)+\left[\frac{T'_z(z^*)+\tau'_s(z^*)\overline{s(z^*)}}{1-T'_z(z^*)-\tau'_s(z^*)\overline{s(z^*)}}+\frac{s'_{inc}(z)\tau_s(z^*)}{1-T'_z(z^*)-\tau'_s(z^*)\overline{s(z^*)}}\right]z^*\zeta^c_z\overline{s(z)}h_z(z^*)\\
&=\int_{z^*}^{\bar z}\left\{(1-g(z))\overline{s(z)}-\frac{\tau_s(z)}{1+\tau_s(z)}\eta_{s|z}(z)\overline{s(z)}-\left[\frac{T'_z(z)+\tau'_s(z)\overline{s(z)}+s'_{inc}(z)\tau_s(z)}{1-T'_z(z)-\tau'_s(z)\overline{s(z)}}\right]z\zeta^c_z s'_{inc}(z)\right\}dH_z(z)\\
&\underbrace{-\frac{\mathbb{V}(s|z^*)}{\overline{s(z^*)}^2}}_{\geq 0}\underbrace{\frac{s'_{inc}(z^*)\tau_s(z^*)}{1-T'_z(z^*)-\tau'_s(z^*)\overline{s(z^*)}}z^*\zeta^c_z\overline{s(z^*)}h_z(z^*)}_{\geq 0} \qquad (242)\\
&-\int_{z^*}^{\bar z}\underbrace{\frac{\mathbb{V}(s|z)}{\overline{s(z)}^2}}_{\geq 0}\left\{\underbrace{\frac{\tau_s(z)}{1+\tau_s(z)}\eta_{s|z}(z)\overline{s(z)}+\frac{s'_{inc}(z)\tau_s(z)}{1-T'_z(z)-\tau'_s(z)s}z\zeta^c_z s'_{inc}(z)}_{\geq 0}\right\}dH_z(z)
\end{aligned}$$

The first two lines correspond to the optimal savings tax formula under unidimensional heterogeneity (Proposition A.3), while the third and fourth lines capture the effect of multidimensional heterogeneity through $\mathbb{V}(s|z)$. Multidimensional heterogeneity adds corrective terms which are unambiguously positive, it thus prescribes lower linear earnings-dependent savings tax rates.

Assuming multidimensional optimality. So far, we have assumed the actual earnings tax schedule is optimal to infer social marginal welfare weights $g(z)$. Implicitly, this assumption is made under unidimensional heterogeneity. Now, if we further assume that the actual earnings earnings tax schedule is optimal under multidimensional heterogeneity, we can use this to simplify the previous exact formula.

Indeed, the formula for optimal earnings tax for a LED tax system under multidimensional heterogeneity is (the proof is omitted because it exactly mirrors that in the unidimensional case with an additional integral over savings s)

$$\mathbb{E}\left[\frac{T'_z(z)+\tau'_s(z)s+s'_{inc}(s,z)\tau_s(z)}{1-T'_z(z)-\tau'_s(z)s}z\zeta^c_z(s,z)\Big|z^*\right]h_z(z^*)=\int_{z\geq z^*}\Big\{\mathbb{E}\left[1-\hat g(s,z)\,|z\right]\Big\}h_z(z)\,dz \qquad (243)$$

The optimal savings tax formula in the presence of multidimensional heterogeneity (Proposition 3) then simplifies to

$$\int_{z\geq z^*}\mathbb{E}\left[\frac{\tau_s(z)}{1+\tau_s(z)}s\zeta^c_{s|z}(s,z)\Big|z\right]h_z(z)\,dz \tag{244}$$
$$=\int_{z\geq z^*}\mathbb{E}\Big[\hat{g}(s,z)-\hat{g}(s,z)\,s\Big|z\Big]h_z(z)\,dz-\int_{z\geq z^*}\mathbb{E}\left[\frac{T'_z(z)+\tau'_s(z)\,s+s'_{inc}(s,z)\,\tau_s(z)}{1-T'_z(z)-\tau'_s(z)\,s}z\zeta^c_z(s,z)s'_{inc}(s,z)\Big|z\right]h_z(z)\,d$$

Further differentiating this equation with respect to z^* we obtain that, at any earnings z^*,

$$\frac{\tau_s(z^*)}{1+\tau_s(z^*)}\mathbb{E}\left[s\zeta^c_{s|z}(s,z)\Big|z^*\right] \tag{245}$$
$$=\mathbb{E}\Big[\hat{g}(s,z)-\hat{g}(s,z)\,s\Big|z^*\Big]-\mathbb{E}\left[\frac{T'_z(z)+\tau'_s(z)\,s+s'_{inc}(s,z)\,\tau_s(z)}{1-T'_z(z)-\tau'_s(z)\,s}z\zeta^c_z(s,z)s'_{inc}(s,z)\Big|z^*\right]$$

Further assuming that welfare weights only depend on earnings, that $\zeta^c_{s|z}$ and ζ^c_z are constant, that η_z is negligible and that $s'_{inc}(s,z)=\frac{s}{\overline{s(z)}}s'_{inc}(z)$ as well as $\eta_{s|z}(s,z)=\frac{s}{\overline{s(z)}}\eta_{s|z}(z)$, we get

$$\frac{\tau_s(z^*)}{1+\tau_s(z^*)}\zeta^c_{s|z}\overline{s(z^*)} \tag{246}$$
$$=g(z^*)\left(1-\overline{s(z^*)}\right)+\frac{\tau_s(z^*)}{1+\tau_s(z^*)}\frac{\eta_{s|z}(z^*)}{\overline{s(z^*)}}\left[\overline{s(z^*)}-\mathbb{V}(s|z^*)-\overline{s(z^*)}^2\right]$$
$$-\mathbb{E}\left[\frac{T'_z(z)+\tau'_s(z)\,s+\frac{s}{\overline{s(z)}}s'_{inc}(z)\,\tau_s(z)}{1-T'_z(z)-\tau'_s(z)\,s}z\zeta^c_z(s,z)\frac{s}{\overline{s(z)}}s'_{inc}(z)\Big|z^*\right]$$

where we have used the fact that $\mathbb{E}\Big[s-s^2\Big|z^*\Big]=\mathbb{E}\Big[s-\left(s-\overline{s(z^*)}\right)^2-2s\overline{s(z^*)}+\overline{s(z^*)}^2\Big|z^*\Big]=\overline{s(z^*)}-\mathbb{V}(s|z^*)-\overline{s(z^*)}^2$.

An exact formula for the optimal LED savings tax in the presence of multidimensional heterogeneity, assuming that the earnings tax is optimal under multidimensional heterogeneity, is thus

$$\frac{\tau_s(z^*)}{1+\tau_s(z^*)}\overline{s(z^*)}\left[\zeta^c_{s|z}-\eta_{s|z}(z^*)\left(\frac{1}{\overline{s(z^*)}}-\frac{\mathbb{V}(s|z^*)}{\overline{s(z^*)}^2}-1\right)\right] \tag{247}$$
$$=g(z^*)\left(1-\overline{s(z^*)}\right)-\mathbb{E}\left[\frac{T'_z(z)+\tau'_s(z)\,s+\frac{s}{\overline{s(z)}}s'_{inc}(z)\,\tau_s(z)}{1-T'_z(z)-\tau'_s(z)\,s}z\zeta^c_z(s,z)\frac{s}{\overline{s(z)}}s'_{inc}(z)\Big|z^*\right].$$

D Details of Tax Systems by Country

In Table 2, we consider five categories of savings subject to various taxation regimes in different countries: (i) wealth, (ii) capital gains, (iii) property, (iv) pensions, and (v) inheritance, which are typically defined in tax codes as follows. First, wealth, which is free from taxation in most advanced economies, is defined as the aggregate value of certain classes of assets, such as real estate, stocks, and bank deposits. Next, capital gains consist of realized gains from financial and real estate invest-

ments, and include interest and dividend payments. Third, property consists of real estate holdings, such as land, private residences, and commercial properties. Fourth, for our purposes, pensions are defined as private retirement savings in dedicated accounts, excluding government transfers to retired individuals, such as Social Security in the United States. Lastly, inheritances—also known as estates—are the collections of assets bequeathed by deceased individuals to living individuals, often relatives.

For each country, we label the tax system applied to each category of savings with the types described in Table 1 or "Other," which encompasses all other tax systems. An additional common simple tax structure is a "composite" tax, in which savings and labor income are not distinguished for the purposes of taxation. Composite taxes are often applied to classes of income for which it is unclear whether the income should be considered capital income or labor income. For example, in a majority of the countries in Table 2, rental income—which requires some active participation from the recipient of the income—is subject to composite taxation.

In the subsections below, we have included additional details about the tax system in each country in Table 2. Note that we characterize tax systems that feature a flat tax on savings above an exempt amount as having a separable nonlinear tax system. In addition, when benefits are withdrawn from pension accounts, they are often subject to the same progressive tax rates as labor income. We characterize these tax systems as separable nonlinear rather than composite since benefits are generally received after retirement from the labor force when the taxpayer's income is primarily composed of savings.

Australia

- **Wealth:** No wealth tax.
- **Capital gains:** Generally a composite tax applies. Gains from certain assets are exempt or discounted.
- **Property:** At the state level, land tax rates are progressive; primary residence land is typically exempt. At the local level, generally flat taxes are assessed on property but the taxes can be nonlinear as well, depending on the locality.
- **Pensions:** A flat tax is assessed on capital gains made within the pension account. A component of pension benefits may be subject to taxation when withdrawn, in which case the lesser of a flax tax or the same progressive tax rates as apply to labor income is assessed.
- **Inheritance:** No inheritance tax.

Austria

- **Wealth:** No wealth tax.

- **Capital gains:** Generally a flat tax is assessed, with the rate depending on the type of asse; taxpayers with lower labor income can opt to apply their labor income tax rate instead. Gains from certain classes of assets are exempt.
- **Property:** Either flat or progressive tax rates are assessed on property, depending on its intended use. Rates vary by municipality.
- **Pensions:** Generally no tax on capital gains made within the pension account. Pension benefits are generally subject to the same progressive tax rates as labor income, with discounts applicable to certain types of withdrawals.
- **Inheritance:** No inheritance tax.

Canada

- **Wealth:** No wealth tax.
- **Capital gains:** For most capital gains, a discount is first applied to the gain and then the discounted gain is added to labor income and taxed progressively. For certain gains, such as interest income, no discount is applied. Lifetime exemptions up to a limit apply to gains from certain classes of assets.
- **Property:** Generally a flat tax is assessed on property, with rates varying by province and locality.
- **Pensions:** No tax on capital gains made within the pension account. Pension benefits are generally subject to the same progressive tax rates as labor income, with exemptions applicable to certain types of withdrawals.
- **Inheritance:** No separate inheritance tax. A final year tax return is prepared for the deceased, including income for that year, that treats all assets as if they have just been sold and applies the relevant taxes (e.g., labor income and capital gains taxes) accordingly.

Denmark

- **Wealth:** No wealth tax.
- **Capital gains:** Progressive taxation with two tax brackets. Gains from certain classes of assets are exempt.
- **Property:** At the national level, property is subject to progressive taxation with two tax brackets. Pensioners under an income threshold can receive tax relief. Land taxes—assessed at the local level—are flat taxes, with rates varying by municipality.

- **Pensions:** A flat tax is assessed on capital gains made within the pension account. Pension benefits are generally subject to the same progressive tax rates as labor income (excluding a labor market surtax), a flat tax, or are exempt from taxation, depending on the type of pension.
- **Inheritance:** Generally a flat tax is assessed on the inheritance above an exemption, with a higher tax rate for more distant relatives. Transfers to spouses and charities are exempt. Inheritances above a certain value are subject to additional taxes.

France

- **Wealth:** No wealth tax.
- **Capital gains:** Different rates—progressive and flat—apply to gains from different classes of assets. Certain low-income individuals are either exempt from taxes or can opt to apply their labor income tax rate, depending on the type of asset. High-income individuals are subject to a surtax. Gains from certain assets are exempt or discounted.
- **Property:** Residence taxes are assessed on property users, while property taxes on developed and undeveloped properties are assessed on owners. Rates are set at the local level and apply to the estimated rental value of the property. Exemptions, reductions, and surcharges may apply depending on the taxpayer's reference income and household composition, certain events, and property characteristics. Surcharges may also apply to higher-value properties. An additional property wealth tax applies at the national level; rates are progressive above an exemption.
- **Pensions:** Generally no tax on capital gains made within the pension account. Pension benefits beyond an exemption are generally subject to the same progressive tax rates as labor income. A flat tax is assessed on certain types of withdrawals, and special rules apply to certain types of accounts.
- **Inheritance:** Either a flat tax or progressive tax rates are assessed on the inheritance above an exemption, with rates and exemptions depending on the relation of the recipient to the deceased and their disability status. Transfers to spouses/civil partners are exempt. Certain shares are required to pass to the deceased's children.

Germany

- **Wealth:** No wealth tax.
- **Capital gains:** Generally a flat tax is assessed on gains above an exemption, but taxpayers with lower labor income can opt to apply their labor income tax rate instead. Gains from certain classes of assets are exempt or subject to special rules.

- **Property:** A flat tax is assessed on property, with rates depending on the class of property and subject to a multiplier, which varies by locality.
- **Pensions:** No tax on capital gains made within the pension account. A portion of pension benefits, which depends on the type of account, is subject to the same progressive tax rates as labor income.
- **Inheritance:** Progressive tax rates are assessed on the inheritance above an exemption, with tax rates and exemptions both depending on the relation of the recipient to the deceased. Pension entitlements are exempt.

Ireland

- **Wealth:** No wealth tax.
- **Capital gains:** A flat tax is assessed on gains above an exemption, with the rate depending on the type of asset. Certain classes of individuals, such as farmers and entrepreneurs, qualify for lower rates and additional exemptions.
- **Property:** Progressive tax rates are assessed on residential properties, with local authorities able to vary the rates to a certain extent. A flat tax is assessed on commercial properties, with rates varying by locality.
- **Pensions:** No tax on capital gains made within the pension account. Depending on the type of withdrawal, pension benefits are either subject to the same progressive tax rates as labor income or different progressive tax rates beyond an exemption. A surtax is assessed on high-value accounts.
- **Inheritance:** A flat tax is assessed on inheritances above an exemption. Exemptions are associated with the recipient and apply to the sum of all inheritances bequeathed to the recipient from certain classes of relatives.

Israel

- **Wealth:** No wealth tax.
- **Capital gains:** Generally a flat tax is assessed on real gains (i.e., the inflationary component of gains is exempt). High-income individuals are subject to a surtax.
- **Property:** Generally the tax increases in the area of the property, with amounts depending on property characteristics and varying by municipality. Tax relief may apply to certain taxpayers, such as new immigrants and low-income individuals, depending on the municipality.
- **Pensions:** No tax on capital gains made within the pension account. Pension benefits are generally subject to the same progressive tax rates as labor income; certain taxpayers qualify for exemptions.

- **Inheritance:** No inheritance tax.

Italy

- **Wealth:** A flat tax is assessed on bank deposits and financial investments held abroad, with exemptions on bank deposits if the average annual account balance is below a certain threshold.
- **Capital gains:** Generally a flat tax is assessed on financial capital gains. For certain real estate capital gains, individuals can choose between separable or composite taxation, either applying a flat tax or their labor income tax rate.
- **Property:** Generally a flat tax is assessed on property, with rates depending on property characteristics and varying by municipality.
- **Pensions:** A flat tax is assessed on capital gains made within the pension account, with the rate depending on the type of asset. Pension benefits are also subject to flat taxes, with rates varying with the duration of the contribution period.
- **Inheritance:** A flat tax is assessed on inheritances, with higher rates for more distant relatives. Different amounts of the inheritance are exempt from taxation for certain close relatives.

Japan

- **Wealth:** No wealth tax.
- **Capital gains:** A flat tax is assessed on gains from certain classes of assets, such as securities and real estate, with the rate depending on the type of asset. Progressive tax rates, composite taxation, exemptions, and discounts apply to gains from different classes of assets.
- **Property:** A flat tax is assessed on property above an exemption, with a lower rate or reduction applicable to certain types of property.
- **Pensions:** No tax on capital gains made within the pension account. Pension benefits are generally subject to progressive tax rates, with the rates depending on the type of withdrawal.
- **Inheritance:** Progressive tax rates are assessed on the inheritance above a general exemption and an exemption that depends on the relation of the recipient to the deceased and their disability status. A surtax applies to more distant relatives. Certain shares are required to pass to certain relatives.

Netherlands

- **Wealth:** A progressive, fictitious estimated return from net assets not intended for daily use is taxed at a flat rate depending on the amount above the exemption.
- **Capital gains:** Gains from a company in which an individual has a substantial stake are subject to a flat tax. Most other capital gains are not subject to taxation.
- **Property:** At the municipal level, a flat tax is assessed on property, with rates depending on property characteristics and varying by municipality. At the national level, progressive tax rates are assessed on the fictitious estimated rental values of primary residences, with substantial deductions applicable to the portion of the tax exceeding the mortgage interest deduction.
- **Pensions:** No tax on capital gains made within the pension account. Pension benefits are generally subject to the same progressive tax rates as labor income, though certain accounts with taxed contributions allow tax-free withdrawals.
- **Inheritance:** Progressive tax rates are assessed on the inheritance above an exemption, with tax rates and exemptions depending on the relation of the recipient to the deceased and their disability status. Additional exemptions apply to certain classes of assets.

New Zealand

- **Wealth:** No wealth tax.
- **Capital gains:** Capital gains from financial assets are generally either subject to composite taxation or are exempt from taxation, depending on the type of gain. Special rules apply to certain classes of assets. Capital gains from real estate are generally subject to composite taxation. Depending on transaction characteristics, gains from the sale of commercial property may be subject to an additional tax, while gains from the sale of residential property may be exempt from taxation.
- **Property:** Generally a fixed fee plus a flat tax is assessed on property, with rates set at the municipal level. Low-income individuals qualify for rebates for owner-occupied residential property.
- **Pensions:** A flat tax is assessed on capital gains made within the pension account, with the rate depending on the type of account; for certain accounts, the rate depends on the taxpayer's labor income in prior years. Pension benefits are generally exempt from taxation.
- **Inheritance:** No inheritance tax.

Norway

- **Wealth:** A flat tax is assessed on wealth above an exemption, with the value of certain classes of assets, such as primary and secondary residences, discounted.
- **Capital gains:** A flat tax is assessed on gains from financial assets above the "risk-free" return (i.e., the counterfactual return on treasury bills of the same value). Gains from certain financial assets, such as dividends, are multiplied by a factor before the tax is assessed. A flat tax is assessed on real estate gains, with exemptions for certain types of property.
- **Property:** A flat tax is assessed on discounted property values, with rates varying by municipality and discounts varying by property type.
- **Pensions:** No tax on capital gains made within the pension account. Pension benefits are generally subject to a lower tax rate than labor income, and taxpayers with smaller benefits qualify for larger tax deductions.
- **Inheritance:** No inheritance tax.

Portugal

- **Wealth:** No wealth tax.
- **Capital gains:** Generally a flat tax is assessed on gains from financial assets, but for certain types of gains, such as interest, low-income individuals can opt to apply their labor income tax rate. For real estate capital gains, a discount is first applied to the gain and then the discounted gain is added to labor income and taxed progressively. Certain classes of real estate are exempt.
- **Property:** Progressive tax rates are assessed on property, with exemptions for certain taxpayers. Rates and exemptions vary based on property characteristics, and an additional exemption applies to low-income individuals.
- **Pensions:** No tax on capital gains made within the pension account, except for dividends, which are generally subject to a flat tax. For different types of withdrawals above an exemption, capital gains are either subject to a flat tax or the same progressive tax rates as labor income when withdrawn. Depending on how contributions were initially taxed and the type of withdrawal, the non-capital gains component of benefits is exempt from taxation, or subject to a flat tax or the same progressive tax rates as labor income on the amount above an exemption.
- **Inheritance:** A flat tax is assessed on the inheritance, with a higher rate for real estate transfers. Transfers to spouses/civil partners, ascendants, and descendants are exempt (except for real estate transfers, which are subject to a low flat tax).

Singapore

- **Wealth:** No wealth tax.
- **Capital gains:** Most capital gains are not subject to taxation. Depending on transaction characteristics, composite taxation may apply.
- **Property:** Progressive tax rates are assessed on the estimated rental value of the property, with rates varying by property type and occupancy status.
- **Pensions:** No tax on capital gains made within the pension account. Pension benefits are generally subject to the same progressive tax rates as labor income; benefits from contributions made before a certain year are exempt from taxation.
- **Inheritance:** No inheritance tax.

South Korea

- **Wealth:** No wealth tax.
- **Capital gains:** Various flat and progressive tax rates are assessed on gains above an exemption; rates and exemptions depend on the type of asset. Gains from certain classes of assets are entirely exempt. Dividends and interest are subject to flat taxation below a certain limit and composite taxation above that limit.
- **Property:** Progressive tax rates are assessed on property, with rates varying by property type.
- **Pensions:** No tax on capital gains made within the pension account. Pension benefits beyond a progressive exemption (i.e, greater portions are exempt at smaller benefit levels) are generally subject to the same progressive tax rates as labor income; the exempt amount may also depend on the type of withdrawal and taxpayer characteristics.
- **Inheritance:** Progressive tax rates are assessed on the inheritance above either a lump-sum or itemized deduction, which depends on the composition of the inheritance and relation of the recipient to the deceased. Transfers to spouses are exempt. The top tax rate increases for controlling shares in a company.

Spain

- **Wealth:** Progressive tax rates are assessed on net assets above an exemption, with an additional exemption for residences.
- **Capital gains:** Progressive tax rates are generally assessed on gains, with exemptions for elderly individuals under certain conditions and for certain real estate gains.

- **Property:** Generally a flat tax is assessed on property, with rates depending on the property type and varying by locality. Exemptions or discounts may apply depending on taxpayer and property characteristics, including taxpayer income.
- **Pensions:** No tax on capital gains made within the pension account. Pension benefits are subject to the same progressive tax rates as labor income.
- **Inheritance:** Progressive tax rates are assessed on the inheritance above an exemption, with tax rates and exemptions depending on the relation of the recipient to the deceased and their disability status. Certain classes of assets, such as family businesses and art collections, are eligible for additional exemptions.

Switzerland

- **Wealth:** A flat tax is assessed on the net value of certain classes of assets and liabilities, with tax rates and exemptions varying by canton.
- **Capital gains:** Progressive tax rates are assessed on gains from real estate, with rates varying by canton. Most capital gains from financial assets are not subject to taxation. Dividends and interest are subject to composite taxation.
- **Property:** Generally a flat tax is imposed on property, with rates varying by canton; a minimum amount per property may apply. For owner-occupied properties not rented out, an estimated rental value is subject to composite taxation.
- **Pensions:** No tax on capital gains made within the pension account. Pension benefits are subject to either the same progressive tax rates as labor income or lower progressive tax rates, depending on the type of withdrawal.
- **Inheritance:** In most cantons, progressive tax rates are assessed on the inheritance and depend on the relation of the recipient to the deceased. Transfers to spouses and children are exempt in most cantons.

Taiwan

- **Wealth:** No wealth tax.
- **Capital gains:** Most capital gains from financial assets are subject to composite taxation; taxpayers can opt for a flat tax to be assessed on dividends, and certain gains are exempt from taxation. A flat tax is assessed on gains from real estate, with the rate depending on the type of asset, and an exemption for primary residences.
- **Property:** Flat or progressive tax rates are assessed on land, depending on its intended use. A flat tax is generally assessed on buildings, with rates depending on their intended use. Certain classes of land and buildings are exempt or subject to reduced rates.

- **Pensions:** No tax on capital gains made within the pension account. Pension benefits beyond an exemption—which depends on the duration of the contribution period—are subject to the same progressive tax rates as labor income.
- **Inheritance:** Progressive tax rates are assessed on the inheritance above an exemption, which depends on the relation of the recipient to the deceased, their disability status, and their age.

United Kingdom

- **Wealth:** No wealth tax.
- **Capital gains:** Either flat or progressive tax rates are assessed on gains, with rates depending on the taxpayer's labor income tax bracket; higher rates generally apply to taxpayers in higher labor income tax brackets. Exemptions for part or all of the gain apply to certain types of assets, such as dividends and primary residences.
- **Property:** Progressive tax rates are assessed on property, with rates varying by locality. Exemptions or discounts may apply to certain taxpayers depending on characteristics, such as age.
- **Pensions:** No tax on capital gains made within the pension account. Pension benefits beyond an exemption are subject to the same progressive tax rates as labor income. An additional flat tax may be imposed on accounts with a value exceeding a lifetime limit, with the tax rate depending on the type of withdrawal.
- **Inheritance:** A flat tax is assessed on the inheritance above an exemption, with larger exemptions for transfers to children. Transfers to spouses/civil partners, charities, and amateur sports clubs are exempt. The tax rate is reduced if a certain share is transferred to charity.

United States

- **Wealth:** No wealth tax.
- **Capital gains:** Gains from "short-term" assets (held for less than a year) are subject to composite taxation. Gains from "long-term" assets are subject to a flat tax, with higher rates for higher-income individuals. Dividends are also subject to either composite taxation or flat taxes that increase with labor income, depending on their source.
- **Property:** Generally a flat tax is assessed on property, with rates varying by state, county, and municipality.
- **Pensions:** No tax on capital gains made within the pension account. Depending on the type of account, benefits are generally either exempt from taxation or subject to the same progressive tax rates as labor income.

- **Inheritance:** Progressive tax rates are assessed on the inheritance above an exemption. Transfers to spouses are generally exempt.