Online Appendix

Sufficient Statistics for Nonlinear Tax Systems With General Across-Income Heterogeneity

Antoine Ferey, Benjamin B. Lockwood, and Dmitry Taubinsky

Table of Contents

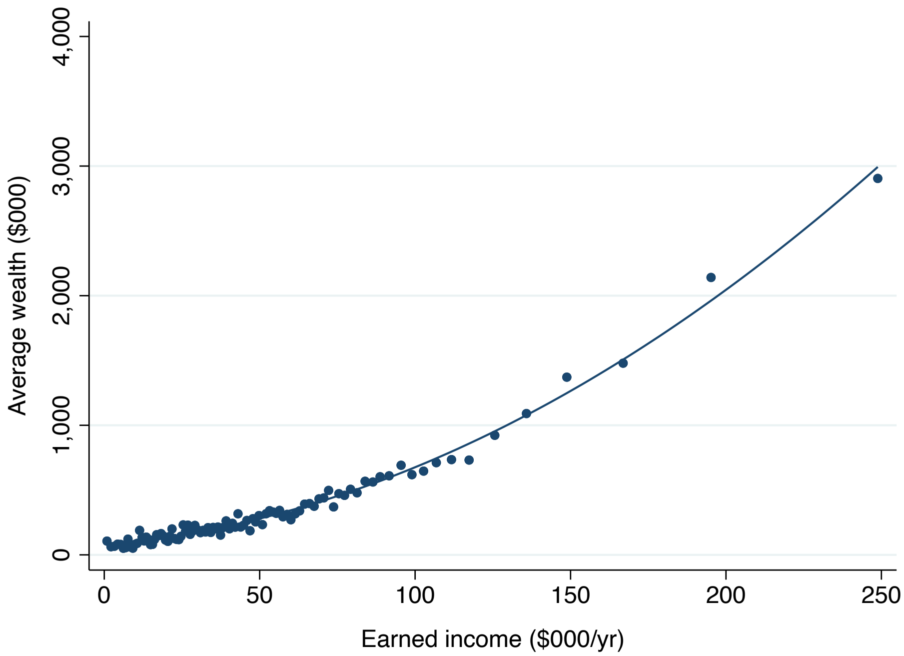
A Supplementary Tables & Figures

Table A1: 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

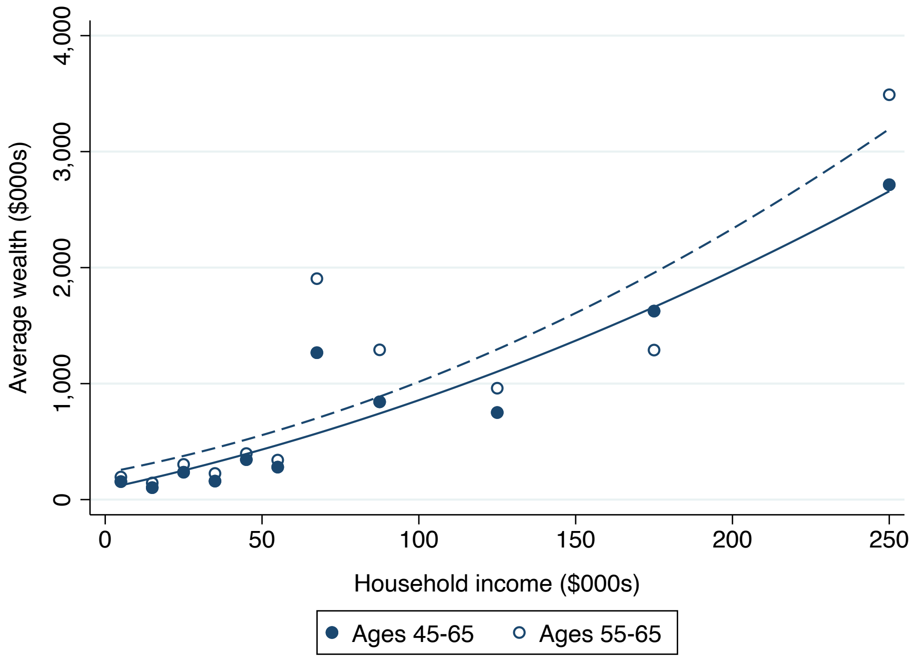
Notes: This table classifies tax systems applied to different savings vehicles across countries in 2020 according to the types of simple tax systems we consider.

Figure A1: Savings Across Incomes 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 among people age 45-65. See Appendix E.A for further details.

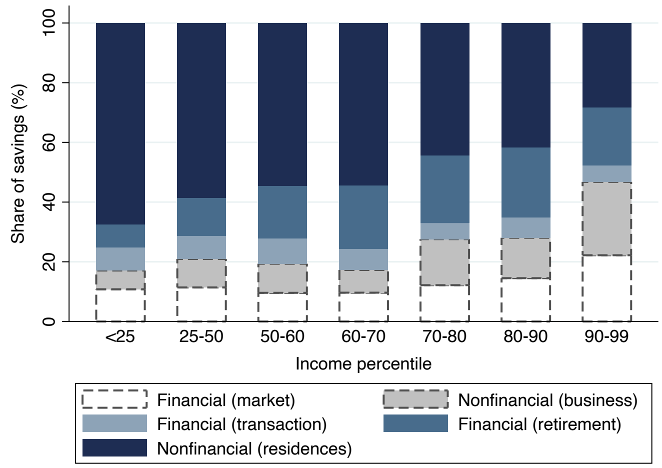
Figure A2: Savings Across Incomes in the United States



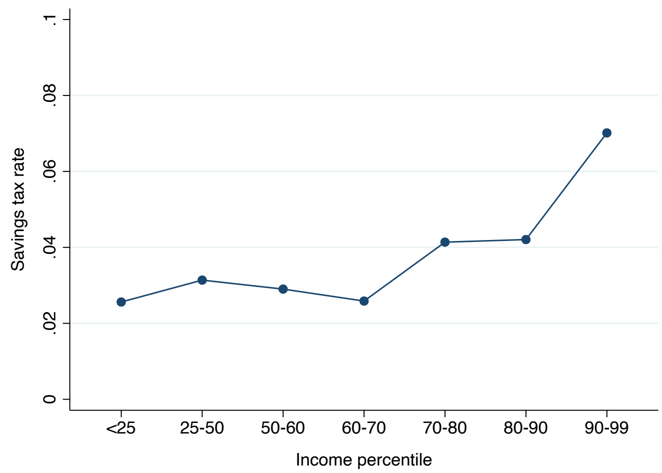
Notes: This figure displays average savings within each earnings bin reported in the Federal Reserve Bank of New York's Survey of Consumer Expectations. Solid points depict averages among respondents age 45-65, corresponding to the middle age bin PSZ. Hollow points depict averages among those aged 55 to 65. See Appendix E.A for further details.

Figure A3: Calibration of Savings Tax Rates Across Incomes in the U.S.

(a) Decomposition of Savings Types: Bricker et al. (2019)



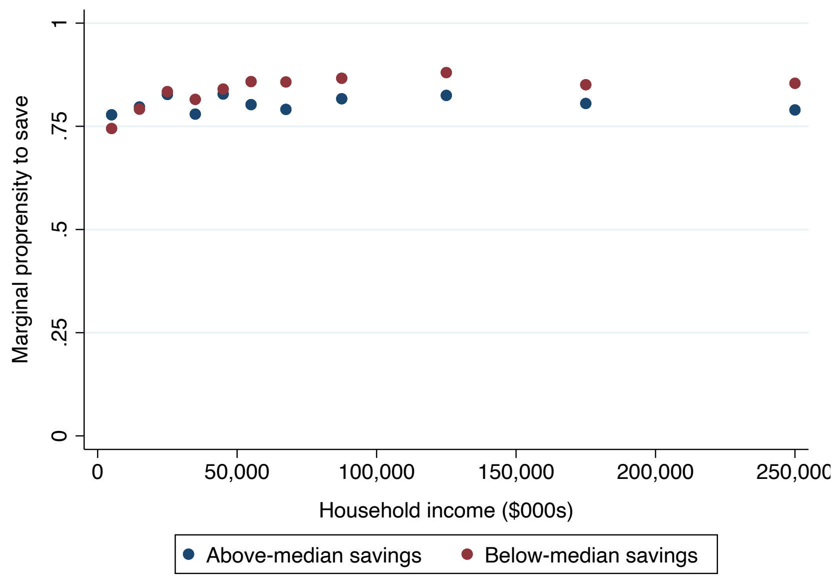
(b) Calibrated Savings Tax Rates in the United States, by Income Percentile



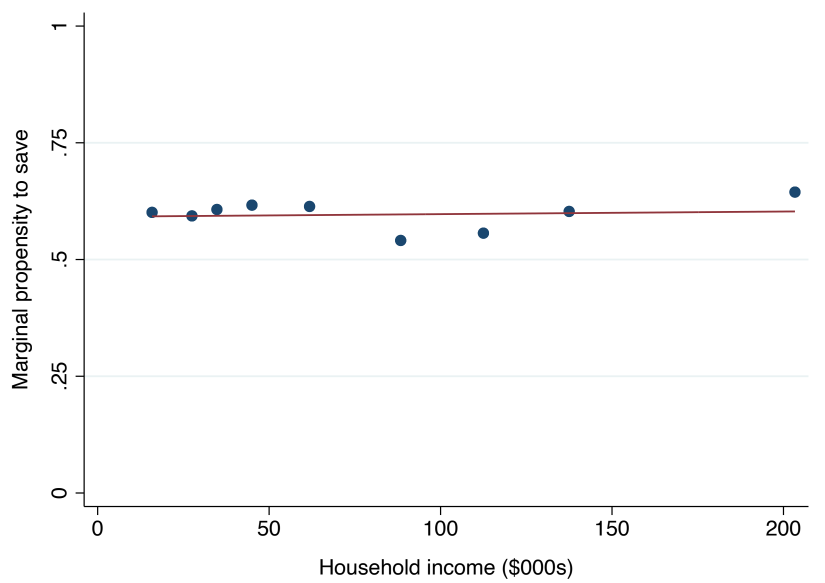
Notes: This figure illustrates the calibration of savings tax rates in the U.S. across the income distribution. Panel (a) plots the composition of asset types in individuals' portfolios across the income distribution, reported by Bricker et al. (2019). Panel (b) plots the implied weighted average savings tax rate in each bin. See Appendix E.A.2 for details.

Figure A4: Marginal Propensities to Save Across Incomes

(a) MPS Out of Net-of-Tax Income (Survey of Consumer Expectations)

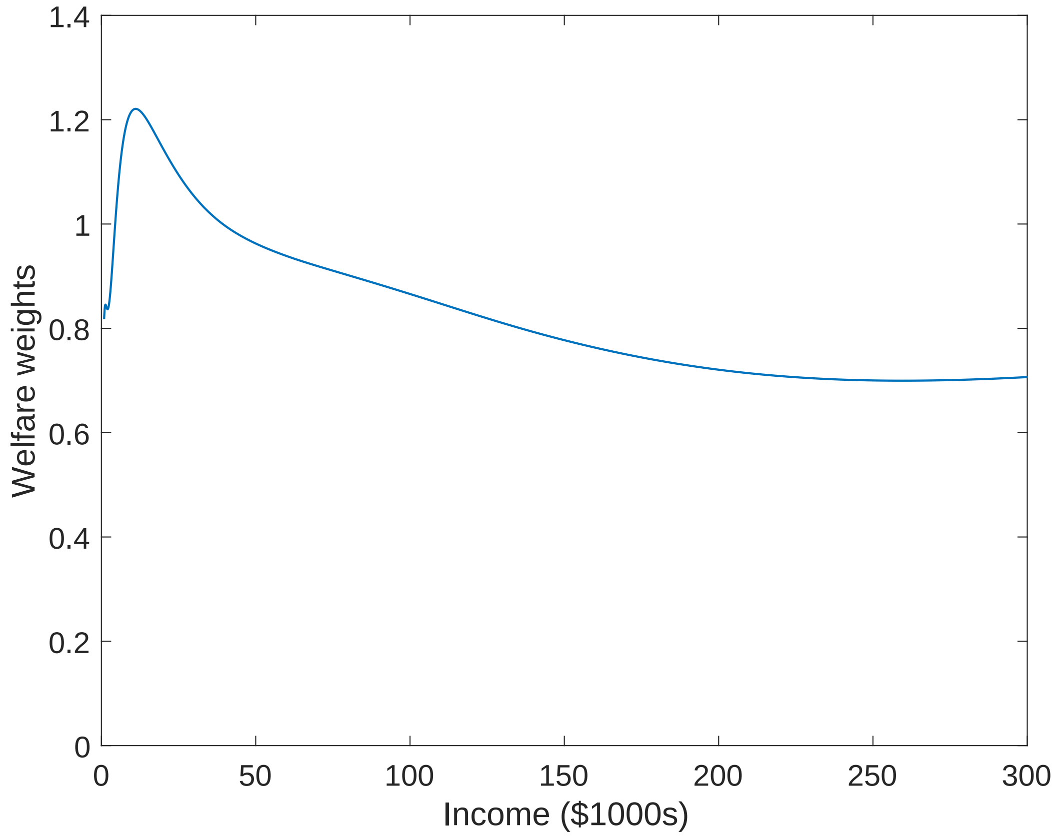


(b) MPS Out of Labor Earnings (Own Survey)



Notes: This figure reports two measures of the marginal propensity to save out of additional income in the U.S. based on survey evidence. In panel (a), the MPS is computed from the Survey of Consumer Expectations. In panel (b), the MPS is computed from the answers to our survey question.

Figure A5: 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. These weights are computed under the assumption that the savings tax is the Pareto-efficient SN schedule reported in Figure 3.

B Supplementary Theoretical Results

B.A Structural characterization of s'_{inc}

In economies with preference heterogeneity, budget heterogeneity, and auxiliary choices, individuals solve

$$\max_{c,s,z,\chi} U(c, \phi_s(s,z,\chi;\theta), \phi_z(s,z,\chi;\theta), \chi;\theta) \; s.t. \; c \leq B(s,z,\chi;\theta) - \mathcal{T}(s,z)$$
$$\iff \max_z \left\{ \max_s \left[\max_\chi U(B(s,z,\chi;\theta) - \mathcal{T}(s,z), \phi_s(s,z,\chi;\theta), \phi_z(s,z,\chi;\theta), \chi;\theta) \right] \right\}. \quad (45)$$

We denote by $\chi(s,z;\theta)$ the solution to the inner problem, $s(z;\theta)$ the solution to the intermediate problem, and $z(\theta)$ the solution to the outer problem. We assume that $\chi(s,z;\theta)$ and $s(z;\theta)$ are interior solutions that satisfy the first-order conditions of these problems.

To keep things tractable, we assume unidimensional heterogeneity in types and maintain the assumption that $z(\theta)$ is strictly increasing to denote $\vartheta(z)$ the type that chooses earnings z. In this setting, we decompose across-income heterogeneity in $s(z) := s(z;\vartheta(z))$ between $s'_{inc}(z) := \frac{\partial s(z;\vartheta(z'))}{\partial z}|_{z'=z}$ and $s'(z) - s'_{inc}(z) := \frac{\partial s(z';\vartheta(z))}{\partial z}|_{z'=z}$ as follows:

Proposition A1. *In economies with preference heterogeneity, budget heterogeneity, and auxiliary choices, sufficient* *statistics $s'_{inc}(z)$ and $s'(z) - s'_{inc}(z)$ are given by*

$$s'_{inc}(z) = -\frac{\mathcal{N}^1_{inc}(z) + \mathcal{N}^2_{inc}(z)}{\mathcal{D}^1(z) + \mathcal{D}^2(z)} \tag{46}$$

$$s'(z) - s'_{inc}(z) = -\frac{\mathcal{N}^1_{het}(z) + \mathcal{N}^2_{het}(z)}{\mathcal{D}^1(z) + \mathcal{D}^2(z)} \tag{47}$$

where terms in the numerators and denominators are

$$\mathcal{N}^1_{inc} := \mathcal{K}_c\Big[B'_z + B'_\chi\frac{\partial\chi}{\partial z} - \mathcal{T}'_z\Big] + \mathcal{K}_s\Big[\frac{\partial\phi_s}{\partial z} + \frac{\partial\phi_s}{\partial\chi}\frac{\partial\chi}{\partial z}\Big] + \mathcal{K}_z\Big[\frac{\partial\phi_z}{\partial z} + \frac{\partial\phi_z}{\partial\chi}\frac{\partial\chi}{\partial z}\Big] + \mathcal{K}_\chi\frac{\partial\chi}{\partial z} \tag{48}$$

$$\mathcal{N}^2_{inc} := U'_c\Big[B''_{sz} + B''_{s\chi}\frac{\partial\chi}{\partial z} - \mathcal{T}''_{sz}\Big] + U'_s\Big[\frac{\partial^2\phi_s}{\partial s\partial z} + \frac{\partial^2\phi_s}{\partial s\partial\chi}\frac{\partial\chi}{\partial z}\Big] + U'_z\Big[\frac{\partial^2\phi_z}{\partial s\partial z} + \frac{\partial^2\phi_z}{\partial s\partial\chi}\frac{\partial\chi}{\partial z}\Big] \tag{49}$$

$$\mathcal{N}^1_{het} := \mathcal{K}_c B'_\chi\frac{\partial\chi}{\partial\theta} + \mathcal{K}_s\Big[\frac{\partial\phi_s}{\partial\chi}\frac{\partial\chi}{\partial\theta} + \frac{\partial\phi_s}{\partial\theta}\Big] + \mathcal{K}_z\Big[\frac{\partial\phi_z}{\partial\chi}\frac{\partial\chi}{\partial\theta} + \frac{\partial\phi_z}{\partial\theta}\Big] + \mathcal{K}_\chi\frac{\partial\chi}{\partial\theta} + \mathcal{K}_\theta \tag{50}$$

$$\mathcal{N}^2_{het} := U'_c B''_{s\chi}\frac{\partial\chi}{\partial\theta} + U'_s\Big[\frac{\partial^2\phi_s}{\partial s\partial\chi}\frac{\partial\chi}{\partial\theta} + \frac{\partial^2\phi_s}{\partial s\partial\theta}\Big] + U'_z\Big[\frac{\partial^2\phi_z}{\partial s\partial\chi}\frac{\partial\chi}{\partial\theta} + \frac{\partial^2\phi_z}{\partial s\partial\theta}\Big] \tag{51}$$

$$\mathcal{D}^1 := \mathcal{K}_c\Big[B'_s + B'_\chi\frac{\partial\chi}{\partial s} - \mathcal{T}'_s\Big] + \mathcal{K}_s\Big[\frac{\partial\phi_s}{\partial s} + \frac{\partial\phi_s}{\partial\chi}\frac{\partial\chi}{\partial s}\Big] + \mathcal{K}_z\Big[\frac{\partial\phi_z}{\partial s} + \frac{\partial\phi_z}{\partial\chi}\frac{\partial\chi}{\partial s}\Big] + \mathcal{K}_\chi\frac{\partial\chi}{\partial s} \tag{52}$$

$$\mathcal{D}^2 := U'_c\Big[B''_{ss} + B''_{s\chi}\frac{\partial\chi}{\partial s} - \mathcal{T}''_{ss}\Big] + U'_s\Big[\frac{\partial^2\phi_s}{(\partial s)^2} + \frac{\partial^2\phi_s}{\partial s\partial\chi}\frac{\partial\chi}{\partial s}\Big] + U'_z\Big[\frac{\partial^2\phi_z}{(\partial s)^2} + \frac{\partial^2\phi_z}{\partial s\partial\chi}\frac{\partial\chi}{\partial s}\Big] \tag{53}$$

with all quantities being evaluated at z, $s(z)$, $\vartheta(z)$, $\chi(z) := \chi(s(z)), z; \vartheta(z))$, as well as $c(z) := B(s(z), z, \chi(z); \vartheta(z)) - \mathcal{T}(s(z), z)$, and where

$$\mathcal{K}_c := U''_{cc}\left(B'_s - \mathcal{T}'_s\right) + U''_{cs}\frac{\partial\phi_s}{\partial s} + U''_{cz}\frac{\partial\phi_z}{\partial s} \tag{54}$$

$$\mathcal{K}_s := U''_{cs}\left(B'_s - \mathcal{T}'_s\right) + U''_{ss}\frac{\partial\phi_s}{\partial s} + U''_{sz}\frac{\partial\phi_z}{\partial s} \tag{55}$$

$$\mathcal{K}_z := U''_{cz}\left(B'_s - \mathcal{T}'_s\right) + U''_{sz}\frac{\partial\phi_s}{\partial s} + U''_{zz}\frac{\partial\phi_z}{\partial s} \tag{56}$$

$$\mathcal{K}_\chi := U''_{c\chi}\left(B'_s - \mathcal{T}'_s\right) + U''_{s\chi}\frac{\partial\phi_s}{\partial s} + U''_{z\chi}\frac{\partial\phi_z}{\partial s} \tag{57}$$

$$\mathcal{K}_\theta := U''_{c\theta}\left(B'_s - \mathcal{T}'_s\right) + U''_{s\theta}\frac{\partial\phi_s}{\partial s} + U''_{z\theta}\frac{\partial\phi_z}{\partial s}. \tag{58}$$

Numerators of $s'_{inc}(z)$ and $s'(z) - s'_{inc}(z)$ are different as they capture direct changes in s, coming from either a change in z or a change in $\vartheta(z)$. Denominators are the same because they capture processes of circular adjustments induced by direct changes in s.

In a simple setting like example (1) with additively separable utility, a separable tax system and preference heterogeneity for s only, the only non-zero term in the numerator of $s'_{inc}(z)$ would be proportional to $\mathcal{K}_c$ capturing changes in the marginal utility of c from changes in z, and the only non-zero term in the numerator of $s'(z) - s'_{inc}(z)$ would be proportional to $\mathcal{K}_\theta$ capturing changes in marginal utility of s from changes in θ.

B.B Optimal Simple Taxes on s, Suboptimal Taxes on z

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.

Unidimensional heterogeneity. When heterogeneity is unidimensional we have the following characterization:

Proposition A2. *Consider a SL/SN/LED tax system with a potentially suboptimal earnings tax $T_z(z)$ and an optimal tax on s (given T_z) for which Conditions 1 and 2-UD hold. Suppose also that in the SN system s is strictly monotonic* *in z. At each bundle* (c^0, s^0, z^0) *chosen by a type* θ*, this tax system satisfies the following optimality conditions for taxes on s:*

$$SL: \quad \frac{\tau_s}{1+\tau_s}\bar{s}\overline{\zeta^c_{s|z}} = \int_z \left\{ s(z)\left(1-\hat{g}(z)\right) - \frac{T_z'(z) + s'_{inc}(z)\,\tau_s}{1-T_z'(z)} z\zeta_z^c(z)\, s'_{inc}(z) \right\} dH_z(z) \tag{59}$$

$$SN: \quad \frac{T_s'(s^0)}{1+T_s'(s^0)} s^0 \zeta^c_{s|z}(s^0) h_s(s^0) = \int_{s\geq s^0} \left(1-\hat{g}(s)\right) dH_s(s) - \frac{T_z'(z^0) + s'_{inc}(z^0) T_s'(s^0)}{1-T_z'(z^0)} z^0 \zeta_z^c(z^0) s'_{inc}(z^0) h_s(s^0) \tag{60}$$

$$LED: \int_{z\geq z^0} \frac{\tau_s(z)}{1+\tau_s(z)} s(z)\zeta^c_{s|z}(z)\, dH_z(z) = \int_{z\geq z^0} \left\{ s(z)\left(1-\hat{g}(z)\right) - \frac{T_z'(z) + \tau_s'(z)s(z) + s'_{inc}(z)\,\tau_s(z)}{1-T_z'(z) - \tau_s'(z)s(z)} z\zeta_z^c(z) s'_{inc}(z) \right\} dH_z(z) - \frac{T_z'(z^0) + \tau_s'(z^0)s^0 + s'_{inc}(z^0)\tau_s(z^0)}{1-T_z'(z^0) - \tau_s'(z)s^0} z^0 \zeta_z^c(z^0)\, s^0\, h_z(z^0) \tag{61}$$

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 sufficient statistic $s'(z) - s'_{inc}(z)$ no longer appears in the formulas. The intuition is that outside of the full optimum, it may still be desirable to tax savings when the earnings tax is suboptimal, although this clashes with Pareto efficiency. Second, $s'_{inc}(z)$ is a key sufficient statistic for optimal savings tax schedules. Indeed, by Lemma 1, a larger $s'_{inc}(z)$ means that savings tax reforms impose higher distortions on earnings and thus generally calls for lower savings tax rate.

Multidimensional heterogeneity. With multidimensional heterogeneity, we have the following characterization:

Proposition A3. *Consider a SL/SN/LED tax system with a potentially suboptimal earnings tax* $T_z(z)$ *and an optimal tax on s (given* T_z*) for which Conditions 1 and 2-MD hold. Suppose also that in the SN system s is strictly monotonic in z. At each bundle* (c^0, s^0, z^0) *chosen by a type* θ*, this tax system satisfies the following optimality conditions for taxes on s:*

$$SL: \quad \frac{\tau_s}{1+\tau_s}\bar{s}\overline{\zeta^c_{s|z}} = \int_z \left\{ \mathbb{E}\left[s\cdot(1-\hat{g}) \middle| z\right] - \mathbb{E}\left[FE_z \cdot s'_{inc} \middle| z\right] \right\} dH_z(z). \tag{62}$$

$$SN: \quad \frac{T_s'(s^0)}{1+T_s'(s^0)} \mathbb{E}[s\zeta^c_{s|z}|s^0] h_s(s^0) = \int_{s\geq s^0} \mathbb{E}\left[1-\hat{g}\middle| s\right] dH_s(s) - \mathbb{E}\left[FE_z\cdot s'_{inc}\middle| s^0\right] h_s(s^0) \tag{63}$$

$$LED: \int_{z\geq z^0} \frac{\tau_s(z)}{1+\tau_s(z)} \mathbb{E}[s\zeta^c_{s|z}|z] dH_z(z) = \int_{z\geq z^0} \mathbb{E}\left[(1-\hat{g})\cdot s\middle| z\right] dH_z(z) - \mathbb{E}\left[FE_z\cdot s\middle| z^0\right] h_z(z^0) - \int_{z\geq z^0} \mathbb{E}\left[FE_z\cdot s'_{inc}\middle| z\right] dH_z(z). \tag{64}$$

B.C Optimal Simple Taxes on z with Multidimensional Heterogeneity

Proposition A4. *Consider a SL/SN/LED tax system with a potentially suboptimal tax on s and an optimal earnings tax* $T_z(z)$ *(given the tax on s) for which Conditions 1 and 2-MD hold. Suppose also that in the SN system s is strictly monotonic in z. At each bundle* (c^0, s^0, z^0) *chosen by a type* θ*, this tax system satisfies the following optimality* *conditions for earnings taxes:*

$$SL: \frac{T_z'\left(z^0\right)}{1-T_z'\left(z^0\right)}\mathbb{E}\left[\zeta_z^c(s,z)\middle|z^0\right] = \frac{1}{z^0 h_z(z^0)}\int_{z\geq z^0}\left\{\mathbb{E}\left[1-\hat{g}\left(s,z\right)|z\right]\right\}dH_z(z) - \mathbb{E}\left[s_{inc}'(s,z)\frac{\tau_s}{1-T_z'\left(z\right)}\zeta_z^c(s,z)\middle|z^0\right] \tag{65}$$

$$SN: \frac{T_z'\left(z^0\right)}{1-T_z'\left(z^0\right)}\mathbb{E}\left[\zeta_z^c(s,z)\middle|z^0\right] = \frac{1}{z^0 h_z(z^0)}\int_{z\geq z^0}\left\{\mathbb{E}\left[1-\hat{g}\left(s,z\right)|z\right]\right\}dH_z(z) - \mathbb{E}\left[s_{inc}'(s,z)\frac{T_s'\left(s\right)}{1-T_z'\left(z\right)}\zeta_z^c(s,z)\middle|z^0\right] \tag{66}$$

$$LED: \mathbb{E}\left[\frac{T_z'\left(z\right)+\tau_s'\left(z\right)s}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s}\zeta_z^c(s,z)\middle|z^0\right] = \frac{1}{z^0 h_z\left(z^0\right)}\int_{z\geq z^0}\left\{\mathbb{E}\left[1-\hat{g}\left(s,z\right)|z\right]\right\}dH_z(z) - \mathbb{E}\left[s_{inc}'(s,z)\frac{\tau_s\left(z\right)}{1-T_z'\left(z\right)-\tau_s'\left(z\right)s}\zeta_z^c(s,z)\middle|z^0\right] \tag{67}$$

B.D Multiple Goods

Suppose individuals consume $n+1$ goods, c and $\boldsymbol{s} = (s_1, s_2, ..., s_n)$. For example, $\boldsymbol{s}$ might correspond to different categories of saving, which the government might choose to tax in different ways. 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 individual of type 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 by $\zeta_{s_i|z}^c(\theta)$, and we define cross-substitution elasticities by $\xi_{s_{j,i}|z}^c(\theta) := -\frac{\mathcal{T}_{s_i}'(\boldsymbol{s}(z;\theta),z)}{s_j(z;\theta)}\frac{\partial s_j(z;\theta)}{\partial \mathcal{T}_{s_i}'(\boldsymbol{s}(z;\theta),z)}\Big|_{z=z(\theta)}$, where $s_j(z;\theta)$ denotes type θ consumption of good j when earning labor income z. We denote causal income effects on good s_j by $s_{j,inc}'(\theta) := \frac{\partial s_j(z;\theta)}{\partial z}\Big|_{z=z(\theta)}$. We continue using $\hat{g}(z)$ to denote the social marginal welfare effect of increasing a z-earner's consumption of c by one unit.[43]

Proposition A5. *Consider an optimal tax system for which Conditions 1 and 2-UD hold. At each bundle (c^0, s^0, z^0) chosen by a type θ, this tax system satisfies:*

$$\frac{\mathcal{T}_{s_i}'(\boldsymbol{s^0}, z^0)}{1+\mathcal{T}_{s_i}'(\boldsymbol{s^0}, z^0)} = \left(s_i'(z^0) - s_{i,inc}'(z^0)\right)\frac{1}{s_i\zeta_{s_i|z}^c(z^0)}\frac{1}{h_z(z^0)}\int_{z\geq z^0}\left(1-\hat{g}(z)\right)dH_z(z) + \underbrace{\sum_{j\neq i}\frac{\mathcal{T}_{s_j}'(\boldsymbol{s^0}, z^0)}{\mathcal{T}_{s_i}'(\boldsymbol{s^0}, z^0)}\frac{s_j^0\xi_{s_{j,i}|z}^c(z^0)}{s_i^0\zeta_{s_i|z}^c(z^0)}}_{\text{Tax diversion ratio}}. \tag{68}$$

Consider a Pareto-efficient tax system for which Conditions 1 and 2-UD hold. At each bundle (c^0, s^0, z^0) chosen by a type θ, this tax system satisfies:

$$\frac{\mathcal{T}_{s_i}'(\boldsymbol{s^0}, z^0)}{1+\mathcal{T}_{s_i}'(\boldsymbol{s^0}, z^0)} = \left(s_i'(z^0) - s_{i,inc}'(z^0)\right)\frac{z^0\zeta_z^c(z^0)}{s_i^0\zeta_{s_i|z}^c(z^0)}\frac{\mathcal{T}_z'(\boldsymbol{s^0}, z^0) + \sum_{j=1}^n s_{j,inc}'(z^0)\mathcal{T}_{s_j}'(\boldsymbol{s^0}, z^0)}{1-\mathcal{T}_z'(\boldsymbol{s^0}, z^0)} + \underbrace{\sum_{j\neq i}\frac{\mathcal{T}_{s_j}'(\boldsymbol{s^0}, z^0)}{\mathcal{T}_{s_i}'(\boldsymbol{s^0}, z^0)}\frac{s_j^0\xi_{s_{j,i}|z}^c(z^0)}{s_i^0\zeta_{s_i|z}^c(z^0)}}_{\text{Tax diversion ratio}}. \tag{69}$$

Proposition A5 features all of the familiar terms of Theorem 1, and includes a novel term that captures the tax implications of substitution effects between the different goods. Intuitively, substituting from s_i to higher-taxed goods

[43] The formula for $\hat{g}(z)$ in this more general setting is in Appendix C.H.

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 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.

B.E 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, 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 tax 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 $\tilde{s}_g(s, z)$ implicitly to satisfy the equation

$$\tilde{s}_g - T_2(\tilde{s}_g, z) = s$$

Note that $\tilde{s}_g$ is generally uniquely defined if we have a system with monotonic realized savings s. Then, the equivalence in tax schedules is given by

$$\mathcal{T}'_s(s, z) = \frac{1}{1+r} \frac{\partial}{\partial s_g} T_2(s_g, z)|_{s_g = \tilde{s}_g} \frac{\partial}{\partial s} \tilde{s}_g \tag{70}$$

and $\mathcal{T}'_z = T'_z$. equation (70) 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 $\tilde{s}_g$ gives

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

and thus

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

from which it follows that

$$\mathcal{T}'_s(s, z) = \frac{1}{1+r} \frac{\frac{\partial}{\partial s_g} T_2(s_g, z)|_{s_g = \tilde{s}_g}}{1 - \frac{\partial}{\partial s_g} T_2(s_g, z)|_{s_g = \tilde{s}_g}}. \tag{71}$$

We can also start with a schedule $\mathcal{T}$ and converts it to the two-period 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 $\tilde{s}(s_g, z)$ to satisfy

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

Then,

$$\begin{aligned} \frac{\partial}{\partial s_g} T_2(s_g, z) &= (1+r)\mathcal{T}'_s(\tilde{s}, z) \frac{\partial}{\partial s_g} \tilde{s} \\ &= \frac{(1+r)\mathcal{T}'_s(\tilde{s}, z)}{1 + (1+r)\mathcal{T}'_s(\tilde{s}, z)} \end{aligned} \tag{72}$$

B.E.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 (72) 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

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

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

B.E.2 Linear tax systems (LED and SL).

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

C Proofs

C.A Proof of Lemma 1 (Earnings Responses to Taxes on s)

Throughout the paper, we characterize earnings responses to (different) savings tax reforms using generalizations of Lemma 1 in Saez (2002). The robust insight in all cases is that a $\Delta\tau$ increase in the marginal tax rate on s induces the same earnings changes (through substitution effects) as a $s_{inc}'(z)\Delta\tau$ increase in earnings tax rate. This is what appears in the body of the text as Lemma 1. In our appendix proofs we use versions of this result that pertains to reforms that have an LED, SL, or SN structure and that allow for multidimensional heterogeneity.

Let

$$V\left(\mathcal{T}(.,z), z; \theta\right) := \max_s\, U\left(z - ps - \mathcal{T}(s,z), s, z; \theta\right) \tag{73}$$

be type θ's indirect utility function at earnings z.

LED reform. Consider a tax reform $\Delta\mathcal{T}_s$ that consists in adding a linear tax rate $\Delta\tau_s\Delta z$ on s for all individuals with earnings z above z^0, phased-in over the earnings bandwidth $\left[z^0, z^0+\Delta z\right]$, with $\Delta\tau_s$ much smaller than Δz:[44]

$$\Delta\mathcal{T}_s(s,z) = \begin{cases} 0 & if\ z \le z^0 \\ \Delta\tau_s\left(z - z^0\right)s & if\ z \in [z^0, z^0+\Delta z] \\ \Delta\tau_s\,\Delta z\, s & if\ z \ge z^0+\Delta z \end{cases} \tag{74}$$

[44] This reform, which is natural to consider for LED tax systems, allows us to derive a sufficient statistics characterization of the optimal smooth tax system (Theorem 1) without the requirement that $s(z)$ is monotonic. If we instead consider an increase in the marginal savings tax rates over a certain bandwidth of savings, which is natural to consider for SN tax systems, we need this extra assumption.

We construct for each type θ a tax reform $\Delta\mathcal{T}_z^\theta$ that affects marginal earnings tax rates, and induces the same earnings response as the initial perturbation $\Delta\mathcal{T}_s$. We define this perturbation for each type θ such that at all earnings z,

$$V(\mathcal{T}(.,z)+\Delta\mathcal{T}_s(.,z),z;\theta)=V(\mathcal{T}(.,z)+\Delta\mathcal{T}_z^\theta(.,z),z;\theta). \tag{75}$$

Then, by construction, the perturbation $\Delta\mathcal{T}_z^\theta$ induces the same earnings response dz as the initial perturbation $\Delta\mathcal{T}_s$. Moreover, both tax reforms must induce the same utility change for type θ. To compute these utility changes, we make use of the envelope theorem.

For types θ with earnings $z(\theta)\in[z^0,z^0+\Delta z]$, this implies

$$\begin{aligned}&U_c'\,\Delta\tau_s\left(z-z^0\right)s\left(z;\theta\right)=U_c'\,\Delta\mathcal{T}_z^\theta(z)\\ \iff&\ \Delta\mathcal{T}_z^\theta(z)=\Delta\tau_s\left(z-z^0\right)s\left(z;\theta\right).\end{aligned} \tag{76}$$

Differentiating both sides with respect to z and letting $\Delta z\to 0$, this implies that in the phase-in region, the reform induces the same earnings change as a small increase $s_{inc}'\left(z\right)\Delta\tau_s$ in the marginal earnings tax rate.

For types θ with earnings $z(\theta)\geq z^0+\Delta z$, this implies

$$\begin{aligned}&U_c'\,\Delta\tau_s\,\Delta z\,s\left(z;\theta\right)=U_c'\,\Delta\mathcal{T}_z^\theta(z)\\ \iff&\ \Delta\mathcal{T}_z^\theta(z)=\Delta\tau_s\,\Delta z\,s\left(z;\theta\right).\end{aligned} \tag{77}$$

That is, above the phase-in region, the reform induces the same earnings changes as a $\Delta\tau_s\Delta z\,s\left(z\right)$ increase in tax liability combined with a $\Delta\tau_s\Delta z\,s_{inc}'\left(z\right)$ increase in the marginal earnings tax rate.[45]

SL reform. Consider a tax reform $\Delta\mathcal{T}_s$ that consists in adding a linear tax rate $\Delta\tau_s$ on s for all individuals. This is a special case of a LED reform. As a result, we directly obtain that this reform induces the same earnings changes as a $\Delta\tau_s\,s\left(z\right)$ increase in tax liability combined with a $\Delta\tau_s\,s_{inc}'\left(z\right)$ increase in the marginal earnings tax rate.

SN reform. Consider a tax reform $\Delta\mathcal{T}_s$ that consists in a small increase $\Delta\tau_s$ in the marginal tax rate on s in a bandwidth $\left[s^0,s^0+\Delta s\right]$, with $\Delta\tau_s$ much smaller than Δs:

$$\Delta\mathcal{T}_s(s,z)=\begin{cases}0 & if\ s\leq s^0\\ \Delta\tau_s(s-s^0) & if\ s\in[s^0,s^0+\Delta s]\\ \Delta\tau_s\Delta s & if\ s\geq s^0+\Delta s\end{cases} \tag{78}$$

We construct for each type θ a perturbation of the earnings tax $\Delta\mathcal{T}_z^\theta$ that induces the same earnings response as the initial perturbation $\Delta\mathcal{T}_s$. Suppose we define this perturbation for each type θ such that at all earnings z,

$$V(\mathcal{T}(.,z)+\Delta\mathcal{T}_s(.,z),z;\theta)=V(\mathcal{T}(.,z)+\Delta\mathcal{T}_z^\theta(.,z),z;\theta). \tag{79}$$

Then, by construction, the perturbation $\Delta\mathcal{T}_z^\theta$ induces the same earnings response dz as the initial perturbation $\Delta\mathcal{T}_s$. Moreover, both tax reforms must induce the same utility change for type θ. To compute these utility changes, we make use of the envelope theorem.

For types θ with $s(z;\theta)\in[s^0,s^0+\Delta s]$, this implies

$$\begin{aligned}&U_c'\Delta\tau_s\left(s\left(z;\theta\right)-s^0\right)=U_c'\Delta\mathcal{T}_z^\theta(z)\\ \iff&\ \Delta\mathcal{T}_z^\theta(z)=\left(s\left(z;\theta\right)-s^0\right)\Delta\tau_s.\end{aligned} \tag{80}$$

Differentiating both sides with respect to z and letting $\Delta s\to 0$, this implies that in the phase-in region, the small increase $\Delta\tau_s$ in the marginal tax rate on s induces the same earnings change as a small increase $s_{inc}'\left(z\right)\Delta\tau_s$ in the marginal earnings tax rate.

[45] This proof transparently extends to the kind of generalized LED reforms we consider in the proof of Theorem 2, where taxes are only increased for individuals with $s\geq s^0$ and the increase in tax liability is proportional to $(s-s^0)$ instead of s.

For types θ with $s(z;\theta) \geq s^0 + \Delta s$, this implies

$$U_c' \, \Delta\tau_s \, \Delta s = U_c' \, \Delta\mathcal{T}_z^\theta(z)$$
$$\iff \Delta\mathcal{T}_z^\theta(z) = \Delta\tau_s \, \Delta s. \tag{81}$$

Thus, a $\Delta\tau_s \, \Delta s$ lump-sum (savings) tax increase induces the same earnings change as a $\Delta\tau_s \, \Delta s$ lump-sum (earnings) tax increase.

C.B Proof of Theorem 1 (Smooth Tax Systems with Unidimensional Heterogeneity)

With unidimensional heterogeneity, our assumptions imply that $z(\theta)$ is a strictly increasing function, we can thus define its inverse by $\vartheta(z)$. This allows us to define consumption of good c as $c(z) := c(z;\vartheta(z))$, consumption of good s as $s(z) := s(z;\vartheta(z))$, and the planner's weights as $\alpha(z) := \alpha(\vartheta(z))$.

In this notation, the problem of the government is to maximize the Lagrangian

$$\mathcal{L} = \int_z \Big[\alpha(z)\, U\Big(c(z), s(z), z; \vartheta(z)\Big) + \lambda\Big(\mathcal{T}(s(z), z) - E\Big)\Big]\, dH_z(z), \tag{82}$$

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

C.B.1 Optimality Condition for Marginal Tax Rates on z

Reform. We consider a small SN-type reform at earnings level z^0 that consists in a small increase $\Delta\tau_z$ of the marginal tax rate on earnings in a small earnings bandwidth Δz:

$$\Delta\mathcal{T}(s,z) = \begin{cases} 0 & if\ z \leq z^0 \\ \Delta\tau_z(z - z^0) & if\ z \in [z^0, z^0 + \Delta z] \\ \Delta\tau_z \Delta z & if\ z \geq z^0 + \Delta z \end{cases} \tag{83}$$

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

- *mechanical effects*:

$$\int_{z \geq z^0} \Big(1 - \frac{\alpha(z)}{\lambda} U_c'\left(c(z), s(z), z; \vartheta(z)\right)\Big)\, \Delta\tau_z \Delta z \, dH_z(z) \tag{84}$$

- *behavioral effects from changes in z*:[46]

$$-\mathcal{T}_z'\left(s\left(z^0\right), z^0\right) \frac{z^0}{1 - \mathcal{T}_z'\left(s\left(z^0\right), z^0\right)} \zeta_z^c(z^0) \Delta\tau_z \, \Delta z h_z(z^0)$$
$$- \int_{z \geq z^0} \mathcal{T}_z'\left(s\left(z\right), z\right) \frac{\eta_z(z)}{1 - \mathcal{T}_z'\left(s\left(z\right), z\right)} \Delta\tau_z \Delta z \, dH_z(z) \tag{85}$$

[46]Note that by definition elasticity concepts include all circularities and adjustments induced by tax reforms such that changes in z and s are given by

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

- *behavioral effects from changes in s:*

$$- \mathcal{T}_s'\left(s(z^0), z^0\right) s_{inc}'(z^0) \left[\frac{z^0}{1-\mathcal{T}_z'\left(s\left(z^0\right), z^0\right)} \zeta_z^c(z^0) \Delta\tau_z\right] \Delta z h_z(z^0)$$
$$- \int_{z \geq z^0} \mathcal{T}_s'\left(s(z), z\right) \left[\frac{\eta_{s|z}(z)}{1+\mathcal{T}_s'\left(s(z), z\right)} + s_{inc}'(z) \frac{\eta_z(z)}{1-\mathcal{T}_z'\left(s\left(z\right), z\right)}\right] \Delta\tau_z \Delta z \, dH_z(z) \quad (86)$$

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

$$\frac{1}{\lambda}\frac{d\mathcal{L}}{\Delta z} = \int_{z \geq z^0} \left(1-\hat{g}(z)\right) \Delta\tau_z dH_z(z)$$
$$- \left(\mathcal{T}_z'\left(s\left(z^0\right), z^0\right) + s_{inc}'(z^0) \mathcal{T}_s'\left(s(z^0), z^0\right)\right) \frac{z^0}{1-\mathcal{T}_z'\left(s\left(z^0\right), z^0\right)} \zeta_z^c(z^0) \Delta\tau_z \, h_z(z^0) \quad (87)$$

where $\hat{g}(z)$, defined in (17), are social marginal welfare weights augmented with the fiscal externalities from income effects,

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

Optimality. A direct implication is a sufficient statistics characterization of the optimal schedule of marginal tax rates on 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

$$\frac{\mathcal{T}_z'\left(s\left(z^0\right), z^0\right)}{1-\mathcal{T}_z'\left(s\left(z^0\right), z^0\right)} = \frac{1}{\zeta_z^c(z^0)} \frac{1}{z^0 h_z(z^0)} \int_{z \geq z^0} \left(1-\hat{g}(z)\right) dH_z(z) - s_{inc}'(z^0) \frac{\mathcal{T}_s'\left(s(z^0), z^0\right)}{1-\mathcal{T}_z'\left(s\left(z^0\right), z^0\right)} \quad (88)$$

which is the optimality condition, equation (18), presented in Theorem 1.

C.B.2 Optimality Condition for Marginal Tax Rates on s

Reform. We consider a small LED-type reform that consists in adding a linear tax rate $\Delta\tau_s \Delta z$ on s for all individuals with earnings z above z^0, phased-in over the earnings bandwidth $\left[z^0, z^0 + \Delta z\right]$:[47]

$$\Delta\mathcal{T}(s, z) = \begin{cases} 0 & if\ z \leq z^0 \\ \Delta\tau_s \left(z - z^0\right) s & if\ z \in [z^0, z^0 + \Delta z] \\ \Delta\tau_s \, \Delta z \, s & if\ z \geq z^0 + \Delta z \end{cases} \quad (89)$$

Let $s^0 = s(z^0)$. We characterize the impact of this reform on the government objective function $\mathcal{L}$ as $\Delta z \to 0$. Normalizing all effects by $1/\lambda$, the reform induces

- *mechanical effects:*

$$\int_{z \geq z^0} \left(1 - \frac{\alpha(z)}{\lambda} U_c'\left(c(z), s(z), z; \theta(z)\right)\right) \Delta\tau_s \Delta z \, s\left(z\right) \, dH_z(z) \quad (90)$$

[47] We use this reform to derive a sufficient statistics characterization of the optimal smooth tax system, without the requirement that $s(z)$ is monotonic. If we instead consider an increase in the marginal savings tax rates over a certain bandwidth of savings, which is natural to consider for SN tax systems, we need this extra assumption.

- *behavioral effects from changes in z:*[48]

$$
\begin{aligned}
&-\mathcal{T}_z'\left(s^0,z^0\right)\left[\frac{z^0\zeta_z^c(z^0)}{1-\mathcal{T}_z'\left(s^0,z^0\right)}\,\Delta\tau_s s^0\right]h_z(z^0)\Delta z\\
&-\int_{z\geq z^0}\mathcal{T}_z'\left(s\left(z\right),z\right)\left[\frac{z\zeta_z^c(z)s_{inc}'(z)}{1-\mathcal{T}_z'\left(s\left(z\right),z\right)}+\frac{\eta_z(z)s(z)}{1-\mathcal{T}_z'\left(s\left(z\right),z\right)}\right]\Delta\tau_s\Delta z\,dH_z(z)
\end{aligned}
\tag{91}
$$

- *behavioral effects from changes in s:*

$$
\begin{aligned}
&-\mathcal{T}_s'\left(s^0,z^0\right)s_{inc}'(z^0)\left[\frac{z^0\zeta_z^c(z^0)}{1-\mathcal{T}_z'\left(s^0,z^0\right)}\,\Delta\tau_s\,s^0\right]h_z(z^0)\Delta z\\
&-\int_{z\geq z^0}\mathcal{T}_s'\left(s(z),z\right)\left[\frac{\zeta_{s|z}^c(z)+\eta_{s|z}(z)}{1+\mathcal{T}_s'\left(s(z),z\right)}\,s(z)+s_{inc}'(z)\left[\frac{z\zeta_z^c(z)s_{inc}'(z)}{1-\mathcal{T}_z'\left(s(z),z\right)}+\frac{\eta_z(z)s(z)}{1-\mathcal{T}_z'\left(s(z),z\right)}\right]\right]\Delta\tau_s\Delta z\,dH_z(z)
\end{aligned}
\tag{92}
$$

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

$$
\begin{aligned}
&\frac{1}{\lambda}\frac{d\mathcal{L}}{\Delta\tau_s\Delta z}\\
&=-\frac{\mathcal{T}_z'\left(s^0,z^0\right)+s_{inc}'(z^0)\mathcal{T}_s'\left(s^0,z^0\right)}{1-\mathcal{T}_z'\left(s^0,z^0\right)}z^0\zeta_z^c(z^0)\,s^0\,h_z(z^0)\\
&+\int_{z\geq z^0}\left\{\left(1-\hat{g}(z)\right)s\left(z\right)-\frac{\mathcal{T}_z'\left(s\left(z\right),z\right)+s_{inc}'(z)\mathcal{T}_s'\left(s(z),z\right)}{1-\mathcal{T}_z'\left(s\left(z\right),z\right)}z\zeta_z^c(z)s_{inc}'(z)-\frac{\mathcal{T}_s'\left(s(z),z\right)}{1+\mathcal{T}_s'\left(s(z),z\right)}s(z)\zeta_{s|z}^c(z)\right\}dH_z(z),
\end{aligned}
\tag{93}
$$

where $\hat{g}(z)$, defined in (17), are social marginal welfare weights augmented with the fiscal externalities from income effects.

Optimality. A direct implication of this result is a sufficient statistics characterization of the optimal marginal tax rates on 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 $s^0=s\left(z^0\right)$ and earnings z^0, the optimal marginal tax rate on s satisfies

$$
\begin{aligned}
&\frac{\mathcal{T}_z'\left(s^0,z^0\right)+s_{inc}'(z^0)\mathcal{T}_z'\left(s^0,z^0\right)}{1-\mathcal{T}_s'\left(s^0,z^0\right)}z^0\zeta_z^c(z^0)\,s^0\,h_z(z^0)\\
&=\int_{z\geq z^0}\left\{\left(1-\hat{g}(z)\right)s\left(z\right)-\frac{\mathcal{T}_z'\left(s(z),z\right)+s_{inc}'(z)\mathcal{T}_z'\left(s(z),z\right)}{1-\mathcal{T}_z'\left(s(z),z\right)}z\zeta_z^c(z)s_{inc}'(z)-\frac{\mathcal{T}_s'\left(s(z),z\right)}{1+\mathcal{T}_s'\left(s(z),z\right)}s(z)\zeta_{s|z}^c(z)\right\}\,dH_z(z).
\end{aligned}
\tag{94}
$$

Using equation (88) for optimal marginal tax rates on z to replace the term on the left-hand side, this formula can be rearranged as

$$
\begin{aligned}
&s(z^0)\int_{z\geq z^0}\left(1-\hat{g}(z)\right)dH_z(z)\\
&=\int_{z\geq z^0}\left\{(1-\hat{g}(z))s(z)-\frac{\mathcal{T}_z'(s(z),z)+s_{inc}'(z)\mathcal{T}_z'(s(z),z)}{1-\mathcal{T}_z'(s(z),z)}z\zeta_z^c(z)s_{inc}'(z)-\frac{\mathcal{T}_s'(s(z),z)}{1+\mathcal{T}_s'(s(z),z)}s(z)\zeta_{s|z}^c(z)\right\}\,dH_z(z).
\end{aligned}
\tag{95}
$$

[48] Applying Lemma 1, changes in z and s at earnings z^0 and above earnings z^0 are respectively

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

Differentiating both sides with respect to z^0 yields

$$s'(z^0)\int_{z\geq z^0}(1-\hat{g}(z))\,dH_z(z) - s^0(1-\hat{g}(z^0))h_z(z^0) \tag{96}$$
$$= -(1-\hat{g}(z^0))s^0 h_z(z^0) + s'_{inc}(z^0)\frac{\mathcal{T}'_z(s^0,z^0)+s'_{inc}(z^0)\mathcal{T}'_s(s^0,z^0)}{1-\mathcal{T}'_z(s^0,z^0)}\zeta^c_z(z^0)z^0h_z(z^0) + \frac{\mathcal{T}'_s(s^0,z^0)}{1+\mathcal{T}'_s(s^0,z^0)}s^0\zeta^c_{s|z}(z^0)h_z(z^0),$$

where both $s^0(1-\hat{g}(z^0))h_z(z^0)$ terms cancel out. Using equation (88) again, the second term on the right-hand side is equal to $s'_{inc}(z^0)\int_{z\geq z^0}(1-\hat{g}(z))\,dH_z(z)$. Rearranging terms, we finally obtain

$$\frac{\mathcal{T}'_s\left(s^0,z^0\right)}{1+\mathcal{T}'_s\left(s^0,z^0\right)}s(z^0)\zeta^c_{s|z}(z^0)h_z(z^0) = \left(s'(z^0)-s'_{inc}(z^0)\right)\int_{z\geq z^0}(1-\hat{g}(z))\,dH_z(z), \tag{97}$$

which is the optimality condition, equation (19), presented in Theorem 1.

C.B.3 Pareto Efficiency Condition

We can combine formulas for optimal marginal tax rates on z and on s to obtain a characterization of Pareto efficiency. Indeed, leveraging the above formula for optimal marginal tax rates on s, and replacing the integral term on the right-hand side by its value from the formula for optimal marginal tax rates on z, equation (88), and simplifying by $h_z(z^0)$ on both sides yields

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

which is the Pareto efficiency condition, equation (20), presented in Theorem 1.

C.C Proof of Theorem 2 (Smooth Tax Systems with Multidimensional Heterogeneity)

With multidimensional heterogeneity, individuals of different types θ may choose the same allocation (s,z) with $c(s,z) = z - s - \mathcal{T}(s,z)$. The problem of the government is thus to maximize the Lagrangian

$$\mathcal{L} = \int_\theta\left\{\alpha(\theta)\,U\Big(c(s(\theta),z(\theta)),s(\theta),z(\theta);\theta\Big) + \lambda\Big(\mathcal{T}(s(\theta),z(\theta)) - E\Big)\right\}dF(\theta), \tag{98}$$

where λ is the social marginal value of public funds.

To transform this integral over types θ into an integral over observables (s,z), let $\Theta(s,z)$ be the set of types choosing allocation (s,z). We then have

$$\mathcal{L} = \int_z\int_s\left\{\mathbb{E}\Big[\alpha(\theta)U\Big(c(s(\theta),z(\theta)),s(\theta),z(\theta);\theta\Big) + \lambda\Big(\mathcal{T}(s(\theta),z(\theta))-E\Big)\Big|\theta\in\Theta(s,z)\Big]\right\}h_{s|z}(s|z)ds\,dH_z(z). \tag{99}$$

C.C.1 Optimality Condition for Marginal Tax Rates on z

Reform. We consider a generalized SN-type reform. That is, we consider a reform that consists in a small increase $\Delta\tau_z$ of the marginal tax rate on earnings in a small bandwidth Δz at earnings level z^0, for individuals with $s\geq s^0$. Formally,

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

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

- *mechanical effects*:

$$\int_{z\geq z^0}\int_{s\geq s^0}\Big(1-\mathbb{E}\Big[\frac{\alpha(\theta)}{\lambda}\,U_c'\Big(c(s,z),s,z;\theta\Big)\Big|\theta\in\Theta(s,z)\Big]\Big)\,\Delta\tau_z\Delta z\,h_{s|z}(s|z)ds\,dH_z(z) \tag{100}$$

- *behavioral effects from changes in z:*[49]

$$\begin{aligned}-\int_{s\geq s^0}\mathcal{T}_z'(s,z^0)\mathbb{E}\left[\frac{z^0\zeta_z^c(\theta)}{1-\mathcal{T}_z'(s,z^0)}\Big|\theta\in\Theta(s,z^0)\right]\Delta\tau_z\,h_{s|z}(s|z^0)ds\,\Delta z h_z(z^0)\\ -\int_{z\geq z^0}\int_{s\geq s^0}\mathcal{T}_z'(s,z)\mathbb{E}\left[\frac{\eta_z(\theta)}{1-\mathcal{T}_z'(s,z)}\Big|\theta\in\Theta(s,z)\right]\Delta\tau_z\Delta z\,h_{s|z}(s|z)ds\,dH_z(z),\end{aligned} \tag{101}$$

  where the first line captures earnings substitution effects in the phase-in region and the second line captures earnings income effects above the phase-in region.

- *behavioral effects from changes in s:*

$$\begin{aligned}&-\int_{s\geq s^0}\mathcal{T}_s'(s,z^0)\mathbb{E}\left[s_{inc}'(\theta)\frac{z^0\zeta_z^c(\theta)}{1-\mathcal{T}_z'(s,z^0)}\Big|\theta\in\Theta(s,z^0)\right]\Delta\tau_z\,h_{s|z}(s|z^0)ds\,\Delta z h_z(z^0)\\ &-\int_{z\geq z^0}\int_{s\geq s^0}\mathcal{T}_s'(s,z)\mathbb{E}\left[\frac{\eta_{s|z}(\theta)}{1+\mathcal{T}_s'(s,z)}+s_{inc}'(\theta)\frac{\eta_z(\theta)}{1-\mathcal{T}_z'(s,z)}\Big|\theta\in\Theta(s,z)\right]\Delta\tau_z\Delta z\,h_{s|z}(s|z)ds\,dH_z(z),\end{aligned} \tag{102}$$

  where the first line captures adjustments in s in the phase-in region, driven by earnings substitution effects, and the second line captures adjustments in s above the phase-in region, driven by income effects.

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

$$\begin{aligned}\frac{1}{\lambda}\frac{d\mathcal{L}}{\Delta z}=&\int_{z\geq z^0}\int_{s\geq s^0}\mathbb{E}\left[1-\hat{g}(\theta)\Big|\theta\in\Theta(s,z)\right]\,h_{s|z}(s|z)ds\,\Delta\tau_z dH_z(z)\\ &-\int_{s\geq s^0}\mathbb{E}\left[FE_z(\theta)\Big|\theta\in\Theta(s,z^0)\right]\Delta\tau_z\,h_{s|z}(s|z^0)ds\,h_z(z^0)\end{aligned} \tag{103}$$

where $\hat{g}(\theta)$ is the social marginal welfare weight augmented with income effects, defined in (17), and where $FE_z(\theta)$ is the fiscal externality from earnings substitution effects, defined in (23).

Optimality. A direct implication is a sufficient statistics characterization of the optimal schedule of marginal tax rates on 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[50]

$$\mathbb{E}\left[FE_z|s\geq s^0,z^0\right](1-H_{s|z}(s^0|z^0))h_z(z^0)=\int_{z\geq z^0}\mathbb{E}\left[1-\hat{g}|s\geq s^0,z\right](1-H_{s|z}(s^0|z))dH_z(z). \tag{104}$$

This is (24) in Theorem 2.

[49]Note that by definition elasticity concepts include all circularities and adjustments induced by tax reforms such that individual changes in z and s are given by

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

[50]Formally, the adjustment factor with conditional densities is $\int_{s\geq s^0}h_{s|z}(s|z)ds$ and not $1-H_{s|z}(s^0|z)=\int_{s>s^0}h_{s|z}(s|z)ds$. However, because we assume away atoms in the distributions of (c,s,z), whether the inequality in the integral is strict or loose becomes irrelevant and we can use $1-H_{s|z}(s^0|z)$ without loss of generality.

C.C.2 Optimality Condition for Marginal Tax Rates on s

Reform. We here consider a generalized LED-type reform. We consider a small reform that consists of increasing the tax rate on s by $\Delta\tau_s \Delta z$ for all individuals with $s \geq s^0$ and $z \geq z^0$, phased-in over the earnings bandwidth $\left[z^0, z^0 + \Delta z\right]$. Formally, the reform is

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

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

- *mechanical effects:*

$$\int_{z \geq z^0} \int_{s \geq s^0} \left(1 - \mathbb{E}\Big[\frac{\alpha(\theta)}{\lambda}\, U_c'\Big(c(s,z), s, z; \theta\Big)\Big|\theta \in \Theta(s,z)\Big]\right)\, \Delta\tau_s \Delta z\,(s - s^0)\, h_{s|z}(s|z) ds\, dH_z(z) \quad (105)$$

- *behavioral effects from changes in z:*[51]

$$\begin{aligned} &- \int_{s \geq s^0} \mathcal{T}_z'(s,z^0) \left[\mathbb{E}\left[\frac{z^0 \zeta_z^c(\theta)}{1 - \mathcal{T}_z'(s^0,z^0)}\Big|\theta \in \Theta(s,z^0)\right]\, \Delta\tau_s (s - s^0)\right] h_{s|z}(s|z^0) ds\, h_z(z^0)\Delta z \quad (106) \\ &- \int_{z \geq z^0} \int_{s \geq s^0} \mathcal{T}_z'(s,z) \mathbb{E}\left[\frac{z\zeta_z^c(\theta) s_{inc}'(\theta)}{1 - \mathcal{T}_z'(s,z)} + \frac{\eta_z(\theta)(s - s^0)}{1 - \mathcal{T}_z'(s,z)}\Big|\theta \in \Theta(s,z^0)\right] \Delta\tau_s \Delta z\, h_{s|z}(s|z) ds\, dH_z(z), \end{aligned}$$

where the first line captures earnings substitution effects in the phase-in region, driven by the increase in marginal tax rates on z induced by the phase-in, and the second line captures earnings changes above the phase-in region, driven by the increase in marginal tax rates on s (Lemma 1) and by the increase in tax liability.

- *behavioral effects from changes in s:*

$$\begin{aligned} &- \int_{s \geq s^0} \mathcal{T}_s'(s,z^0) \mathbb{E}\left[s_{inc}'(\theta) \frac{z^0 \zeta_z^c(\theta)}{1 - \mathcal{T}_z'(s,z^0)}\Big|\theta \in \Theta(s,z^0)\right]\, \Delta\tau_s\,(s - s^0) h_{s|z}(s|z^0) ds\, h_z(z^0)\Delta z \quad (107) \\ &- \int_{z \geq z^0} \int_{s \geq s^0} \mathcal{T}_s'(s,z) \mathbb{E}\left[\frac{s\zeta_{s|z}^c(\theta) + \eta_{s|z}(\theta)(s - s^0)}{1 + \mathcal{T}_s'(s,z)}\Big|\theta \in \Theta(s,z^0)\right] \Delta\tau_s \Delta z\, h_{s|z}(s|z) ds\, dH_z(z) \\ &- \int_{z \geq z^0} \int_{s \geq s^0} \mathcal{T}_s'(s,z) \mathbb{E}\left[s_{inc}'(\theta)\left(\frac{z\zeta_z^c(\theta) s_{inc}'(\theta)}{1 - \mathcal{T}_z'(s,z)} + \frac{\eta_z(\theta)(s - s^0)}{1 - \mathcal{T}_z'(s,z)}\right)\Big|\theta \in \Theta(s,z^0)\right] \Delta\tau_s \Delta z\, h_{s|z}(s|z) ds\, dH_z(z), \end{aligned}$$

where the first line captures adjustments in s in the phase-in region, driven by earnings substitution effects, and the second and third line capture adjustments in s above the phase-in region, respectively driven by the increase in marginal tax rates on s and in tax liability and by earnings changes.

[51] Applying Lemma 1, individual changes in z and s for types with $z \in [z^0, z^0 + \Delta z]$ and $s \geq s^0$, and for types with $z \geq z^0 + \Delta z$ and $s \geq s^0$ are respectively

$$\begin{cases} dz = -\frac{z^0 \zeta_z^c(\theta)}{1 - \mathcal{T}_z'}\, \Delta\tau_s\,(s - s^0) \\ ds = s_{inc}'(\theta) dz \end{cases} \quad \text{and} \quad \begin{cases} dz = -\frac{z\zeta_z^c(\theta)}{1 - \mathcal{T}_z'}\, \Delta\tau_s \Delta z\, s_{inc}'(\theta) - \frac{\eta_z(\theta)}{1 - \mathcal{T}_z'}\, \Delta\tau_s \Delta z\,(s - s^0) \\ ds = -\frac{s\zeta_{s|z}^c(\theta)}{1 + \mathcal{T}_s'}\, \Delta\tau_s \Delta z - \frac{\eta_s(\theta)}{1 + \mathcal{T}_s'} \Delta\tau_s \Delta z\,(s - s^0) + s_{inc}'(\theta) dz \end{cases}$$

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

$$\begin{aligned}
&\frac{1}{\lambda}\frac{d\mathcal{L}}{\Delta\tau_s\Delta z} \\
&= -\int_{s\geq s^0} \mathbb{E}\left[FE_z(\theta)\middle|\theta\in\Theta(s,z^0)\right]\,(s-s^0)\,h_{s|z}(s|z^0)ds\,h_z(z^0) \qquad (108)\\
&+\int_{z\geq z^0}\int_{s\geq s^0}\left(1-\mathbb{E}\left[\hat{g}(\theta)\middle|\theta\in\Theta(s,z)\right]\right)(s-s^0)\,h_{s|z}(s|z)ds\,dH_z(z)\\
&-\int_{z\geq z^0}\int_{s\geq s^0}\left\{\mathbb{E}\left[FE_z(\theta)s'_{inc}(\theta)+\frac{\mathcal{T}'_s(s,z)}{1+\mathcal{T}'_s(s,z)}s\zeta^c_{s|z}(\theta)\middle|\theta\in\Theta(s,z)\right]\right\}h_{s|z}(s|z)ds\,dH_z(z),
\end{aligned}$$

where $\hat{g}(\theta)$ represents the average of social marginal welfare weight augmented with income effects, defined in (17), and where $FE_z(\theta)$ is the fiscal externality from earnings substitution effects, defined in (23).

Optimality. A direct implication of this result is a sufficient statistics characterization of the optimal marginal tax rates on s. Indeed, at the optimum, the reform should have a zero impact on the government objective, $d\mathcal{L}=0$. As a result, at each earnings z^0 and savings s^0, the optimal marginal tax rate on s satisfies the following condition,

$$\begin{aligned}
&\mathbb{E}\left[FE_z\cdot(s-s^0)\middle|s\geq s^0,z^0\right](1-H_{s|z}(s^0|z^0))h_z(z^0) \qquad (109)\\
&+\int_{z\geq z^0}\mathbb{E}\left[FE_z\cdot s'_{inc}+\frac{\mathcal{T}'_s(s,z)}{1+\mathcal{T}'_s(s,z)}s\zeta^c_{s|z}\middle|s\geq s^0,z\right](1-H_{s|z}(s^0|z))dH_z(z)\\
&=\int_{z\geq z^0}\mathbb{E}\left[(1-\hat{g})(s-s^0)\middle|s\geq s^0,z\right](1-H_{s|z}(s^0|z))dH_z(z).
\end{aligned}$$

We use the fact that

$$\begin{aligned}
\mathbb{E}\left[FE_z\cdot(s-s^0)\middle|s\geq s^0,z^0\right]&=\mathbb{E}\left[FE_z\middle|s\geq s^0,z^0\right]\left(\mathbb{E}\left[s\middle|s\geq s^0,z^0\right]-s^0\right)+Cov\left[FE_z,s\middle|s\geq s^0,z^0\right]\\
\mathbb{E}\left[(1-\hat{g})(s-s^0)\middle|s\geq s^0,z^0\right]&=\mathbb{E}\left[1-\hat{g}\middle|s\geq s^0,z^0\right]\left(\mathbb{E}\left[s\middle|s\geq s^0,z^0\right]-s^0\right)-Cov\left[\hat{g},s\middle|s\geq s^0,z^0\right]\\
\mathbb{E}\left[FE_z\cdot s'_{inc}\middle|s\geq s^0,z\right]&=\mathbb{E}\left[FE_z\middle|s\geq s^0,z\right]\mathbb{E}\left[s'_{inc}\middle|s\geq s^0,z\right]+Cov\left[FE_z,s'_{inc}\middle|s\geq s^0,z\right]
\end{aligned}$$

to obtain

$$\begin{aligned}
&\left\{\left(\mathbb{E}\left[s\middle|s\geq s^0,z^0\right]-s^0\right)\mathbb{E}\left[FE_z\middle|s\geq s^0,z^0\right]+Cov\left[FE_z,s\middle|s\geq s^0,z^0\right]\right\}(1-H_{s|z}(s^0|z^0))h_z(z^0)\\
&+\int_{z\geq z^0}\left\{\mathbb{E}\left[s'_{inc}\middle|s\geq s^0,z\right]\mathbb{E}\left[FE_z\middle|s\geq s^0,z\right]+Cov\left[FE_z,s'_{inc}\middle|s\geq s^0,z\right]\right\}(1-H_{s|z}(s^0|z))dH_z(z)\\
&+\int_{z\geq z^0}\mathbb{E}\left[\frac{\mathcal{T}'_s(s,z)}{1+\mathcal{T}'_s(s,z)}s\zeta^c_{s|z}\middle|s\geq s^0,z\right](1-H_{s|z}(s^0|z))dH_z(z) \qquad (110)\\
&=\int_{z\geq z^0}\left\{\left(\mathbb{E}\left[s\middle|s\geq s^0,z^0\right]-s^0\right)\mathbb{E}\left[1-\hat{g}\middle|s\geq s^0,z^0\right]-Cov\left[\hat{g},s\middle|s\geq s^0,z^0\right]\right\}(1-H_{s|z}(s^0|z))dH_z(z).
\end{aligned}$$

Now, we can use integration by parts to rewrite the first right-hand side term of (110) as

$$\begin{aligned}
&\int_{z\geq z^0}\left(\mathbb{E}\left[s\middle|s\geq s^0,z\right]-s^0\right)\mathbb{E}\left[1-\hat{g}\middle|s\geq s^0,z\right](1-H_{s|z}(s^0|z))h_z(z)dz \qquad (111)\\
&=-\int_{z\geq z^0}\left(\mathbb{E}\left[s\middle|s\geq s^0,z\right]-s^0\right)\frac{d}{dz}\left[\int_{y\geq z}\mathbb{E}\left[1-\hat{g}|s\geq s^0,y\right](1-H_{s|z}(s^0|y))h_z(y)dy\right]\\
&=\left(\mathbb{E}\left[s\middle|s\geq s^0,z^0\right]-s^0\right)\int_{z\geq z^0}\mathbb{E}\left[1-\hat{g}|s\geq s^0,z\right](1-H_{s|z}(s^0|z))h_z(z)dz\\
&\quad+\int_{z\geq z^0}\frac{d}{dz}\left\{\mathbb{E}\left[s\middle|s\geq s^0,z\right]\right\}\left[\int_{y\geq z}\mathbb{E}\left[1-\hat{g}|s\geq s^0,y\right](1-H_{s|z}(s^0|y))h_z(y)dy\right]dz.
\end{aligned}$$

Using the optimality of marginal tax rates on z,

$$\mathbb{E}\left[FE_z|s\geq s^0,z\right](1-H_{s|z}(s^0|z))h_z(z)=\int_{y\geq z}\mathbb{E}\left[1-\hat{g}|s\geq s^0,y\right](1-H_{s|z}(s^0|y))h_z(y)dy,$$

and simplifying by $\left(\mathbb{E}\left[s\middle|s\geq s^0,z\right]-s^0\right)\mathbb{E}\left[FE_z|s\geq s^0,z^0\right](1-H_{s|z}(s^0|z^0))h_z(z^0)$ on both sides of (110), we obtain

$$\begin{aligned}
&\int_{z\geq z^0}\mathbb{E}\left[\frac{\mathcal{T}_s'(s,z)}{1+\mathcal{T}_s'(s,z)}s\zeta_{s|z}^c\middle|s\geq s^0,z\right](1-H_{s|z}(s^0|z))dH_z(z) \qquad (112)\\
&=\int_{z\geq z^0}\left\{\left[\frac{d}{dz}\left(\mathbb{E}\left[s\middle|s\geq s^0,z\right]\right)-\mathbb{E}\left[s_{inc}'\middle|s\geq s^0,z\right]\right]\left[\int_{y\geq z}\mathbb{E}\left[1-\hat{g}\middle|s\geq s^0,y\right](1-H_{s|z}(s^0|y))h_z(y)dy\right]\right\}dz\\
&-\int_{z\geq z^0}Cov\left[\hat{g},s\middle|s\geq s^0,z\right](1-H_{s|z}(s^0|z))dH_z(z)-\int_{z\geq z^0}Cov\left[FE_z,s_{inc}'\middle|s\geq s^0,z\right](1-H_{s|z}(s^0|z))dH_z(z)\\
&-Cov\left[FE_z,s\middle|s\geq s^0,z^0\right](1-H_{s|z}(s^0|z^0))h_z(z^0).
\end{aligned}$$

This is the optimality condition in integral form. If we differentiate it with respect to z^0, we get

$$\begin{aligned}
&\mathbb{E}\left[\frac{\mathcal{T}_s'(s,z)}{1+\mathcal{T}_s'(s,z)}s\zeta_{s|z}^c\middle|s\geq s^0,z^0\right](1-H_{s|z}(s^0|z^0))h_z(z^0) \qquad (113)\\
&=\left[\frac{d}{dz^0}\left(\mathbb{E}\left[s\middle|s\geq s^0,z^0\right]\right)-\mathbb{E}\left[s_{inc}'\middle|s\geq s^0,z^0\right]\right]\int_{z\geq z^0}\mathbb{E}\left[1-\hat{g}\middle|s\geq s^0,z\right](1-H_{s|z}(s^0|z))dH_z(z)\\
&-Cov\left[\hat{g},s\middle|s\geq s^0,z^0\right](1-H_{s|z}(s^0|z^0))h_z(z^0)-Cov\left[FE_z,s_{inc}'\middle|s\geq s^0,z\right](1-H_{s|z}(s^0|z^0))h_z(z^0)\\
&+\frac{d}{dz^0}\left(Cov\left[FE_z,s\middle|s\geq s^0,z^0\right](1-H_{s|z}(s^0|z^0))h_z(z^0)\right).
\end{aligned}$$

This is (25) in Theorem 2.

C.D Proof of Proposition 2 (Measurement of Causal Income Effects)

Here, we derive different expressions for the empirical measurement of the sufficient statistic $s_{inc}'(\theta)$.

Case 1. If individuals' preferences are weakly separable between the utility of consumption $u(.)$ and the disutility to work $k(.)$, type θ's problem is written as

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

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

$$-(p+\mathcal{T}_s'(s,z))u_c'(z-ps-\mathcal{T}(s,z),s;\theta)+u_s'(z-ps-\mathcal{T}(s,z),s;\theta)=0.$$

Differentiating in a first step this equation with respect to savings s and earnings z yields

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

Differentiating in a second step this equation with respect to savings s and disposable income y yields

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

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

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

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 $y = z - ps - \mathcal{T}\left(s,z\right)$.

Case 2. If individuals' 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

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

where $\xi^s_w\left(\theta\right) \equiv \frac{w(\theta)}{s(\theta)}\frac{\partial s(\theta)}{\partial w(\theta)}$ is individuals' elasticity of savings with respect to their wage rate, and $\xi^h_w\left(\theta\right) \equiv \frac{w(\theta)}{h(\theta)}\frac{\partial h(\theta)}{\partial w(\theta)}$ 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 $\varsigma^c_s := -\frac{1-\mathcal{T}'_z}{s}\frac{\partial s}{\partial \mathcal{T}'_z}$ and $\zeta^c_z := -\frac{1-\mathcal{T}'_z}{z}\frac{\partial z}{\partial \mathcal{T}'_z}$, we then have

$$\begin{aligned}
\frac{\partial s}{\partial \mathcal{T}'_z} &= \frac{\partial s\left(z;\theta\right)}{\partial z}\frac{\partial z}{\mathcal{T}'_z} \\
\iff \left(-\frac{s}{1-\mathcal{T}'_z}\varsigma^c_s\left(\theta\right)\right) &= s'_{inc}\left(\theta\right)\left(-\frac{z}{1-\mathcal{T}'_z}\zeta^c_z\left(\theta\right)\right) \\
\iff s'_{inc}(\theta) &= \frac{s(\theta)}{z(\theta)}\frac{\varsigma^c_s(\theta)}{\zeta^c_z(\theta)}.
\end{aligned}$$

C.E Proof of Proposition A1 (Structural characterization of s'_{inc} and s'_{het})

In economies with preference heterogeneity, budget heterogeneity, and auxiliary choices, $s(z;\theta)$ solves

$$\max_s U\bigg(B\big(s,z,\chi(s,z;\theta);\theta\big) - \mathcal{T}(s,z), \phi_s\big(s,z,\chi(s,z;\theta);\theta\big), \phi_z\big(s,z,\chi(s,z;\theta);\theta\big), \chi(s,z;\theta);\theta\bigg) \tag{114}$$

where $\chi(s, z; \theta)$ denotes utility-maximizing auxiliary choices. As a result, applying the envelope theorem to changes in χ, $s(z;\theta)$ is defined by the following first-order condition

$$U_c'(.)\big[B_s'(s(z;\theta), z, \chi(s(z;\theta), z;\theta);\theta) - \mathcal{T}_s'(s(z;\theta), z)\big] + U_s'(.)\frac{\partial \phi_s(s, z, \chi(s, z;\theta);\theta)}{\partial s}\Big|_{s=s(z;\theta)} + U_z'(.)\frac{\partial \phi_z(s, z, \chi(s, z;\theta);\theta)}{\partial s}\Big|_{s=s(z;\theta)} = 0. \quad (115)$$

Now, to compute $s_{inc}' = \frac{\partial s(z;\theta)}{\partial z}$, we differentiate this first-order condition with respect to z while holding θ fixed:

$$\begin{aligned}
&\Big[\underbrace{U_{cc}''(.)(B_s' - \mathcal{T}_s') + U_{cs}''(.)\frac{\partial \phi_s}{\partial s} + U_{cz}''(.)\frac{\partial \phi_z}{\partial s}}_{\mathcal{K}_c}\Big]\Big[B_s'\frac{\partial s(z;\theta)}{\partial z} + B_z' + B_\chi'\Big(\frac{\partial \chi}{\partial s}\frac{\partial s(z;\theta)}{\partial z} + \frac{\partial \chi}{\partial z}\Big) - \mathcal{T}_s'\frac{\partial s(z;\theta)}{\partial z} - \mathcal{T}_z'\Big] \\
&+ \Big[\underbrace{U_{cs}''(.)(B_s' - \mathcal{T}_s') + U_{ss}''(.)\frac{\partial \phi_s}{\partial s} + U_{sz}''(.)\frac{\partial \phi_z}{\partial s}}_{\mathcal{K}_s}\Big]\Big[\frac{\partial \phi_s}{\partial s}\frac{\partial s(z;\theta)}{\partial z} + \frac{\partial \phi_s}{\partial z} + \frac{\partial \phi_s}{\partial \chi}\Big(\frac{\partial \chi}{\partial s}\frac{\partial s(z;\theta)}{\partial z} + \frac{\partial \chi}{\partial z}\Big)\Big] \\
&+ \Big[\underbrace{U_{cz}''(.)(B_s' - \mathcal{T}_s') + U_{sz}''(.)\frac{\partial \phi_s}{\partial s} + U_{zz}''(.)\frac{\partial \phi_z}{\partial s}}_{\mathcal{K}_z}\Big]\Big[\frac{\partial \phi_z}{\partial s}\frac{\partial s(z;\theta)}{\partial z} + \frac{\partial \phi_z}{\partial z} + \frac{\partial \phi_z}{\partial \chi}\Big(\frac{\partial \chi}{\partial s}\frac{\partial s(z;\theta)}{\partial z} + \frac{\partial \chi}{\partial z}\Big)\Big] \\
&+ \Big[\underbrace{U_{c\chi}''(.)(B_s' - \mathcal{T}_s') + U_{s\chi}''(.)\frac{\partial \phi_s}{\partial s} + U_{z\chi}''(.)\frac{\partial \phi_z}{\partial s}}_{\mathcal{K}_\chi}\Big]\Big[\frac{\partial \chi}{\partial s}\frac{\partial s(z;\theta)}{\partial z} + \frac{\partial \chi}{\partial z}\Big] \\
&+ U_c'\Big[B_{ss}''\frac{\partial s(z;\theta)}{\partial z} + B_{sz}'' + B_{s\chi}''\Big(\frac{\partial \chi}{\partial s}\frac{\partial s(z;\theta)}{\partial z} + \frac{\partial \chi}{\partial z}\Big) - \mathcal{T}_{ss}''\frac{\partial s(z;\theta)}{\partial z} - \mathcal{T}_{sz}''\Big] \\
&+ U_s'\Big[\frac{\partial^2 \phi_s}{(\partial s)^2}\frac{\partial s(z;\theta)}{\partial z} + \frac{\partial^2 \phi_s}{\partial s \partial z} + \frac{\partial^2 \phi_s}{\partial s \partial \chi}\Big(\frac{\partial \chi}{\partial s}\frac{\partial s(z;\theta)}{\partial z} + \frac{\partial \chi}{\partial z}\Big)\Big] \\
&+ U_z'\Big[\frac{\partial^2 \phi_z}{(\partial s)^2}\frac{\partial s(z;\theta)}{\partial z} + \frac{\partial^2 \phi_z}{\partial s \partial z} + \frac{\partial^2 \phi_z}{\partial s \partial \chi}\Big(\frac{\partial \chi}{\partial s}\frac{\partial s(z;\theta)}{\partial z} + \frac{\partial \chi}{\partial z}\Big)\Big] = 0. \quad (116)
\end{aligned}$$

Rearranging terms yields

$$\frac{\partial s(z;\theta)}{\partial z} = -\frac{\mathcal{K}_c\left[B_z' + B_\chi'\frac{\partial \chi}{\partial z} - \mathcal{T}_z'\right] + \mathcal{K}_s\left[\frac{\partial \phi_s}{\partial z} + \frac{\partial \phi_s}{\partial \chi}\frac{\partial \chi}{\partial z}\right] + \mathcal{K}_z\left[\frac{\partial \phi_z}{\partial z} + \frac{\partial \phi_z}{\partial \chi}\frac{\partial \chi}{\partial z}\right] + \mathcal{K}_\chi\left[\frac{\partial \chi}{\partial z}\right] + ...}{\mathcal{K}_c\left[B_s' + B_\chi'\frac{\partial \chi}{\partial s} - \mathcal{T}_s'\right] + \mathcal{K}_s\left[\frac{\partial \phi_s}{\partial s} + \frac{\partial \phi_s}{\partial \chi}\frac{\partial \chi}{\partial s}\right] + \mathcal{K}_z\left[\frac{\partial \phi_z}{\partial s} + \frac{\partial \phi_z}{\partial \chi}\frac{\partial \chi}{\partial s}\right] + \mathcal{K}_\chi\left[\frac{\partial \chi}{\partial s}\right] + ...}$$
$$\frac{... + U_c'\left[B_{sz}'' + B_{s\chi}''\frac{\partial \chi}{\partial z} - \mathcal{T}_{sz}''\right] + U_s'\left[\frac{\partial^2 \phi_s}{\partial s \partial z} + \frac{\partial^2 \phi_s}{\partial s \partial \chi}\frac{\partial \chi}{\partial z}\right] + U_z'\left[\frac{\partial^2 \phi_z}{\partial s \partial z} + \frac{\partial^2 \phi_z}{\partial s \partial \chi}\frac{\partial \chi}{\partial z}\right]}{... + U_c'\left[B_{ss}'' + B_{s\chi}''\frac{\partial \chi}{\partial s} - \mathcal{T}_{ss}''\right] + U_s'\left[\frac{\partial^2 \phi_s}{(\partial s)^2} + \frac{\partial^2 \phi_s}{\partial s \partial \chi}\frac{\partial \chi}{\partial s}\right] + U_z'\left[\frac{\partial^2 \phi_z}{(\partial s)^2} + \frac{\partial^2 \phi_z}{\partial s \partial \chi}\frac{\partial \chi}{\partial s}\right]}. \quad (117)$$

Similarly, to compute $\frac{\partial s(z;\theta)}{\partial \theta}$, we differentiate the first-order condition for $s(z;\theta)$ with respect to θ while holding

z fixed:

$$\underbrace{\left[U''_{cc}(.)\left(B'_s-\mathcal{T}'_s\right)+U''_{cs}(.)\frac{\partial\phi_s}{\partial s}+U''_{cz}(.)\frac{\partial\phi_z}{\partial s}\right]}_{\mathcal{K}_c}\left[B'_s\frac{\partial s(z;\theta)}{\partial\theta}+B'_\chi\Big(\frac{\partial\chi}{\partial s}\frac{\partial s(z;\theta)}{\partial\theta}+\frac{\partial\chi}{\partial\theta}\Big)-\mathcal{T}'_s\frac{\partial s(z;\theta)}{\partial\theta}\right]$$
$$+\underbrace{\left[U''_{cs}(.)\left(B'_s-\mathcal{T}'_s\right)+U''_{ss}(.)\frac{\partial\phi_s}{\partial s}+U''_{sz}(.)\frac{\partial\phi_z}{\partial s}\right]}_{\mathcal{K}_s}\left[\frac{\partial\phi_s}{\partial s}\frac{\partial s(z;\theta)}{\partial\theta}+\frac{\partial\phi_s}{\partial\chi}\Big(\frac{\partial\chi}{\partial s}\frac{\partial s(z;\theta)}{\partial\theta}+\frac{\partial\chi}{\partial\theta}\Big)+\frac{\partial\phi_s}{\partial\theta}\right]$$
$$+\underbrace{\left[U''_{cz}(.)\left(B'_s-\mathcal{T}'_s\right)+U''_{sz}(.)\frac{\partial\phi_s}{\partial s}+U''_{zz}(.)\frac{\partial\phi_z}{\partial s}\right]}_{\mathcal{K}_z}\left[\frac{\partial\phi_z}{\partial s}\frac{\partial s(z;\theta)}{\partial\theta}+\frac{\partial\phi_z}{\partial\chi}\Big(\frac{\partial\chi}{\partial s}\frac{\partial s(z;\theta)}{\partial\theta}+\frac{\partial\chi}{\partial\theta}\Big)+\frac{\partial\phi_z}{\partial\theta}\right]$$
$$+\underbrace{\left[U''_{c\chi}(.)\left(B'_s-\mathcal{T}'_s\right)+U''_{s\chi}(.)\frac{\partial\phi_s}{\partial s}+U''_{z\chi}(.)\frac{\partial\phi_z}{\partial s}\right]}_{\mathcal{K}_\chi}\left[\frac{\partial\chi}{\partial s}\frac{\partial s(z;\theta)}{\partial\theta}+\frac{\partial\chi}{\partial\theta}\right]$$
$$+\underbrace{\left[U''_{c\theta}(.)\left(B'_s-\mathcal{T}'_s\right)+U''_{s\theta}(.)\frac{\partial\phi_s}{\partial s}+U''_{z\theta}(.)\frac{\partial\phi_z}{\partial s}\right]}_{\mathcal{K}_\theta}$$
$$+U'_c\Big[B''_{ss}\frac{\partial s(z;\theta)}{\partial\theta}+B''_{s\chi}\Big(\frac{\partial\chi}{\partial s}\frac{\partial s(z;\theta)}{\partial\theta}+\frac{\partial\chi}{\partial\theta}\Big)-\mathcal{T}''_{ss}\frac{\partial s(z;\theta)}{\partial\theta}\Big]$$
$$+U'_s\Big[\frac{\partial^2\phi_s}{(\partial s)^2}\frac{\partial s(z;\theta)}{\partial\theta}+\frac{\partial^2\phi_s}{\partial s\partial\chi}\Big(\frac{\partial\chi}{\partial s}\frac{\partial s(z;\theta)}{\partial\theta}+\frac{\partial\chi}{\partial\theta}\Big)+\frac{\partial^2\phi_s}{\partial s\partial\theta}\Big]$$
$$+U'_z\Big[\frac{\partial^2\phi_z}{(\partial s)^2}\frac{\partial s(z;\theta)}{\partial\theta}+\frac{\partial^2\phi_z}{\partial s\partial\chi}\Big(\frac{\partial\chi}{\partial s}\frac{\partial s(z;\theta)}{\partial\theta}+\frac{\partial\chi}{\partial\theta}\Big)+\frac{\partial^2\phi_z}{\partial s\partial\theta}\Big]=0. \tag{118}$$

Rearranging terms yields

$$\frac{\partial s(z;\theta)}{\partial\theta}=-\frac{\mathcal{K}_c\left[B'_\chi\frac{\partial\chi}{\partial\theta}\right]+\mathcal{K}_s\left[\frac{\partial\phi_s}{\partial\chi}\frac{\partial\chi}{\partial\theta}+\frac{\partial\phi_s}{\partial\theta}\right]+\mathcal{K}_z\left[\frac{\partial\phi_z}{\partial\chi}\frac{\partial\chi}{\partial\theta}+\frac{\partial\phi_z}{\partial\theta}\right]+\mathcal{K}_\chi\left[\frac{\partial\chi}{\partial\theta}\right]+\mathcal{K}_\theta+...}{\mathcal{K}_c\left[B'_s+B'_\chi\frac{\partial\chi}{\partial s}-\mathcal{T}'_s\right]+\mathcal{K}_s\left[\frac{\partial\phi_s}{\partial s}+\frac{\partial\phi_s}{\partial\chi}\frac{\partial\chi}{\partial s}\right]+\mathcal{K}_z\left[\frac{\partial\phi_z}{\partial s}+\frac{\partial\phi_z}{\partial\chi}\frac{\partial\chi}{\partial s}\right]+\mathcal{K}_\chi\left[\frac{\partial\chi}{\partial s}\right]+...}$$
$$\frac{...+U'_c\left[B''_{s\chi}\frac{\partial\chi}{\partial\theta}\right]+U'_s\left[\frac{\partial^2\phi_s}{\partial s\partial\chi}\frac{\partial\chi}{\partial\theta}+\frac{\partial^2\phi_s}{\partial s\partial\theta}\right]+U'_z\left[\frac{\partial^2\phi_z}{\partial s\partial\chi}\frac{\partial\chi}{\partial\theta}+\frac{\partial^2\phi_z}{\partial s\partial\theta}\right]}{...+U'_c\left[B''_{ss}+B''_{s\chi}\frac{\partial\chi}{\partial s}-\mathcal{T}''_{ss}\right]+U'_s\left[\frac{\partial^2\phi_s}{(\partial s)^2}+\frac{\partial^2\phi_s}{\partial s\partial\chi}\frac{\partial\chi}{\partial s}\right]+U'_z\left[\frac{\partial^2\phi_z}{(\partial s)^2}+\frac{\partial^2\phi_z}{\partial s\partial\chi}\frac{\partial\chi}{\partial s}\right]}. \tag{119}$$

C.F Proof of Proposition 3 & A2 (Simple Tax Systems with Unidimensional Heterogeneity)

All simple tax systems that we consider feature a nonlinear earnings tax $T_z(z)$. The derivation of optimal earnings tax formulas for simple tax systems with unidimensional heterogeneity thus parallels that for general smooth tax systems as it uses the same Saez (2001) reform increasing marginal tax rates on z around a given earnings level z^0 (see Appendix C.B.1). As a result, the formula for optimal marginal tax rates on z, equation (18), continues to hold for simple tax systems. The rest of this section details the proofs for optimal marginal tax rates on s and Pareto efficiency in the different simple tax systems that we consider.

C.F.1 SL tax system

SL tax reform. When the government uses a linear tax on s such that $\mathcal{T}(s,z)=\tau_s\,s+T_z(z)$, we consider a small reform of the linear tax rate τ_s that consists in a small increase $\Delta\tau_s$. For an individual with earnings z, this reform increases tax liability by $\Delta\tau_s\,s(z)$ and increases the marginal tax rate on s by $\Delta\tau_s$.

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

- *mechanical effects:*

$$\int_z \left(1 - \frac{\alpha(z)}{\lambda} U_c'\left(c(z), s(z), z; \vartheta(z)\right)\right) \Delta\tau_s\, s(z)\, dH_z(z) \tag{120}$$

- *behavioral effects from changes in z:*[52]

$$-\int_z T_z'(z)\left[\frac{z\zeta_z^c(z)}{1-T_z'(z)}\,\Delta\tau_s\, s_{inc}'(z) + \frac{\eta_z(z)}{1-T_z'(z)}\,\Delta\tau_s\, s(z)\right] dH_z(z) \tag{122}$$

- *behavioral effects from changes in s:*

$$-\int_z \tau_s\left[\frac{s(z)\zeta_{s|z}^c(z)}{1+\tau_s}\Delta\tau_s + \frac{\eta_{s|z}(z)}{1+\tau_s}\Delta\tau_s\, s\left(z\right)\right] dH_z(z)$$
$$-\int_z \tau_s s_{inc}'(z)\left[\frac{z\zeta_z^c(z)}{1-T_z'}\,\Delta\tau_s\, s_{inc}'(z) + \frac{\eta_z(z)}{1-T_z'}\,\Delta\tau_s\, s\left(z\right)\right] dH_z(z) \tag{123}$$

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

$$\frac{d\mathcal{L}}{\lambda} = \int_z \left\{ s(z)\left(1-\hat{g}(z)\right) - \frac{T_z'(z) + s_{inc}'(z)\tau_s}{1-T_z'(z)} z\zeta_z^c(z)\, s_{inc}'(z) - \frac{\tau_s}{1+\tau_s} s(z)\zeta_{s|z}^c(z) \right\} \Delta\tau_s\, dH_z(z), \tag{124}$$

with social marginal welfare weights augmented with the fiscal impact of income effects, $\hat{g}(z)$, defined in (17).

Optimal SL tax rate on s. A direct implication of this result is a sufficient statistics characterization of the optimal linear tax rate τ_s. Indeed, at the optimum, the reform should have a zero impact on the government objective, meaning that the optimal τ_s satisfies

$$\frac{\tau_s}{1+\tau_s}\int_z s(z)\zeta_{s|z}^c(z) dH_z(z) = \int_z \left\{ s(z)\left(1-\hat{g}(z)\right) - \frac{T_z'(z) + \tau_s\, s_{inc}'(z)}{1-T_z'(z)} z\zeta_z^c(z)\, s_{inc}'(z) \right\} dH_z(z). \tag{125}$$

Note that here $\int_z \frac{s(z)}{1+\tau_s}\zeta_{s|z}^c(z) dH_z(z)$ is the aggregate population response to a change in τ_s. Defining $\overline{\zeta_{s|z}^c}$ as the aggregate elasticity of $\bar{s} := \int_z s(z) dH_z(z)$, we can rewrite this term as $\frac{\bar{s}}{1+\tau_s}\overline{\zeta_{s|z}^c}$. This yields equation (59) in Proposition A2, and it holds for any (potentially suboptimal) nonlinear earnings tax schedule $T_z\left(z\right)$.

Now, assume that the earnings tax schedule is optimal. Equation (19) applied to SL tax systems then implies that at each earnings z,

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

Using this expression, we obtain

$$\frac{\tau_s}{1+\tau_s}\bar{s}\overline{\zeta_{s|z}^c} = \int_z \left\{ s(z)\left(1-\hat{g}(z)\right) \right\} dH_z(z) - \int_z \left\{ s_{inc}'(z) \int_{x\geq z} \left(1-\hat{g}(x)\right) dH_z(x) \right\} dz. \tag{126}$$

By integration by part, we have

$$\int_z \left\{ s(z)\left(1-\hat{g}(z)\right) \right\} dH_z(z) = \int_z \left\{ s'(z) \int_{x\geq z} \left(1-\hat{g}(x)\right) dH_z(x) \right\} dz$$

[52] Applying Lemma 1, changes in z and s are here given by

$$\begin{cases} dz = -\frac{z\zeta_z^c(z)}{1-T_z'(z)}\,\Delta\tau_s\, s_{inc}'(z) - \frac{\eta_z(z)}{1-T_z'(z)}\,\Delta\tau_s\, s(z) \\ ds = -\frac{s(z)\zeta_{s|z}^c(z)}{1+\tau_s}\,\Delta\tau_s - \frac{\eta_{s|z}(z)}{1+\tau_s}\Delta\tau_s\, s(z) + s_{inc}'(z) dz \end{cases} \tag{121}$$

which yields equation (27) in Proposition 3:

$$\frac{\tau_s}{1+\tau_s}\bar{s}\overline{\zeta^c_{s|z}} = \int_z (s'(z) - s'_{inc}(z)) \left[\int_{x\geq z} (1-\hat{g}(x))\, h_z(x) dx\right] dz. \tag{127}$$

To derive an alternative expression of this result, we use again integration by part,

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

to obtain

$$\begin{aligned}
\frac{\tau_s}{1+\tau_s}\bar{s}\overline{\zeta^c_{s|z}} &= \int_z \left[\int_{x\leq z} (s'(x) - s'_{inc}(x))\, dx\right] (1-\hat{g}(z)) h_z(z) dz \\
&= \int_z \left[\int_{x\leq z} (s'(x) - s'_{inc}(x))\, dx\right] h_z(z) dz - \int_z \left[\int_{x\leq z} (s'(x) - s'_{inc}(x))\, dx\right] \hat{g}(z) h_z(z) dz \\
&= -Cov\left[\int_{x\leq z} (s'(x) - s'_{inc}(x))\, dx \,,\, \hat{g}(z)\right] \qquad (128)
\end{aligned}$$

where the last line follows from the definition of the covariance, $Cov(X,Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$, and the fact that $\int_z \hat{g}(z) h_z(z) dz = 1$. This is equation (28) in Proposition 3.

Pareto efficiency for SL tax systems. There are at least two methods to derive a condition for Pareto efficiency. A first "constructive" proof is to combine tax reforms in a way that neutralizes all lump-sum changes in tax liability, thereby offsetting all utility changes. A second "analytical" proof is to rely on the previous expressions obtained integration by parts. The working paper version of this paper features a "constructive" proof and we here provide a shorter "analytical" proof.

Starting from equation (27) in Proposition 3 characterizing the optimal τ_s through

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

and noting as before that optimal marginal tax rates on z in a SL system satisfy

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

we immediately obtain equation (31) in Proposition 3:

$$\frac{\tau_s}{1+\tau_s}\bar{s}\overline{\zeta^c_{s|z}} = \int_z (s'(z) - s'_{inc}(z))\, \frac{T'_z(z) + s'_{inc}(z)\tau_s}{1-T'_z(z)} z\zeta^c_z(z) dH_z(z). \tag{129}$$

C.F.2 SN tax systems

SN tax reform. When the government uses a SN tax system such that $\mathcal{T}(s,z) = T_s(s) + T_z(z)$, we consider a small reform of the tax on s at $s^0 = s(\theta^0)$ that consists in a small increase $\Delta\tau_s$ of the marginal tax rate on s in a small bandwidth Δs. Formally,

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

Since we assume there is a strictly increasing mapping between z and s, we denote z^0 the earnings level such that $s^0 = s(z^0)$.[53] We characterize the impact of this reform on the government objective function $\mathcal{L}$ as $\Delta s \to 0$. Normalizing all effects by $1/\lambda$, the reform induces

- *mechanical effects:*

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

- *behavioral effects from changes in z:*[54]

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

- *behavioral effects from changes in s:*

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

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

$$\frac{1}{\lambda}\frac{d\mathcal{L}}{\Delta s} = s'\left(z^0\right) \int_{z\geq z^0} (1-\hat{g}(z))\, \Delta\tau_s \, dH_z(z) - \Bigg\{ \mathcal{T}_s'\left(s^0, z^0\right) \frac{s^0}{1+\mathcal{T}_s'(s^0,z^0)} \zeta_{s|z}^c(z^0)$$
$$+ \left[\mathcal{T}_z'(s^0,z^0) + s_{inc}'(z^0)\mathcal{T}_s'(s^0,z^0)\right] \frac{z^0}{1-\mathcal{T}_z'(s^0,z^0)} \zeta_z^c(z^0) s_{inc}'(z^0) \Bigg\} \Delta\tau_s \, h_z(z^0). \tag{131}$$

Optimal SN tax rates on s. A direct implication of this result is a sufficient statistics characterization of the optimal marginal tax rates on 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 $s^0 = s\left(z^0\right)$ the optimal marginal tax rate on s satisfies

$$\frac{\mathcal{T}_s'(s^0,z^0)}{1+\mathcal{T}_s'(s^0,z^0)} s^0 \zeta_{s|z}^c(z^0) h_z(z^0) = s'\left(z^0\right) \int_{z\geq z^0} (1-\hat{g}(z))\, dH_z(z)$$
$$- s_{inc}'(z^0) \frac{\mathcal{T}_z'(s^0,z^0) + s_{inc}'(z^0)\mathcal{T}_s'(s^0,z^0)}{1-\mathcal{T}_z'(s^0,z^0)} z^0 \zeta_z^c(z^0) h_z(z^0) \tag{132}$$

which is equation (60) in Proposition A2, recognizing that $\mathcal{T}_z'(s,z) = T_z'(z)$ and $\mathcal{T}_s'(s,z) = T_s'(s)$ and that we can equivalently write this condition in terms of the distribution of s. This characterization holds for any (potentially suboptimal) nonlinear earnings tax schedule $T_z(z)$.

Now, further assume that the earnings tax schedule is optimal. Equation (18) applied to SN tax systems then

[53] Our sufficient statistic characterization of optimal SN tax systems relies on strict monotonicity of the function $s(z)$. Hence, it is also valid if we assume a strictly decreasing mapping $s(z)$. It can be extended to weakly monotonic $s(z)$ (i.e., non-decreasing or non-increasing) with slight modifications.

[54] Applying Lemma 1, changes in z and s are here given by

$$\begin{cases} dz = -\frac{z}{1-T_z'} \zeta_z^c(z) \Delta T_z^{\theta\prime} - \frac{\eta_z(z)}{1-T_z'} \Delta T_z^\theta \\ ds = -\frac{s(z)}{1+T_s'} \zeta_{s|z}^c(z) \Delta T_s' - \frac{\eta_{s|z}(z)}{1+T_s'} \Delta T_s + s_{inc}'(z) dz \end{cases} \tag{130}$$

where T_z^θ is a $s_{inc}'(z)\Delta\tau_s$ increase in the marginal earnings tax rate when $s \in [s^0, s^0+\Delta s]$, and a $\Delta\tau_s\Delta s$ increase in tax liability when $s \geq s^0 + \Delta s$. Moreover, the mass of individuals in the bandwidth is $\Delta s\, h_s(s(z^0)) = \Delta s\, \frac{h_z(z^0)}{s'(z^0)}$.

implies that at each earnings z^0,

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

Using this expression to substitute the last term yields equation (29) in Proposition 3:

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

$$\iff \frac{T_s'(s^0)}{1 + T_s'(s^0)} s^0 \zeta_{s|z}^c(s^0) h_s(s^0) s'(z^0) = \left(s'(z^0) - s_{inc}'(z^0)\right) \int_{s \geq s^0} (1 - \hat{g}(s))\, dH_s(s) \tag{134}$$

where we have used in the last line that $h_s(s(z^0)) = \frac{h_z(z^0)}{s'(z^0)}$ as well as

$$\int_{z \geq z^0} (1 - \hat{g}(z))\, h_z(z) dz = \int_{s(z) \geq s(z^0)} (1 - \hat{g}(s(z)))\, h_s(s) s'(z) dz = \int_{s \geq s^0} (1 - \hat{g}(s))\, h_s(s) ds. \tag{135}$$

Pareto efficiency for SN tax systems. We can combine formulas for optimal marginal tax rates on s and z to obtain a characterization of Pareto efficiency (as we did for SL). Indeed, leveraging the previous optimal formula for marginal tax rates on s, and replacing the integral term by its value given from the optimal formula for marginal earnings tax rates yields

$$\frac{T_s'(s(z^0))}{1 + T_s'(s(z^0))} s(z^0) \zeta_{s|z}^c(z^0) = \left(s'(z^0) - s_{inc}'(z^0)\right) \frac{T_z'(z^0) + s_{inc}'(z^0)T_s'(s(z^0))}{1 - T_z'(z^0)} z^0 \zeta_z^c(z^0) \tag{136}$$

which is the Pareto efficiency condition (32) presented in Proposition 3.

C.F.3 LED tax systems

The particular linear reforms considered in the sufficient statistics characterization of optimal marginal tax rates on s for general smooth tax systems $\mathcal{T}(s, z)$ are also available for LED tax systems. As a result, the derivation of optimal marginal tax rates on s in LED tax systems is identical to the derivation for general smooth tax systems (see Appendix C.B.2), and the optimality formula in equation (19) continues to hold. This, in turn, implies that the Pareto efficiency condition in equation (20) also holds, thereby proving all sufficient statistics characterizations for LED tax systems.

C.G Proof of Proposition 4, A3 & A4 (Simple Tax Systems with Multidimensional Heterogeneity)

All simple tax systems that we consider feature a nonlinear earnings tax $T_z(z)$. We can thus consider a reform that increases marginal tax rates on z around a given earnings level z^0, independent of individuals' level of s. In our derivations of the optimal marginal tax rates on z for general smooth tax systems, we use a similar reform except that it only affects individuals with s higher than a given s_0 (see Appendix C.C.1). In other words, we here consider a particular case where $s_0 = 0$ such that $H_{s|z}(s^0|z^0) = 0$. This leads to the following characterization of optimal marginal tax rates on z:

$$\mathbb{E}\left[FE_z|z^0\right] h_z(z^0) = \int_{z \geq z^0} \mathbb{E}\left[1 - \hat{g}|z\right] dH_z(z). \tag{137}$$

Optimality conditions in Proposition A4 directly follow, replacing FE_z, $\mathcal{T}_s'$, and $\mathcal{T}_z'$ by their respective values in SL, SN, and LED tax systems.

The rest of this section details the derivations of conditions for optimal tax rates on s in the different simple tax systems that we consider.

C.G.1 Separable linear (SL) tax system

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

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

- *mechanical effects*

$$\int_z\int_s\left(1-\mathbb{E}\Big[\frac{\alpha(\theta)}{\lambda}\,U'_c\Big(c(s,z),s,z;\theta\Big)\Big|\theta\in\Theta(s,z)\Big]\right)\;\Delta\tau_s s\,h_{s|z}(s|z)ds\,dH_z(z) \tag{138}$$

- *behavioral effects from changes in z*[55]

$$-\int_z\int_s T'_z(z)\mathbb{E}\left[\frac{z\zeta^c_z(\theta)s'_{inc}(\theta)}{1-T'_z(z)}+\frac{\eta_z(\theta)\,s}{1-T'_z(z)}\Big|\theta\in\Theta(s,z)\right]\Delta\tau_s\,h_{s|z}(s|z)ds\,dH_z(z) \tag{140}$$

- *behavioral effects from changes in s*

$$\begin{aligned}&-\int_z\int_s\tau_s\,\mathbb{E}\Big[\frac{s\zeta^c_{s|z}(\theta)+\eta_{s|z}(\theta)s}{1+\tau_s}\Big|\theta\in\Theta(s,z)\Big]\Delta\tau_s h_{s|z}(s|z)ds\,dH_z(z)\\&\quad-\int_z\int_s\tau_s\,\mathbb{E}\Big[s'_{inc}(\theta)\Big(\frac{z\zeta^c_z(\theta)s'_{inc}(\theta)}{1-T'_z(z)}+\frac{\eta_z(\theta)s}{1-T'_z(z)}\Big)\Big|\theta\in\Theta(s,z)\Big]\Delta\tau_s h_{s|z}(s|z)ds\,dH_z(z)\end{aligned} \tag{141}$$

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

$$\begin{aligned}\frac{1}{\lambda}\frac{d\mathcal{L}}{\Delta\tau_s\Delta z}=&\int_z\int_s\Big(1-\mathbb{E}\Big[\hat{g}(\theta)\Big|\theta\in\Theta(s,z)\Big]\Big)\,s\,h_{s|z}(s|z)ds\,dH_z(z)\\&-\int_z\int_s\left\{\mathbb{E}\left[\frac{T'_z(z)+s'_{inc}(\theta)\tau_s}{1-T'_z(z)}z\zeta^c_z(\theta)s'_{inc}(\theta)+\frac{\tau_s}{1+\tau_s}s\zeta^c_{s|z}(\theta)\Big|\theta\in\Theta(s,z)\right]\right\}h_{s|z}(s|z)ds\,dH_z(z)\end{aligned} \tag{142}$$

where $\hat{g}(s,z):=\mathbb{E}\left[\hat{g}(\theta)|\theta\in\Theta(s,z)\right]$ represents the average of social marginal welfare weight augmented with income effects, defined in (17).

Optimal SL tax rate on s. Characterizing the optimal linear tax rate τ_s through $d\mathcal{L}=0$, it satisfies

$$\frac{\tau_s}{1+\tau_s}\int_z\mathbb{E}\left[s\zeta^c_{s|z}\Big|z\right]dH_z(z)=\int_z\mathbb{E}\Big[(1-\hat{g})\cdot s\Big|z\Big]dH_z(z)-\int_z\mathbb{E}\left[FE_z\cdot s'_{inc}\Big|z\right]dH_z(z). \tag{143}$$

where, noting $\frac{1}{1+\tau_s}\int_z\mathbb{E}\left[s\zeta^c_{s|z}\Big|z\right]dH_z(z)$ is the aggregate population response to a change in τ_s, we define $\overline{\zeta^c_{s|z}}$ as the aggregate elasticity of $\overline{s}:=\int_s s dH_s(s)$ to rewrite this term as $\frac{\overline{s}}{1+\tau_s}\overline{\zeta^c_{s|z}}$. This yeilds (62) in Proposition A3, valid for any (potentially suboptimal) nonlinear earnings tax schedule $T_z(z)$.where $FE_z(\theta)$ is the fiscal externality from earnings substitution effects defined in (23). Using

$$\begin{aligned}\mathbb{E}\left[(1-\hat{g})\cdot s\Big|z\right]&=\mathbb{E}\left[1-\hat{g}\Big|z\right]\mathbb{E}\left[s\Big|z\right]-Cov\left[\hat{g},s\Big|z\right],\\\mathbb{E}\left[FE_z\cdot s'_{inc}\Big|z\right]&=\mathbb{E}\left[FE_z\Big|z\right]\mathbb{E}\left[s'_{inc}\Big|z\right]+Cov\left[FE_z,s'_{inc}\Big|z\right],\end{aligned}$$

[55] Applying Lemma 1, changes in z and s are here given by

$$\begin{cases}dz=-\frac{z\zeta^c_z(\theta)}{1-T'_z(z)}\,\Delta\tau_s\,s'_{inc}(\theta)-\frac{\eta_z(\theta)}{1-T'_z(z)}\,\Delta\tau_s\,s\\ds=-\frac{s\zeta^c_{s|z}(\theta)}{1+\tau_s}\,\Delta\tau_s-\frac{\eta_{s|z}(\theta)}{1+\tau_s}\Delta\tau_s\,s+s'_{inc}(\theta)dz\end{cases} \tag{139}$$

this gives

$$\frac{\tau_s}{1+\tau_s}\bar{s}\overline{\zeta^c_{s|z}} = \int_z \left\{ \bar{s}(z)\mathbb{E}\left[1-\hat{g}\Big|z\right] - Cov\left[\hat{g}, s\Big|z\right] - \overline{s'_{inc}}(z)\mathbb{E}\left[FE_z\Big|z\right] - Cov\left[FE_z, s'_{inc}\Big|z\right]\right\} dH_z(z). \quad (144)$$

Assuming that the nonlinear earnings tax schedule $T_z(z)$ is optimal, we can use the condition for optimal marginal tax rates on z,

$$\mathbb{E}\left[FE_z|z\right] h_z(z) = \int_{y\geq z} \mathbb{E}\left[1-\hat{g}|y\right] dH_z(y). \quad (145)$$

and use the fact that by integration by parts,

$$\int_z \bar{s}(z)\mathbb{E}\left[1-\hat{g}|z\right] h_z(z)dz = \int_z \bar{s}'(z)\left[\int_{y\geq z} \mathbb{E}\left[1-\hat{g}|y\right] dH_z(y)\right] dz$$

to obtain equation (34) in Proposition 4:

$$\frac{\tau_s}{1+\tau_s} = \frac{1}{\bar{s}\overline{\zeta^c_{s|z}}}\left\{\int_z \left[\left(\bar{s}'(z) - \overline{s'_{inc}}(z)\right)\int_{y\geq z} \mathbb{E}\left[1-\hat{g}\Big|y\right] dH_z(y)\right] dz \right. \quad (146)$$
$$\left. - \int_z \left[Cov\left[\hat{g}, s\Big|z\right] + Cov\left[FE_z, s'_{inc}\Big|z\right]\right] dH_z(z)\right\}$$

C.G.2 Separable nonlinear (SN) tax system

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

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

- *mechanical effects*

$$\int_{s\geq s^0}\int_z \left(1 - \mathbb{E}\Big[\frac{\alpha(\theta)}{\lambda} U'_c\Big(c(s,z), s, z; \theta\Big)\Big|\theta \in \Theta(s,z)\Big]\right) \Delta\tau_s \Delta s\, h_{s|z}(s|z) ds\, dH_z(z) \quad (147)$$

- *behavioral effects from changes in z*[56]

$$-\int_z T'_z(z)\,\mathbb{E}\left[\frac{z\zeta^c_z(\theta)}{1-T'_z(z)} s'_{inc}(\theta)\Big|\theta\in\Theta(s^0,z)\right]\Delta\tau_s\Delta s\, h_{s|z}(s^0|z)\, dH_z(z)$$
$$-\int_{s\geq s^0}\int_z T'_z(z)\,\mathbb{E}\left[\frac{\eta_z(\theta)}{1-T'_z(z)}\Big|\theta\in\Theta(s^0,z)\right]\Delta\tau_s\Delta s\, h_{s|z}(s|z) ds\, dH_z(z) \quad (149)$$

[56]Applying Lemma 1, changes in z and s are here given by

$$\begin{cases} dz = -\frac{z}{1-T'_z}\zeta^c_z(\theta)\Delta T^{\theta\prime}_z - \frac{\eta_z(\theta)}{1-T'_z}\Delta T^\theta_z \\ ds = -\frac{s}{1+T'_s}\zeta^c_{s|z}(\theta)\Delta T'_s - \frac{\eta_{s|z}(\theta)}{1+T'_s}\Delta T_s + s'_{inc}(\theta)dz \end{cases} \quad (148)$$

where the reform ΔT^θ_z is a $s'_{inc}(\theta)\Delta\tau_s$ increase in the marginal earnings tax rate when $s\in[s^0, s^0+\Delta s]$, and a $\Delta\tau_s\Delta s$ increase in tax liability when $s\geq s^0+\Delta s$.

- *behavioral effects from changes in s*

$$- T_s'(s^0)\int_z \mathbb{E}\left[\frac{s^0\zeta^c_{s|z}(\theta)}{1+T_s'(s^0)} + s'_{inc}(\theta)\frac{z\zeta^c_z(\theta)}{1-T_z'(z)}s'_{inc}(\theta)\Big|\theta\in\Theta(s^0,z)\right]\Delta\tau_s\Delta s\, h_{s|z}(s^0|z)ds\,dH_z(z)$$
$$- \int_{s\geq s^0}\int_z T_s'(s)\mathbb{E}\left[\frac{\eta_{s|z}(\theta)}{1+T_s'(s)} + s'_{inc}(\theta)\frac{\eta_z(\theta)}{1-T_z'(z)}\Big|\theta\in\Theta(s,z)\right]\Delta\tau_s\Delta s\, h_{s|z}(s|z)ds\,dH_z(z) \quad (150)$$

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

$$\frac{1}{\lambda}\frac{d\mathcal{L}}{\Delta s\Delta\tau_s} = \int_{s\geq s^0}\int_z\left\{\mathbb{E}\left[1-\hat{g}(\theta)\Big|\theta\in\Theta(s^0,z)\right]\right\}h_{s|z}(s|z)dsdH_z(z) \quad (151)$$
$$- \int_z T_z'(z)\,\mathbb{E}\left[\frac{z\zeta^c_z(\theta)s'_{inc}(\theta)}{1-T_z'(z)}\Big|\theta\in\Theta(s^0,z)\right]h_{s|z}(s^0|z)\,dH_z(z)$$
$$- T_s'(s^0)\int_z \mathbb{E}\left[\frac{s^0\zeta^c_{s|z}(\theta)}{1+T_s'(s^0)} + s'_{inc}(\theta)\frac{z\zeta^c_z(\theta)s'_{inc}(\theta)}{1-T_z'(z)}\Big|\theta\in\Theta(s^0,z)\right]h_{s|z}(s^0|z)ds\,dH_z(z)$$

where $\hat{g}(\theta)$, defined in (17), represent social marginal welfare weights augmented with the fiscal impact of income effects.

Optimal SN tax rate on s. Characterizing the optimal marginal tax rate on s, through $\frac{d\mathcal{L}}{\Delta s\Delta\tau_s} = 0$, it satisfies at each savings s^0,

$$\frac{T_s'(s^0)}{1+T_s'(s^0)}\mathbb{E}\left[s\zeta^c_{s|z}\Big|s^0\right]h_s(s^0) = \int_{s\geq s^0}\mathbb{E}\left[1-\hat{g}\Big|s\right]dH_s(s) - \mathbb{E}\left[FE_z\cdot s'_{inc}\Big|s^0\right]h_s(s^0) \quad (152)$$

where $FE_z(\theta)$ is the fiscal externality from earnings substitution effects defined in (23). This is equation (63) in Proposition A3, valid for any (potentially suboptimal) nonlinear earnings tax schedule $T_z(z)$.

Using

$$\mathbb{E}\left[FE_z\cdot s'_{inc}\Big|s\right] = \mathbb{E}\left[FE_z\Big|s\right]\mathbb{E}\left[s'_{inc}\Big|s\right] + Cov\left[FE_z, s'_{inc}\Big|s\right], \quad (153)$$

we can rewrite this as

$$\frac{T_s'(s^0)}{1+T_s'(s^0)} = \frac{1}{\mathbb{E}\left[s\zeta^c_{s|z}\Big|s^0\right]}\left\{\frac{1}{h_s(s^0)}\int_{s\geq s^0}\mathbb{E}\left[1-\hat{g}\Big|s\right]dH_s(s)\right. \quad (154)$$
$$\left. - \overline{s'_{inc}}(s^0)\mathbb{E}\left[FE_z\Big|s^0\right] - Cov\left[FE_z, s'_{inc}\Big|s^0\right]\right\}.$$

This expression cannot be simplified further using the optimality condition for marginal tax rates on z. It is equation (35) in Proposition 4.

C.G.3 Linear earnings-dependent (LED) tax system

In our derivations of the optimal marginal tax rates on s for general smooth tax systems (Appendix C.C.2), we use a reform that consists of increasing the tax rate on s by $\Delta\tau_s\Delta z$ for all individuals with $s\geq s^0$ and $z\geq z^0$, phased-in over the earnings bandwidth $\left[z^0, z^0+\Delta z\right]$. In a LED system, we impose a linear tax rate on s for all individuals with the same earnings z, regardless of their level s. We thus consider a particular case of the previous reform where $s_0=0$ such that $H_{s|z}(s^0|z)=0$. This leads to the following characterization of optimal LED tax rates on s,

$$\mathbb{E}\left[FE_z\cdot s\Big|z^0\right]h_z(z^0) + \int_{z\geq z^0}\mathbb{E}\left[FE_z\cdot s'_{inc} + \frac{\tau_s(z)}{1+\tau_s(z)}s\zeta^c_{s|z}\Big|z\right]dH_z(z) = \int_{z\geq z^0}\mathbb{E}\left[(1-\hat{g})\cdot s\Big|z\right]dH_z(z). \quad (155)$$

which we can rewrite as

$$\int_{z\geq z^0} \mathbb{E}\left[\frac{\tau_s(z)}{1+\tau_s(z)} s\zeta^c_{s|z}\Big|z\right] dH_z(z) = \int_{z\geq z^0} \mathbb{E}\left[(1-\hat{g})\cdot s\Big|z\right] dH_z(z) - \mathbb{E}\left[FE_z\cdot s\Big|z^0\right] h_z(z^0) - \int_{z\geq z^0} \mathbb{E}\left[FE_z\cdot s'_{inc}\Big|z\right] dH_z(z) \tag{156}$$

This is equation 64 in Proposition A3, valid for any (potentially suboptimal) nonlinear earnings tax schedule $T_z(z)$.

Assuming the nonlinear earnings tax is optimal, we can follow similar steps as in our derivations of the optimal marginal tax rates on s for general smooth tax systems (Appendix C.C.2), to obtain

$$\frac{\tau_s(z^0)}{1+\tau_s(z^0)} = \frac{1}{\mathbb{E}[s\zeta^c_{s|z}|z^0]}\Bigg\{\left(\bar{s}'(z^0) - \overline{s'_{inc}}(z^0)\right)\frac{1}{h_z(z^0)}\int_{z\geq z^0}\mathbb{E}\left[1-\hat{g}\Big|z\right] dH_z(z) - Cov\left[\hat{g}, s\Big|z^0\right] - Cov\left[FE_z, s'_{inc}\Big|z^0\right] + \frac{1}{h_z(z^0)}\frac{d}{dz^0}\left(Cov\left[FE_z, s\Big|z^0\right]\, h_z(z^0)\right)\Bigg\}. \tag{157}$$

which is equation (67) in Proposition 4.

C.H Proof of Proposition A5 (Multiple Goods)

C.H.1 Setting and definitions

The problem of the government is to maximize the following Lagrangian

$$\mathcal{L} = \int_z \left\{\alpha(z)\, U\Big(z - \mathcal{T}(\boldsymbol{s}(z), z) - \sum_{i=1}^n s_i(z), \boldsymbol{s}(z), z; \vartheta(z)\Big) + \lambda\mathcal{T}(\boldsymbol{s}(z), z) - E\right\} dH_z(z) \tag{158}$$

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

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

$$\zeta^c_{s_i|z}(\theta) := -\frac{1+\mathcal{T}'_{s_i}(\boldsymbol{s}(z;\theta), z)}{s_i(z;\theta)}\frac{\partial s_i(z;\theta)}{\partial \mathcal{T}'_{s_i}(\boldsymbol{s}(z;\theta), z)}\Big|_{z=z(\theta)} \tag{159}$$

$$\xi^c_{s_{j,i}|z}(\theta) := \frac{\mathcal{T}'_{s_i}(\boldsymbol{s}(z;\theta), z)}{s_j(z;\theta)}\frac{\partial s_j(z;\theta)}{\partial \mathcal{T}'_{s_i}(\boldsymbol{s}(z;\theta), z)}\Big|_{z=z(\theta)} \tag{160}$$

and in terms of the following statistics,

$$s'_{i,inc}(\theta) := \frac{\partial s_i(z;\theta)}{\partial z}\Big|_{z=z(\theta)} \tag{161}$$

$$\hat{g}(\theta) := \left[\alpha(z)\frac{U'_c(z)}{\lambda} - \left(\mathcal{T}'_z(\boldsymbol{s}(z), z) + \sum_{i=1}^n s'_{i,inc}(z)\mathcal{T}'_{s_i}(\boldsymbol{s}(z), z)\right)\frac{\partial z(.)}{\partial \mathcal{T}} - \sum_{i=1}^n \mathcal{T}'_{s_i}(\boldsymbol{s}(z), z)\frac{\partial s_i(.)}{\partial \mathcal{T}}\right]\Big|_{z=z(\theta)}. \tag{162}$$

C.H.2 Optimal marginal tax rates on earnings z

We consider a small reform at earnings level z^0 that consists in a small increase $\Delta\tau_z$ of the marginal earnings tax rate $\mathcal{T}'_z$ in a small bandwidth Δz. The impact of this reform on the Lagrangian as $\Delta z \to 0$ is

$$\begin{aligned}\frac{1}{\lambda}\frac{d\mathcal{L}}{\Delta\tau_z\Delta z} &= \int_{x\geq z^0}\Big(1-\alpha(x)\frac{U'_c(x)}{\lambda}\Big)dH_z(x) \\ &+ \mathcal{T}'_z\left(\boldsymbol{s}(z^0),z^0\right)\frac{\partial z(.)}{\partial\mathcal{T}'_z}\Big|_{z=z^0}h_z(z^0) + \int_{x\geq z_0}\mathcal{T}'_z\left(\boldsymbol{s}(x),x\right)\frac{\partial z(.)}{\partial\mathcal{T}}\Big|_{z=x}dH_z(x) \\ &+ \sum_{i=1}^{n}\mathcal{T}'_{s_i}\left(\boldsymbol{s}(z^0),z^0\right)s'_{i,inc}(z^0)\frac{\partial z(.)}{\partial\mathcal{T}'_z}\Big|_{z=z^0}h_z(z^0) \\ &+ \int_{x\geq z_0}\sum_{i=1}^{n}\mathcal{T}'_{s_i}\left(\boldsymbol{s}(x),x\right)\left[\frac{\partial s_i(.)}{\partial\mathcal{T}}\Big|_{z=x} + s'_{i,inc}(x)\frac{\partial z(.)}{\partial\mathcal{T}}\Big|_{z=x}\right]dH_z(x).\end{aligned} \tag{163}$$

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

$$-\left[\mathcal{T}'_z\left(\boldsymbol{s}(z^0),z^0\right) + \sum_{i=1}^{n}\mathcal{T}'_{s_i}\left(\boldsymbol{s}(z^0),z^0\right)s'_{i,inc}(z^0)\right]\frac{\partial z(.)}{\partial\mathcal{T}'_z}\Big|_{z=z^0}h_z(z_0) = \int_{x\geq z_0}\Big(1-\hat{g}(x)\Big)dH_z(x). \tag{164}$$

C.H.3 Optimal marginal tax rates on good i

We consider a small reform at earnings level z^0 that consists in adding a linear tax rate $\Delta\tau_s\Delta z$ on s_i for all individuals with earnings z above z^0, phased-in over the earnings bandwidth $\left[z^0, z^0+\Delta z\right]$. In the bandwidth $\left[z^0, z^0+\Delta z\right]$, this reform induces labor supply distortions on earnings z. At earnings $z \geq z^0+\Delta z$, this reform 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}.[57]

The impact of this reform on the Lagrangian as $\Delta z \to 0$ is

$$\begin{aligned}\frac{1}{\lambda}\frac{d\mathcal{L}}{\Delta\tau_s\Delta z} &= \int_{x\geq z^0}\Big(1-\alpha(x)\frac{U'_c(x)}{\lambda}\Big)s_i(x)dH_z(x) \\ &+ \mathcal{T}'_z\left(\boldsymbol{s}(z^0),z^0\right)\frac{\partial z(.)}{\partial\mathcal{T}'_z}\Big|_{z=z^0}s_i(z^0)h_z(z^0) + \int_{x\geq z^0}\mathcal{T}'_z\left(\boldsymbol{s}(x),x\right)\left[\frac{\partial z(.)}{\partial\mathcal{T}'_z}s'_{i,inc}(x) + \frac{\partial z(.)}{\partial\mathcal{T}}\Big|_{z=x}s_i(x)\right]dH_z(z) \\ &+ \sum_{j=1}^{n}\mathcal{T}'_{s_j}\left(\boldsymbol{s}(z^0),z^0\right)\left[s'_{j,inc}(z^0)\frac{\partial z(.)}{\partial\mathcal{T}'_z}\Big|_{z=z^0}s_i(z^0)\right]h_z(z^0) \\ &+ \int_{x\geq z^0}\sum_{j=1}^{n}\mathcal{T}'_{s_j}\left(\boldsymbol{s}(x),x\right)\left\{\frac{\partial s_j(.)}{\partial\mathcal{T}'_{s_i}}\Big|_{z=x} + \frac{\partial s_j(.)}{\partial\mathcal{T}}\Big|_{z=x}s_i(x) + s'_{j,inc}(x)\left[\frac{\partial z(.)}{\partial\mathcal{T}'_z}s'_{i,inc}(x) + \frac{\partial z(.)}{\partial\mathcal{T}}\Big|_{z=x}s_i(x)\right]\right\}dH_z(z)\end{aligned} \tag{165}$$

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

[57] Applying Lemma 1, which still holds in this setting, changes in z and s_j at earnings z^0 and above earnings z^0 are respectively

$$\begin{cases} dz = \frac{\partial z(.)}{\partial\mathcal{T}'_z}\Delta\tau_s\, s_i(z^0) \\ ds_j = s'_{j,inc}(z^0)dz \end{cases} \quad \text{and} \quad \begin{cases} dz = \frac{\partial z(.)}{\partial\mathcal{T}'_z}\Delta\tau_s\Delta z\, s'_{i,inc}(z) + \frac{\partial z(.)}{\partial\mathcal{T}}\Delta\tau_s\Delta z\, s_i(z) \\ ds_j = \frac{\partial s_j(.)}{\partial\mathcal{T}'_{s_i}}\Delta\tau_s\Delta z + \frac{\partial s_j(.)}{\partial\mathcal{T}}\Delta\tau_s\Delta z\, s_i(z) + s'_{j,inc}(z)dz \end{cases}$$

fiscal impacts of income effects $\hat{g}(x)$, we obtain

$$-\left[\mathcal{T}_z'\left(\boldsymbol{s}(z^0),z^0\right)+\sum_{j=1}^{n}\mathcal{T}_{s_j}'\left(\boldsymbol{s}(z^0),z^0\right)s_{j,inc}'(z^0)\right]\frac{\partial z(.)}{\partial \mathcal{T}_z'}\Big|_{z=z^0}s_i(z^0)h_z(z^0)=\int_{x\geq z^0}\Big(1-\hat{g}(x)\Big)s_i(x)dH_z(z)$$
$$+\int_{x\geq z^0}\left\{\left[\mathcal{T}_z'\left(\boldsymbol{s}(x),x\right)+\sum_{j=1}^{n}\mathcal{T}_{s_j}'\left(\boldsymbol{s}(x),x\right)s_{j,inc}'(x)\right]\frac{\partial z(.)}{\partial \mathcal{T}_z'}\Big|_{z=x}s_{i,inc}'(x)+\sum_{j=1}^{n}\mathcal{T}_{s_j}'\left(\boldsymbol{s}(x),x\right)\frac{\partial s_j(.)}{\partial \mathcal{T}_{s_i}'}\Big|_{z=x}\right\}dH_z(z). \tag{166}$$

C.H.4 Deriving Proposition A5

For any good i, we combine the optimality condition for marginal tax rates on earnings z with the one for marginal tax rates on good i to obtain

$$s_i(z^0)\int_{x\geq z_0}\Big(1-\hat{g}(x)\Big)dH_z(x)=\int_{x\geq z^0}\Big(1-\hat{g}(x)\Big)s_i(x)dH_z(z) \tag{167}$$
$$+\int_{x\geq z^0}\left[\left[\mathcal{T}_z'\left(\boldsymbol{s}(x),x\right)+\sum_{j=1}^{n}\mathcal{T}_{s_j}'\left(\boldsymbol{s}(x),x\right)s_{j,inc}'(x)\right]\frac{\partial z(.)}{\partial \mathcal{T}_z'}\Big|_{z=x}s_{i,inc}'(x)+\sum_{j=1}^{n}\mathcal{T}_{s_j}'\left(\boldsymbol{s}(x),x\right)\frac{\partial s_j(.)}{\partial \mathcal{T}_{s_i}'}\Big|_{z=x}\right]dH_z(z)$$

such that differentiating with respect to earnings z^0 gives after simplification

$$s_i'(z^0)\int_{x\geq z_0}\Big(1-\hat{g}(x)\Big)dH_z(x) \tag{168}$$
$$=-\left[\left[\mathcal{T}_z'\left(\boldsymbol{s}(z^0),z^0\right)+\sum_{j=1}^{n}\mathcal{T}_{s_j}'\left(\boldsymbol{s}(z^0),z^0\right)s_{j,inc}'(z^0)\right]\frac{\partial z(.)}{\partial \mathcal{T}_z'}\Big|_{z=z^0}s_{i,inc}'(z^0)+\sum_{j=1}^{n}\mathcal{T}_{s_j}'\left(\boldsymbol{s}(z^0),z^0\right)\frac{\partial s_j(.)}{\partial \mathcal{T}_{s_i}'}\Big|_{z=z^0}\right]h_z(z^0).$$

Making use of the optimality condition for marginal earnings tax rates, we can substitute the first term on the right-hand side to obtain

$$-\sum_{j=1}^{n}\mathcal{T}_{s_j}'\left(\boldsymbol{s}(z^0),z^0\right)\frac{\partial s_j(.)}{\partial \mathcal{T}_{s_i}'}\Big|_{z=z^0}=\left[s_i'(z^0)-s_{i,inc}'(z^0)\right]\frac{1}{h_z(z^0)}\int_{z\geq z_0}\Big(1-\hat{g}(z)\Big)dH_z(z). \tag{169}$$

Isolating the term relative to $\mathcal{T}_{s_i}'\left(\boldsymbol{s}(z^0),z^0\right)$ on the left-hand side yields the following optimal tax formula

$$-\mathcal{T}_{s_i}'\left(\boldsymbol{s}(z^0),z^0\right)\frac{\partial s_i(.)}{\partial \mathcal{T}_{s_i}'}\Big|_{z=z^0}=\left[s_i'(z^0)-s_{i,inc}'(z^0)\right]\frac{1}{h_z(z^0)}\int_{z\geq z^0}\Big(1-\hat{g}(z)\Big)dH_z(z)+\sum_{j\neq i}\mathcal{T}_{s_j}'\left(z^0\right)\frac{\partial s_j(.)}{\partial \mathcal{T}_{s_i}'}\Big|_{z=z^0} \tag{170}$$

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 $\zeta_{s_i|z}^c$ and the cross elasticity $\xi_{s_{j,i}|z}^c$ to finally obtain

$$\frac{\mathcal{T}_{s_i}'\left(\boldsymbol{s}(z^0),z^0\right)}{1+\mathcal{T}_{s_i}'\left(\boldsymbol{s}(z^0),z^0\right)}=\left[s_i'(z^0)-s_{i,inc}'(z^0)\right]\frac{1}{s_i(z^0)\zeta_{s_i|z}^c(z^0)}\frac{1}{h_z(z^0)}\int_{z\geq z^0}\Big(1-\hat{g}(z)\Big)dH_z(z)$$
$$+\sum_{j\neq i}\frac{\mathcal{T}_{s_j}'\left(\boldsymbol{s}(z^0),z^0\right)}{\mathcal{T}_{s_i}'\left(\boldsymbol{s}(z^0),z^0\right)}\frac{s_j(z^0)\xi_{s_{j,i}|z}^c(z^0)}{s_i(z^0)\zeta_{s_i|z}^c(z^0)} \tag{171}$$

which is the first condition stated in Proposition A5.

To derive the second condition stated in Proposition A5, we substitute the first term on the right-hand side using the optimality condition for marginal tax rates on earnings z to directly obtain

$$\frac{\mathcal{T}'_{s_i}\left(\boldsymbol{s}(z^0),z^0\right)}{1+\mathcal{T}'_{s_i}\left(\boldsymbol{s}(z^0),z^0\right)} = \left[s'_i(z^0)-s'_{i,inc}(z^0)\right]\frac{\mathcal{T}'_z\left(\boldsymbol{s}(z^0),z^0\right)+\sum_{j=1}^n \mathcal{T}'_{s_j}\left(\boldsymbol{s}(z^0),z^0\right)s'_{j,inc}(z^0)}{1-\mathcal{T}'_z\left(\boldsymbol{s}(z^0),z^0\right)}\frac{z^0\zeta_z^c(z^0)}{s_i(z^0)\zeta^c_{s_i|z}(z^0)}$$
$$+\sum_{j\neq i}\frac{\mathcal{T}'_{s_j}\left(\boldsymbol{s}(z^0),z^0\right)}{\mathcal{T}'_{s_i}\left(\boldsymbol{s}(z^0),z^0\right)}\frac{s_j(z^0)\xi^c_{s_{j,i}|z}(z^0)}{s_i(z^0)\zeta^c_{s_i|z}(z^0)}. \tag{172}$$

This completes the proof of Proposition A5.

C.I Proof of Proposition 5 (Bequest Taxation and Behavioral Biases)

C.I.1 Setting

We here provide a sufficient statistics characterization of a smooth tax system $\mathcal{T}(s,z)$ under the following additively separable representation of individuals' preferences

$$U\big(c,s,z;\theta\big)=u\left(c;\theta\right)-k\left(z;\theta\right)+\beta\left(\theta\right)v\left(s;\theta\right), \tag{173}$$

and for a utilitarian government that maximizes

$$\int_\theta\big[U\left(c\left(\theta\right),s\left(\theta\right),z\left(\theta\right);\theta\right)+\nu\left(\theta\right)v\left(s\left(\theta\right);\theta\right)\big]dF\left(\theta\right), \tag{174}$$

where $\nu(\theta)$ parametrizes the degree of misalignment on the valuation of s.

Using the mapping between types θ and earnings z under unidimensional heterogeneity, the Lagrangian of the problem is

$$\mathcal{L}=\int_z\Big[U\left(c\left(z\right),s\left(z\right),z;\vartheta\left(z\right)\right)+\nu\left(z\right)v\left(s\left(z\right);\vartheta\left(z\right)\right)+\lambda\left(\mathcal{T}\left(s,z\right)-E\right)\Big]dH_z\left(z\right). \tag{175}$$

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.

C.I.2 Optimal marginal tax rates on z.

A small reform at earnings z^0 that consists in a small increase $\Delta\tau_z$ of the marginal earnings tax rate in a small bandwidth Δz has the following effect as $\Delta z\to 0$,

$$\frac{1}{\lambda}\frac{d\mathcal{L}}{\Delta\tau_z\Delta z}=\int_{z\geq z^0}\left(1-\hat{g}(z)\right)\,dH_z(z) \tag{176}$$
$$-\left(\mathcal{T}'_z(s(z^0),z^0)+s'_{inc}(z^0)\Big(\mathcal{T}'_s(s(z^0),z^0)+\nu(z^0)\frac{v'(s(z^0))}{\lambda}\Big)\right)\frac{z^0}{1-\mathcal{T}'_z(s(z^0),z^0)}\zeta^c_z(z^0)\,h_z(z^0).$$

In this context, social marginal welfare weights augmented with income effects $\hat{g}(z)$ are equal to

$$\hat{g}(z)=\frac{u'\left(c(z)\right)}{\lambda}+\Big(\mathcal{T}'_s\left(s(z),z\right)+\nu(z)\frac{v'\left(s(z)\right)}{\lambda}\Big)\frac{\eta_{s|z}(z)}{1+\mathcal{T}'_s\left(s(z),z\right)} \tag{177}$$

and we can use individuals' first-order condition for s, $(1+\mathcal{T}'_s)\,u'\left(c\right)=\beta v'\left(s\right)$, to express the misalignment wedge in terms of the social marginal welfare weights $g\left(z\right):=\frac{u'(c(z))}{\lambda}$ as

$$\nu(z)\frac{v'\left(s(z)\right)}{\lambda}=\frac{\nu(z)}{\beta(z)}g(z)\left(1+\mathcal{T}'_s\right). \tag{178}$$

The optimal schedule of marginal earnings tax rates is thus characterized by

$$\frac{\mathcal{T}_z'\left(s(z^0), z^0\right)}{1-\mathcal{T}_z'\left(s(z^0), z^0\right)} = \frac{1}{\zeta_z^c(z^0)}\frac{1}{z^0 h_z(z^0)}\int_{z\geq z^0}(1-\hat{g}(z))\,dH_z(z) \\ - s_{inc}'(z^0)\frac{\mathcal{T}_s'\left(s(z^0), z^0\right)}{1-\mathcal{T}_z'\left(s(z^0), z^0\right)} - s_{inc}'(z^0)\frac{\nu(z^0)}{\beta(z^0)}g(z^0)\frac{1+\mathcal{T}_s'\left(s(z^0), z^0\right)}{1-\mathcal{T}_z'\left(s(z^0), z^0\right)}. \tag{179}$$

C.I.3 Optimal marginal tax rates on s.

A small reform at earnings level z^0 that consists in adding a linear tax rate $\Delta\tau_s\Delta z$ on s for all individuals with earnings z above z^0, phased-in over the earnings bandwidth $\left[z^0, z^0+\Delta z\right]$, has the following effect as $\Delta s\to 0$,

$$\frac{1}{\lambda}\frac{d\mathcal{L}}{\Delta\tau_s\Delta z} \tag{180}$$
$$= -\left[\mathcal{T}_z'(s^0, z^0) + s_{inc}'(z^0)\Big(\mathcal{T}_s'(s^0, z^0) + \nu(z^0)\frac{v'(s(z^0))}{\lambda}\Big)\right]\frac{z^0}{1-\mathcal{T}_z'(s^0, z^0)}\zeta_z^c(z^0)\, s^0\, h_z(z^0) \\ + \int_{z\geq z^0}\left\{(1-\hat{g}(z))s(z) - \left[\mathcal{T}_s'(s(z), z) + \nu(z)\frac{v'(s(z))}{\lambda}\right]\frac{s(z)\zeta_{s|z}^c(z)}{1+\mathcal{T}_s'(s(z), z)}\right\}dH_z(z) \\ - \int_{z\geq z^0}\left\{\left[\mathcal{T}_z'(s(z), z) + s_{inc}'(z)\Big(\mathcal{T}_s'(s(z), z) + \nu(z)\frac{v'(s(z))}{\lambda}\Big)\right]\frac{z\zeta_z^c(z)}{1-\mathcal{T}_z'(s(z), z)}s_{inc}'(z)\right\}dH_z(z).$$

We characterize optimal taxes through $d\mathcal{L}=0$. Replacing the misalignment wedge by its expression in terms of social marginal welfare weights $g(z)$, we obtain that the optimal schedule of marginal tax rates on s is characterized by

$$\left[\mathcal{T}_z'(s^0, z^0) + s_{inc}'(z^0)\Big(\mathcal{T}_s'(s^0, z^0) + \frac{\nu(z^0)}{\beta(z^0)}g(z)\left(1+\mathcal{T}_s'(s^0, z^0)\right)\Big)\right]\frac{z^0}{1-\mathcal{T}_z'(s^0, z^0)}\zeta_z^c(z^0)\, s^0\, h_z(z^0) \tag{181}$$
$$= \int_{z\geq z^0}\left\{(1-\hat{g}(z))s(z) - \left[\mathcal{T}_s'(s(z), z) + \frac{\nu(z)}{\beta(z)}g(z)\left(1+\mathcal{T}_s'(s(z), z)\right)\right]\frac{s(z)\zeta_{s|z}^c(z)}{1+\mathcal{T}_s'(s(z), z)}\right\}dH_z(z) \\ - \int_{z\geq z^0}\left\{\left[\mathcal{T}_z'(s(z), z) + s_{inc}'(z)\Big(\mathcal{T}_s'(s(z), z) + \frac{\nu(z)}{\beta(z)}g(z)\left(1+\mathcal{T}_s'(s(z), z)\right)\Big)\right]\frac{z\zeta_z^c(z)}{1-\mathcal{T}_z'}s_{inc}'(z)\right\}dH_z(z).$$

C.I.4 Deriving Proposition 5

Combining optimality conditions for marginal tax rates on z and s yields

$$s(z^0)\int_{z\geq z^0}(1-\hat{g}(z))\,dH_z(z) = \int_{z\geq z^0}(1-\hat{g}(z))s(z)dH_z(z) \tag{182}$$
$$- \int_{z\geq z^0}\left\{\left[\mathcal{T}_z'(s(z), z) + s_{inc}'(z)\left(\mathcal{T}_s'(s(z), z) + \frac{\nu(z)}{\beta(z)}g(z)\left(1+\mathcal{T}_s'(s(z), z)\right)\right)\right]\frac{z\zeta_z^c(z)}{1-\mathcal{T}_z'}s_{inc}'(z)\right\}dH_z(z) \\ - \int_{z\geq z^0}\left\{\left[\mathcal{T}_s'(s(z), z) + \frac{\nu(z)}{\beta(z)}g(z)\left(1+\mathcal{T}_s'(s(z), z)\right)\right]\frac{s(z)\zeta_{s|z}^c(z)}{1+\mathcal{T}_s'(s(z), z)}\right\}dH_z(z).$$

Differentiating with respect to z^0, we obtain after simplification

$$s'(z^0)\int_{z\geq z^0}(1-\hat{g}(z))\,dH_z(z) \tag{183}$$
$$=\left\{\left[\mathcal{T}_z'(s(z^0),z^0)+s_{inc}'(z^0)\Big(\mathcal{T}_s'(s(z^0),z^0)+\frac{\nu(z^0)}{\beta(z^0)}g(z^0)\left(1+\mathcal{T}_s'(s(z^0),z^0)\right)\Big)\right]\frac{z^0\zeta_z^c(z^0)}{1-\mathcal{T}_z'}s_{inc}'(z^0)\right\}h_z(z^0)$$
$$+\left\{\left[\mathcal{T}_s'(s(z^0),z^0)+\frac{\nu(z^0)}{\beta(z^0)}g(z^0)\left(1+\mathcal{T}_s'(s(z^0),z^0)\right)\right]\frac{s(z^0)\zeta_{s|z}^c(z^0)}{1+\mathcal{T}_s'(s(z^0),z^0)}\right\}h_z(z^0).$$

Substituting the first term on the right-hand side by its expression from the optimality condition for marginal tax rates on z, and rearranging we obtain

$$\frac{\mathcal{T}_s'(s(z^0),z^0)}{1+\mathcal{T}_s'(s(z^0),z^0)}+\frac{\nu(z^0)}{\beta(z^0)}g(z^0)=\Big(s'(z^0)-s_{inc}'(z^0)\Big)\frac{1}{s(z^0)\zeta_{s|z}^c(z^0)}\frac{1}{h_z(z^0)}\int_{z\geq z^0}(1-\hat{g}(z))\,dH_z(z) \tag{184}$$

which is the first optimality condition in Proposition 5.

Conversely, substituting the term on the left-hand side by its expression from the optimality condition for marginal tax rates on z, and rearranging we obtain

$$\left[\mathcal{T}_s'(s(z^0),z^0)+\frac{\nu(z^0)}{\beta(z^0)}g(z^0)\left(1+\mathcal{T}_s'(s(z^0),z^0)\right)\right]\frac{s(z^0)\zeta_{s|z}^c(z^0)}{1+\mathcal{T}_s'(s(z^0),z^0)} \tag{185}$$
$$=\Big(s'(z^0)-s_{inc}'(z^0)\Big)\left[\mathcal{T}_z'(s(z^0),z^0)+s_{inc}'(z^0)\Big(\mathcal{T}_s'(s(z^0),z^0)+\frac{\nu(z^0)}{\beta(z^0)}g(z^0)\left(1+\mathcal{T}_s'(s(z^0),z^0)\right)\Big)\right]\frac{z^0\zeta_z^c(z^0)}{1-\mathcal{T}_z'(s(z^0),z^0)}$$

which is the second optimality condition in Proposition 5.

C.J Proof of Proposition 6 (Multidimensional Range with Heterogeneous Prices)

C.J.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 particular, we consider simple tax systems of the SN type, where the tax on s is nonlinear but independent of earnings z such that $T_2(s,z)=T_2(s)$, and of the LED type, where the tax on s is linear but earnings-dependent such that $T_2(s,z)=\tau_s(z)\,s$.

In this setting, we can write type θ's problem as

$$\max_{c,s,z}\ U(c,s,z;\theta)\text{ s.t. }c+p(z,\theta)s\leq z-T_1(z)-p(z,\theta)T_2(s,z) \tag{186}$$
$$\iff \max_z\left\{\max_s\ U\Big(z-T_1(z)-p(z,\theta)\left(s+T_2(s,z)\right),s,z;\theta\Big)\right\} \tag{187}$$

where the inner problem leads to consumption choices $c\,(z;\theta)$ and $s\,(z;\theta)$, and the outer problem leads to an earnings choice $z\,(\theta)$. Assuming $z\,(\theta)$ continues to be a bijective mapping, we again denote $\vartheta\,(z)$ its inverse. 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.[58]

Let λ_1 and λ_2 be the marginal values of public funds associated with the resource constraints

$$\int_z T_1(z)dH_z(z)\geq E_1 \tag{188}$$
$$\int_z T_2(s(z),z)dH_z(z)\geq E_2. \tag{189}$$

[58]When taking derivatives, the presence of these two arguments is implicit. For instance, a total derivative corresponds to $\frac{dp}{dz}:=\frac{\partial p}{\partial z}+\frac{\partial p}{\partial\theta}\frac{\partial\theta}{\partial z}$, whereas a partial derivative $\frac{\partial p}{\partial z}$ represents variation in only the first argument.

The problem of the government is to maximize the Lagrangian

$$\mathcal{L} = \int_z \Big\{ \alpha(z)\, U\Big(z - T_1(z) - p(z)\big(s(z) + T_2(s(z), z)\big), s(z), z; \vartheta(z)\Big) + \lambda_1 T_1(z) + \lambda_2 T_2(s(z), z) - E_1 - E_2 \Big\} dH_z(z). \quad (190)$$

C.J.2 Adapting Lemma 1

Lemma A1. *For a type $\theta = \vartheta(z)$, we have that:*
(1a) *a small increase $\Delta\tau_z$ in the marginal tax rate $\frac{\partial T_2}{\partial z}$ generates the same earnings change as a small increase $p(z)\Delta\tau_z$ in the marginal tax rate $\frac{\partial T_1}{\partial z}$.*
(1b) *a small increase $\Delta\tau_s$ in the marginal tax rate $\frac{\partial T_2}{\partial s}$ generates the same earnings change as a small increase $p(z)s'_{inc}(z)\Delta\tau_s$ in the marginal tax rate $\frac{\partial T_1}{\partial z}$.*
(2) *a small increase ΔT in the T_2 tax liability faced by type $\theta = \vartheta(z)$ generates the same earnings change as a small increase $p(z)\Delta T$ in the T_1 tax liability.*

Proof. We first derive an abstract characterization that we then apply to different tax reforms.

Let type θ indirect utility function at earnings z be

$$V\left(T_1(z), T_2(., z), z; \theta\right) := \max_s\ U\Big(z - T_1(z) - p(z,\theta)\left(s + T_2(s,z)\right), s, z; \theta\Big). \quad (191)$$

Consider a small reform $\Delta T_2(s, z)$ of T_2, and construct for each type θ a perturbation $\Delta T_1^\theta(z)$ 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,

$$V(T_1(z) + \Delta T_1^\theta(z), T_2(., z), z; \theta) = V(T_1(z), T_2(., z) + \Delta T_2(., z), z; \theta). \quad (192)$$

Then, by construction, the perturbation $\Delta T_1^\theta(z)$ induces the same earnings response dz as the initial perturbation $\Delta T_2(., z)$. Moreover, both tax reforms must induce the same utility change for type θ. Applying the envelope theorem yields

$$-U'_c(z;\theta) \cdot \Delta T_1^\theta(z) = -U'_c(z;\theta)\, p(z,\theta) \cdot \Delta T_2\left(s(z;\theta), z\right) \quad (193)$$

such that finally, the perturbation $\Delta T_1^\theta(z)$ is

$$\Delta T_1^\theta(z) = p(z,\theta) \cdot \Delta T_2\left(s(z;\theta), z\right). \quad (194)$$

and we can now apply this abstract characterization to different tax reforms.

(1a) Consider a small increase $\Delta\tau_z$ in the marginal tax rate $\frac{\partial T_2}{\partial z}$ over a small bandwidth of income $\left[z^0, z^0 + \Delta z\right]$. Then, for any type θ such that $z(\theta) \in \left[z^0, z^0 + \Delta z\right]$, we have $\Delta T_2\left(s(z;\theta), z\right) = \Delta\tau_z\left(z - z^0\right)$ such that $\Delta T_1^\theta(z) = p(z,\theta)\Delta\tau_z\left(z - z^0\right)$ and differentiating with respect to z we get

$$\left(\Delta T_1^\theta(z)\right)' = \frac{\partial p(z,\theta)}{\partial z}\Delta\tau_z\left(z - z^0\right) + p(z,\theta)\Delta\tau_z. \quad (195)$$

At the limit $\Delta z \to 0$ such that $z \to z^0$, a small increase $\Delta\tau_z$ in the marginal tax rate $\frac{\partial T_2}{\partial z}$ generates the same earnings change as a small increase $p(z)\Delta\tau_z$ in the marginal tax rate $T_1'(z)$.

(1b) Consider a small increase $\Delta\tau_s$ in the marginal tax rate $\frac{\partial T_2}{\partial s}$ over a small bandwidth of savings $\left[s^0, s^0 + \Delta s\right]$. Then, for any type θ such that $s(\theta) \in \left[s^0, s^0 + \Delta s\right]$, we have $\Delta T_2\left(s(z;\theta), z\right) = \Delta\tau_s\left(s(z;\theta) - s^0\right)$ such that $\Delta T_1^\theta(z) = p(z,\theta)\Delta\tau_z\left(s(z;\theta) - s^0\right)$ and differentiating with respect to z we get

$$\left(\Delta T_1^\theta(z)\right)' = \frac{\partial p(z,\theta)}{\partial z}\Delta\tau_z\left(s(z;\theta) - s^0\right) + p(z,\theta)\Delta\tau_z s'_{inc}(z). \quad (196)$$

At the limit $\Delta s \to 0$ such that $s \to s^0$, a small increase $\Delta\tau_s$ in the marginal tax rate $\frac{\partial T_2}{\partial s}$ generates the same earnings change as a small increase $p(z)s'_{inc}(z)\Delta\tau_z$ in the marginal tax rate $T_1'(z)$.

(2) Consider a small lump-sum increase ΔT in the T_2 tax liability for a type θ who earns z, we then have $\Delta T_1^{\theta}(z) = p(z,\theta)\Delta T$ such that the equivalent reform is no longer a lump-sum increase. Hence, a small increase ΔT in the T_2 tax liability faced by a type $\vartheta(z)$ generates the same earnings change as a small increase $p(z)\Delta T$ in the T_1 tax liability. □

C.J.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:

$$\int_{x\geq 0}\left\{-\alpha(x)U_c'(x)+\lambda_1+\left(\lambda_1 T_1'(x)+\lambda_2\frac{\partial T_2}{\partial z}+s_{inc}'(x)\lambda_2\frac{\partial T_2}{\partial s}\right)\frac{\partial z(.)}{\partial T_1}+\lambda_2\frac{\partial T_2}{\partial s}\frac{\partial s(.)}{\partial T_1}\right\}dH_z(x)=0 \tag{197}$$

$$\int_{x\geq 0}\left\{-\alpha(x)p(x)U_c'(x)+\lambda_2+\left(\lambda_1 T_1'(x)+\lambda_2\frac{\partial T_2}{\partial z}+s_{inc}'(x)\lambda_2\frac{\partial T_2}{\partial s}\right)\frac{\partial z(.)}{\partial T_2}+\lambda_2\frac{\partial T_2}{\partial s}\frac{\partial s(.)}{\partial T_2}\right\}dH_z(x)=0 \tag{198}$$

where $z(.)$ and $s(.)$ denote, with a slight abuse of notation, 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 A1, $\frac{\partial z(.)}{\partial T_2} = \frac{\partial z(.)}{\partial T_1}p(z) + \frac{\partial z(.)}{\partial T_1'}\frac{\partial p}{\partial z}$ and that $\frac{\partial s(.)}{\partial T_2} = \frac{\partial s(.)}{\partial T_1}p(z)$ to obtain

$$\int_{x\geq 0}\left\{1-\left[\alpha(x)\frac{U_c'(x)}{\lambda_1}-\left(T_1'(x)+\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial z}+s_{inc}'(x)\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\right)\frac{\partial z(.)}{\partial T_1}-\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\frac{\partial s(.)}{\partial T_1}\right]\right\}dH_z(z)=0 \tag{199}$$

$$\begin{aligned}\int_{x\geq 0}&\left\{\frac{\lambda_2}{\lambda_1}-p(x)\left[\alpha(x)\frac{U_c'(x)}{\lambda_1}-\left(T_1'(x)+\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial z}+s_{inc}'(x)\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\right)\frac{\partial z(.)}{\partial T_1}-\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\frac{\partial s(.)}{\partial T_1}\right]\right.\\ &\left.+\left(T_1'(x)+\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial z}+s_{inc}'(x)\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\right)\frac{\partial z(.)}{\partial T_1'}\frac{\partial p}{\partial z}\right\}dH_z(x)=0.\end{aligned} \tag{200}$$

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

$$\hat{g}(x) := \alpha(x)\frac{U_c'(x)}{\lambda_1}-\left(T_1'(x)+\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial z}+s_{inc}'(x)\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\right)\frac{\partial z(.)}{\partial T_1}-\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\frac{\partial s(.)}{\partial T_1} \tag{201}$$

$$\varphi(x) := \left(T_1'(x)+\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial z}+s_{inc}'(x)\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\right)\frac{\partial z(.)}{\partial T_1'}\frac{\partial p}{\partial z} \tag{202}$$

where all partial derivatives are evaluated at x, we finally obtain

$$\overline{\hat{g}} := \int_{x\geq 0}\hat{g}(x)\,dH_z(x)=1 \tag{203}$$

$$\overline{\hat{g}p-\varphi} := \int_{x\geq 0}\Big(\hat{g}(x)p(x)-\varphi(x)\Big)\,dH_z(x)=\frac{\lambda_2}{\lambda_1}. \tag{204}$$

C.J.4 Optimal tax rates on z

We consider a small reform at earnings level z^0 that consists in a small increase $\Delta\tau_z$ of the marginal earnings tax rate $T_1'(z)$ in a small bandwidth Δz. The impact on the Lagrangian is as $\Delta z \to 0$,

$$\frac{d\mathcal{L}}{\Delta\tau_z\Delta z} = \int_{x\geq z^0}\Big(\lambda_1 - \alpha(x)U_c'(x)\Big)dH_z(x) \tag{205}$$
$$+ \left[\lambda_1 T_1'\left(z^0\right) + \lambda_2\frac{\partial T_2}{\partial z}\Big|_{z=z^0}\right]\frac{\partial z(.)}{\partial T_1'\left(z^0\right)}h_z(z^0) + \int_{x\geq z^0}\left[\lambda_1 T_1'\left(x\right) + \lambda_2\frac{\partial T_2}{\partial z}\right]\frac{\partial z(.)}{\partial T_1}dH_z(x)$$
$$+ \lambda_2\frac{\partial T_2}{\partial s}\Big|_{z=z^0}s_{inc}'(z^0)\frac{\partial z(.)}{\partial T_1'\left(z^0\right)}h_z(z^0) + \int_{x\geq z^0}\lambda_2\frac{\partial T_2}{\partial s}\left[\frac{\partial s(.)}{\partial T_1} + s_{inc}'(x)\frac{\partial z(.)}{\partial T_1}\right]dH_z(x).$$

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

$$-\left[T_1'(z^0) + \frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial z}\Big|_{z^0} + s_{inc}'(z^0)\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\Big|_{z^0}\right]\frac{\partial z(.)}{\partial T_1'(z^0)} = \frac{1}{h_z(z^0)}\int_{x\geq z^0}\Big[1 - \hat{g}(x)\Big]dH_z(x). \tag{206}$$

C.J.5 Optimal tax rates on s

SN tax system. We consider a small reform at $s^0 = s(z^0)$ that consists in a small increase $\Delta\tau_s$ of $\frac{\partial T_2}{\partial s}$, the marginal tax rate on s, in a small bandwidth Δs. Using Lemma 2, we characterize the impact of the reform on the Lagrangian as $\Delta s \to 0$

$$\frac{d\mathcal{L}}{\Delta\tau_s\Delta s} = \int_{x\geq z^0}\Big(\lambda_2 - \alpha(x)p(x)U_c'(x)\Big)dH_z(x) \tag{207}$$
$$+ \left[\lambda_1 T_1'(z^0) + \lambda_2\frac{\partial T_2}{\partial z}\Big|_{z=z^0}\right]\frac{\partial z(.)}{\partial T_1'\left(z^0\right)}s_{inc}'(z^0)p(z^0)\frac{h_z(z^0)}{s'(z^0)}$$
$$+ \int_{x\geq z^0}\left[\lambda_1 T_1'(x) + \lambda_2\frac{\partial T_2}{\partial z}\right]\left(\frac{\partial z(.)}{\partial T_1}p(x) + \frac{\partial z(.)}{\partial T_1'\left(x\right)}\frac{\partial p}{\partial z}\right)dH_z(x)$$
$$+ \lambda_2\frac{\partial T_2}{\partial s}\Big|_{z=z^0}\left[\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\Big|_{z=z^0}\right)} + s_{inc}'(z^0)\frac{\partial z(.)}{\partial T_1'(z^0)}s_{inc}'(z^0)p(z^0)\right]\frac{h_z(z^0)}{s'(z^0)}$$
$$+ \int_{x\geq z^0}\lambda_2\frac{\partial T_2}{\partial s}\left[\frac{\partial s(.)}{\partial T_2} + s_{inc}'(x)\left(\frac{\partial z(.)}{\partial T_1}p(x) + \frac{\partial z(.)}{\partial T_1'\left(x\right)}\frac{\partial p}{\partial z}\right)\right]dH_z(x).$$

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 $s^0 = s(z^0)$:

$$-\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\Big|_{z^0}\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\Big|_{z^0}\right)}h_z(z^0) = s'(z^0)\int_{x\geq z^0}\left\{\frac{\lambda_2}{\lambda_1} - \hat{g}(x)p(x) + \varphi(x)\right\}dH_z(x) \tag{208}$$
$$+ \left[T_1'(z^0) + \frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial z}\Big|_{z^0} + s_{inc}'(z^0)\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\Big|_{z^0}\right]\frac{\partial z(.)}{\partial T_1'\left(z^0\right)}s_{inc}'(z^0)p(z^0)h_z(z^0)$$

LED tax system. We consider a small reform at $s^0 = s(z^0)$ that consists in a small increase $\Delta\tau_s$ of the linear savings tax rate $\tau_s(z)$ phased in over the earnings bandwidth $[z^0, z^0 + \Delta z]$. Using Lemma (2), we characterize the impact of the reform on the Lagrangian as $\Delta z \to 0$

$$\frac{d\mathcal{L}}{\Delta\tau_s\Delta z} = \int_{x\geq z^0}\Big(\lambda_2 - \alpha(x)p(x)U_c'(x)\Big)s(x)\,dH_z(x) \tag{209}$$
$$+\left(\lambda_1 T_1'(z^0)+\lambda_2\tau_s'(z^0)s(z^0)\right)\frac{\partial z(.)}{\partial T_1'(z^0)}p(z^0)s(z^0)\,h_z(z^0)$$
$$+\int_{x\geq z^0}\left(\lambda_1 T_1'(x)+\lambda_2\tau_s'(z^0)s(z^0)\right)\left[\frac{\partial z(.)}{\partial T_1}p(x)s(x)+\frac{\partial z(.)}{\partial T_1'(x)}\left(\frac{\partial p}{\partial z}s(x)+p(x)s_{inc}'(x)\right)\right]dH_z(x)$$
$$+\lambda_2\tau_s(z^0)s_{inc}'(z^0)\left[\frac{\partial z(.)}{\partial T_1'(z^0)}p(z^0)s(z^0)\right]h_z(z^0)$$
$$+\int_{x\geq z^0}\lambda_2\tau_s(x)\left[\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\big|_x\right)}+\frac{\partial s(.)}{\partial T_1}p(x)s(x)+s_{inc}'(x)\left[\frac{\partial z(.)}{\partial T_1}p(x)s(x)+\frac{\partial z(.)}{\partial T_1'(x)}\left(\frac{\partial p}{\partial z}s(x)+p(x)s_{inc}'(x)\right)\right]\right]dH_z(x)$$

since the reform triggers for individuals at z^0 changes in earnings z equivalent to those induced by a $p(z)\,\Delta\tau_s\,s(z)$ increase in $T_1'(z^0)$, and for individuals above z^0 an increase in tax liability equivalent to a $p(z)\,\Delta\tau_s\,\Delta z\,s(z)$ increase in T_1 and a change in marginal earnings tax rates equivalent to a $\left(\frac{\partial p}{\partial z}\,s(z)+p(z)\,s_{inc}'(z)\right)\Delta\tau_s\,\Delta z$ increase in $T_1'(z)$, in addition to the $\Delta\tau_s\Delta z$ increase in the linear tax rate on s.

We characterize optimal taxes through $d\mathcal{L}=0$. Renormalizing by λ_1, we can plug in $\hat{g}(x)$ and $\varphi(x)$ to obtain the following optimality condition for linear earnings-dependent tax rates on s at each earnings z^0

$$-\left(T_1'(z^0)+\frac{\lambda_2}{\lambda_1}\tau_s'(z^0)s(z^0)+\frac{\lambda_2}{\lambda_1}s_{inc}'(z^0)\tau_s(z^0)\right)\frac{\partial z(.)}{\partial T_1'(z^0)}p(z^0)s(z^0)\,h_z(z^0)$$
$$=\int_{x\geq z^0}\left\{\Big(\frac{\lambda_2}{\lambda_1}-\hat{g}(x)p(x)+\varphi(x)\Big)s(x)+\frac{\lambda_2}{\lambda_1}\tau_s(x)\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\big|_x\right)}\right\}dH_z(x) \tag{210}$$
$$+\int_{x\geq z^0}\left(T_1'(x)+\frac{\lambda_2}{\lambda_1}\tau_s'(x)s(x)+\frac{\lambda_2}{\lambda_1}s_{inc}'(x)\tau_s(x)\right)\frac{\partial z(.)}{\partial T_1'(x)}p(x)s_{inc}'(x)\,dH_z(x)$$

C.J.6 Deriving Proposition 6

SN tax system. A two-part SN tax system $\{T_1(z),T_2(s)\}$ thus satisfies two optimality conditions: the optimality condition in equation (206) for $T_1'(z)$ and the optimality condition in equation (208) for $T_2'(s)$. Combining these two conditions, we get that at each earnings z^0, the optimal SN tax system satisfies

$$-\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\Big|_{z^0}\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\big|_{z^0}\right)}=\frac{s'(z^0)}{h_z(z^0)}\int_{x\geq z^0}\left\{\frac{\lambda_2}{\lambda_1}-\hat{g}(x)p(x)+\varphi(x)\right\}dH_z(x) \tag{211}$$
$$-p(z^0)\frac{s_{inc}'(z^0)}{h_z(z^0)}\int_{x\geq z^0}\Big[1-\hat{g}(x)\Big]dH_z(x)$$

Adding and subtracting $p(z^0)\frac{s'(z^0)}{h_z(z^0)}\int_{x\geq z^0}\Big[1-\hat{g}(x)\Big]\,dH_z(x)$ yields

$$-\frac{\lambda_2}{\lambda_1}\frac{\partial T_2}{\partial s}\Big|_{z^0}\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\big|_{z^0}\right)}=p(z^0)\frac{s'(z^0)-s_{inc}'(z^0)}{h_z(z^0)}\int_{x\geq z^0}\Big[1-\hat{g}(x)\Big]\,dH_z(x) \tag{212}$$
$$+\frac{s'(z^0)}{h_z(z^0)}\int_{x\geq z^0}\left\{\frac{\lambda_2}{\lambda_1}-\hat{g}(x)p(x)+\varphi(x)\right\}dH_z(x)-p(z^0)\frac{s'(z^0)}{h_z(z^0)}\int_{x\geq z^0}\Big[1-\hat{g}(x)\Big]\,dH_z(x).$$

Defining $\zeta^c_{s|z}(z) = -\frac{1+p\frac{\partial T_2}{\partial s}\big|_{z^0}}{s}\frac{\partial s(.)}{p\,\partial\left(\frac{\partial T_2}{\partial s}\big|_{z^0}\right)}$ such that $\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\big|_{z^0}\right)} = -\frac{p\,s}{1+p\frac{\partial T_2}{\partial s}\big|_{z^0}}\zeta^c_{s|z}(z)$, we get[59]

$$\begin{aligned} &\frac{\overline{\hat{g}p - \varphi\frac{\partial T_2}{\partial s}\big|_{z^0}}}{1+p(z^0)\frac{\partial T_2}{\partial s}\big|_{z^0}} \\ &= \frac{1}{s(z^0)\zeta^c_{s|z}(z^0)}\frac{1}{h_z(z^0)}\left\{\left(s'(z^0) - s'_{inc}(z^0)\right)\int_{x\geq z^0}\Big[1-\hat{g}(x)\Big]\,dH_z(x) + \frac{s'(z^0)}{p(z^0)}\Big[\Psi(z^0)+\Phi(z^0)\Big]\right\} \end{aligned} \tag{213}$$

where we use $\overline{\hat{g}p-\varphi} = \frac{\lambda_2}{\lambda_1}$ and $\overline{\hat{g}(x)} = 1$ to obtain the additional terms

$$\begin{aligned} \Psi(z^0) :=& \int_{z\geq z^0}\left[\overline{\hat{g}p} - \hat{g}(z)p(z)\right]dH_z(z) - p(z^0)\int_{z\geq z^0}\Big[\bar{\hat{g}} - \hat{g}(z)\Big]dH_z(z) \qquad (214)\\ =& \int_{z\geq z^0}\left[\int_{x\geq 0}\hat{g}(x)p(x)dH_z(x) - \hat{g}(z)p(z)\right]dH_z(z) - p(z^0)\int_{z\geq z^0}\left[\int_{x\geq 0}\hat{g}(x)dH_z(x)-\hat{g}(z)\right]dH_z(z)\\ =& \left(1-H_z(z^0)\right)\int_{x\geq 0}\hat{g}(x)p(x)dH_z(x) - \int_{z\geq z^0}\hat{g}(z)p(z)dH_z(z)\\ &\quad - p(z^0)\left(1-H_z(z^0)\right)\int_{x\geq 0}\hat{g}(x)dH_z(x) - p(z^0)\int_{z\geq z^0}\hat{g}(z)dH_z(z)\\ =& (1-H_z(z^0))\int_{x\geq 0}\hat{g}(x)\Big(p(x)-p(z^0)\Big)dH_z(x) - \int_{x\geq z^0}\hat{g}(x)\Big(p(x)-p(z^0)\Big)dH_z(x)\\ =& (1-H_z(z^0))\int_{x\leq z^0}\hat{g}(x)\Big(p(x)-p(z^0)\Big)dH_z(z) + H_z(z^0)\int_{x\geq z^0}\hat{g}(x)\Big(p(z^0)-p(x)\Big)dH_z(x) \qquad (215)\\ \Phi(z^0) :=& \int_{x\geq z^0}\Big[\varphi(x) - \overline{\varphi(x)}\Big]\,dH_z(x) \qquad (216) \end{aligned}$$

which proves the optimal formula for SN tax systems in Proposition 6.

LED tax system. A two-part LED tax system $\{T_1(z), \tau_s(z)s\}$ thus satisfies two optimality conditions: the optimality condition in equation (206) for $T_1'(z)$ and the optimality condition in equation (210) for $\tau_s(z)$. Combining these two conditions, we get that at each earnings z^0 the optimal LED tax system satisfies

$$\begin{aligned} &p(z^0)s(z^0)\int_{x\geq z^0}\Big[1-\hat{g}(x)\Big]\,dH_z(x) \\ &= \int_{x\geq z^0}\left\{\Big(\frac{\lambda_2}{\lambda_1} - \hat{g}(x)p(x) + \varphi(x)\Big)s(x) + \frac{\lambda_2}{\lambda_1}\tau_s(x)\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\big|_x\right)}\right\}dH_z(x)\\ &+ \int_{x\geq z^0}\left(T_1'(x) + \frac{\lambda_2}{\lambda_1}\tau_s'(x)s(x) + \frac{\lambda_2}{\lambda_1}s'_{inc}(x)\tau_s(x)\right)\frac{\partial z(.)}{\partial T_1'(x)}p(x)s'_{inc}(x)\,dH_z(x). \end{aligned} \tag{217}$$

[59]With homogeneous p, a SN savings tax levied in period-1 dollars $T_s(s)$ is simply equal to $T_s(s) = pT_2(s)$. As a result, this elasticity definition ensures that $\zeta^c_{s|z}(z)$ coincides with the elasticity concept introduced before:

$$\zeta^c_{s|z}(z) = -\frac{1+T_s'(s)}{s}\frac{\partial s(.)}{\partial T_s'(s)} = -\frac{1+pT_2'(s)}{s}\frac{\partial s(.)}{p\partial T_2'(s)}.$$

Differentiating with respect to z^0 yields

$$\begin{aligned}&\Big(p'(z^0)s(z^0)+p(z^0)s'(z^0)\Big)\int_{x\geq z^0}\Big[1-\hat{g}(x)\Big]\,dH_z(x)-p(z^0)s(z^0)\Big[1-\hat{g}(z^0)\Big]\,h_z(z^0)\\&=-\left\{\Big(\frac{\lambda_2}{\lambda_1}-\hat{g}(z^0)p(z^0)+\varphi(z^0)\Big)s(z^0)+\frac{\lambda_2}{\lambda_1}\tau_s(z^0)\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\big|_{z^0}\right)}\right\}h_z(z^0)\\&-\left(T_1'(z^0)+\frac{\lambda_2}{\lambda_1}\tau_s'(z^0)s(z^0)+\frac{\lambda_2}{\lambda_1}s_{inc}'(z^0)\tau_s(z^0)\right)\frac{\partial z(.)}{\partial T_1'(z^0)}p(z^0)s_{inc}'(z^0)\,h_z(z^0).\end{aligned}\tag{218}$$

Using the optimality condition in equation (206) for $T_1'(z)$, the last term is equal to $p(z^0)s_{inc}'(z^0)\int_{x\geq z^0}\left[1-\hat{g}(x)\right]\,dH_z(x)$ at the optimum such that

$$\begin{aligned}&-\frac{\lambda_2}{\lambda_1}\tau_s(z^0)\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\big|_{z^0}\right)}\,h_z(z^0)\\&=p(z^0)s_{het}'(z^0)\int_{x\geq z^0}\Big[1-\hat{g}(x)\Big]\,dH_z(x)+p'(z^0)s(z^0)\int_{x\geq z^0}\Big[1-\hat{g}(x)\Big]\,dH_z(x)\\&+\left\{\frac{\lambda_2}{\lambda_1}-\Big(\hat{g}(z^0)p(z^0)-\varphi(z^0)\Big)-p(z^0)\Big[1-\hat{g}(z^0)\Big]\right\}s(z^0)h_z(z^0).\end{aligned}\tag{219}$$

We can now plug in the elasticity $\frac{\partial s(.)}{\partial\left(\frac{\partial T_2}{\partial s}\big|_{z^0}\right)}=-\frac{p(z^0)s(z^0)}{1+p(z^0)\frac{\partial T_2}{\partial s}\big|_{z^0}}\zeta_{s|z}^c(z^0)$ with $\frac{\partial T_2}{\partial s}\big|_{z^0}=\tau_s(z^0)$ and use the fact that $\overline{\hat{g}p-\varphi}=\frac{\lambda_2}{\lambda_1}$ and $\overline{\hat{g}}=1$ to obtain

$$\begin{aligned}&\overline{\hat{g}p-\varphi}\frac{\tau_s(z^0)}{1+p(z^0)\tau_s(z^0)}\\&=\frac{1}{s(z^0)\zeta_{s|z}^c(z^0)}\frac{1}{h_z(z^0)}\left\{\Big(s'(z^0)-s_{inc}'(z^0)\Big)\int_{x\geq z^0}\Big[1-\hat{g}(x)\Big]\,dH_z(x)+\frac{p'(z^0)}{p(z^0)}s(z^0)\int_{x\geq z^0}\Big[1-\hat{g}(x)\Big]\,dH_z(x)\right\}\\&+\frac{1}{p(z^0)}\frac{1}{\zeta_{s|z}^c(z^0)}\left\{\left[\overline{\hat{g}p}-p(z^0)\overline{\hat{g}}\right]-\left[\overline{\varphi}-\varphi(z^0)\right]\right\}\end{aligned}\tag{220}$$

which proves the optimal formula for LED tax systems in Proposition 6.

D Implementation with Smooth and Simple Tax Systems

D.A Implementation with Smooth Tax Systems

We consider the case where $\theta\in\mathbb{R}$.

Formally, we say that there is across-ability preference heterogeneity for consumption bundles when the marginal rate of substitution between c and s varies with earnings ability. We denote this marginal rate of substitution by $\mathcal{S}(c,s,z;\theta):=\frac{U_s'(c,s,z;\theta)}{U_c'(c,s,z;\theta)}$ and similarly let $\mathcal{Z}(c,s,z;\theta):=\frac{U_z'(c,s,z;\theta)}{U_c'(c,s,z;\theta)}$ be the marginal rate of substitution between consumption c and earnings z. Denoting $\mathcal{S}_\theta'(c,s,z;\theta_0):=\frac{\partial}{\partial\theta}\mathcal{S}(c,s,z;\theta)|_{\theta=\theta_0}$, we formally define preference heterogeneity as follows:

Definition 1. *There is across-ability preference heterogeneity for consumption bundles if some individuals prefer different (c,s) bundles conditional on having the same earnings z; i.e.,*

$$\exists\theta_0,\forall\,(c,s,z)\,,\ \mathcal{S}_\theta'(c,s,z;\theta_0)\neq 0.\tag{221}$$

For instance, in example (1), $\mathcal{S}_\theta'(c,s,z;\theta)>0$ whenever $\delta(\theta)$ covaries positively with $w(\theta)$. Such across-ability preference heterogeneity in consumption bundles is the focus of our baseline results, and for the rest of the paper we will refer to it simply as "preference heterogeneity."

Our implementation result in the presence of preference heterogeneity relies on commonly-assumed regularity conditions.

Assumption A1. *In the optimal incentive-compatible allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_\theta$, we assume that: (i) $z(\theta)$ is a smooth function of θ, (ii) c and s are smooth functions of z and (iii) 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 $\theta' \neq \theta$.*

Assumption A2. *In the optimal incentive-compatible allocation $\mathcal{A} = \{(c(\theta), s(\theta), z(\theta))\}_\theta$, $z(\theta) \neq z(\theta')$ if $\theta \neq \theta'$.*

Assumption A1 is a standard assumption required to apply optimal control methods to characterize the optimal allocation. The more demanding assumption is assumption A2.[60] Lemma A2 presents an extended Spence-Mirrlees condition that can justify assumption A2.

Lemma A2. *Suppose that when $z'(\theta) \neq 0$, the following extended Spence-Mirrlees condition holds at the optimal incentive-compatible allocation $\mathcal{A}$ for any type $\tilde{\theta}$:*

$$\mathcal{S}'_\theta\left(c(\theta), s(\theta), z(\theta); \tilde{\theta}\right) \frac{s'(\theta)}{z'(\theta)} + \mathcal{Z}'_\theta\left(c(\theta), s(\theta), z(\theta); \tilde{\theta}\right) > 0. \tag{222}$$

Suppose, moreover, that assumption A1 holds. Then earnings $z(\theta)$ are strictly increasing with type θ in the optimal incentive-compatible allocation $\mathcal{A}$.

This strict monotonicity property allows us to define the function $\vartheta(z)$, which maps each earnings level z to the type to which it is assigned in the optimal incentive-compatible allocation.

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

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

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

Theorem A1. *Under Assumptions 1 and A1, an optimal incentive-compatible allocation is implementable by a smooth tax system. In this smooth tax system, individuals' 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—for example, by defining $\mathcal{T}(s(\theta), z(\theta)) = z(\theta) - c(\theta) - s(\theta)$ for $\theta \in \Theta$ and letting $\mathcal{T}(s, z) \to \infty$ 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 allows individuals to locally adjust s and z to points not chosen by any other type in the optimal allocation, and thus the set of possible deviations is much larger than when the optimal mechanism can simply disallow certain allocations.

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

$$\mathcal{T}'_s\left(s\left(\theta\right), z\left(\theta\right)\right) = \mathcal{S}\left(c(\theta), s(\theta), z(\theta); \theta\right) - p \tag{223}$$

$$\mathcal{T}'_z\left(s\left(\theta\right), z\left(\theta\right)\right) = \mathcal{Z}\left(c(\theta), s(\theta), z(\theta); \theta\right) + 1. \tag{224}$$

In the presence of preference heterogeneity, individuals' temptation to deviate from their assigned allocation $(c(\theta), s(\theta), z(\theta))$ are stronger under a smooth tax system than under a direct mechanism. For example, suppose that higher types θ have a stronger relative preference for s. If they deviate downward to some other earnings level $z(\theta') < z(\theta)$, then under a

[60] Assuming that $z(\theta) \neq z(\theta')$ for $\theta \neq \theta'$ is effectively equivalent to assuming that z is strictly increasing in θ. Plainly, the strict monotonicity of z implies $z(\theta) \neq z(\theta')$ for $\theta \neq \theta'$. To go in the other direction, simply note that if $z(\theta) \neq z(\theta')$ for $\theta \neq \theta'$ then there is a suitable reordering of θ that does not affect any other behaviors but that ensures that z is strictly monotonic in θ.

direct mechanism they will be forced to choose $s(\theta')$. Under a smooth tax system, however, the deviating type θ will choose $s' > s(\theta')$ at earnings level $z(\theta')$, making this *double deviation* more appealing.

Tax implementation results that involve multidimensional consumption bundles and multidimensional tax systems typically avoid the difficulties associated with double deviations by ruling out the type of preference heterogeneity that we consider here.[61] Thus, to our knowledge, our proof of Theorem A1 is different from typical implementation proofs in the optimal tax literature. The proof, contained in Appendix D.D, 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 successive elements are increasingly convex around the bundles $(s(\theta), z(\theta))$ offered in the optimal mechanism.

In the second step of the proof, we show that for each type θ there exists a k sufficiently large such that this type's second-order conditions hold at $(c(\theta), s(\theta), z(\theta))$. 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 Appendix Lemmas A3 and A4, which characterize individuals' budget constraints and second derivatives of indirect utility functions for any tax system $\mathcal{T}$ that preserves individuals' first-order conditions.

In the third step, we show that there exists a sufficiently large k such that for *all* types θ, $(c(\theta), s(\theta), z(\theta))$ is a *global* optimum 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-Weierstrass 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.

D.B Implementation with 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 $\mathcal{A} = \{(c^*(\theta), s^*(\theta), z^*(\theta))\}_\theta$ when heterogeneity is unidimensional, $\Theta \subset \mathbb{R}$.

First, we define candidate SN and LED tax systems that satisfy type-specific feasibility and individuals' first-order conditions at the optimal allocation. Second, in Proposition A6, we present sufficient conditions under which these SN and LED tax systems also satisfy individuals' second-order conditions at the optimal allocation, implying local optimality. Third, in Proposition A7, 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.

There are interesting differences between SN and LED tax systems in their ability to implement the optimal allocation. Under our baseline assumptions, we have shown that $z^*(\theta)$ is strictly increasing with type (Lemma A2). However, $s^*(\theta)$ may not be monotonic. When the optimal incentive-compatible allocation $\mathcal{A}$ features a monotonic $s^*(\theta)$, we show that implementation by a SN tax system requires relatively weaker conditions than implementation by a LED tax system. However, when the optimal incentive-compatible allocation $\mathcal{A}$ features non-monotonicity in $s^*(\theta)$, we show that a LED tax system may be able to implement the optimal allocation, whereas a SN tax system cannot – the candidate SN tax system is not even well defined. Hence, all implementation results for SN tax systems are made under the following assumption:

Assumption A3. *When the SN system is studied, $s^*(\theta)$ is assumed strictly monotonic (increasing or decreasing) in type θ.*

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

[61] As pointed out by Kocherlakota (2005), Werning (2010), and others, smooth tax systems can also generate double deviations in dynamic settings where there is a discrete set of types. Werning (2010) provides a general implementation proof for a dynamic setting where productivity is smoothly distributed. The setting studied by Werning (2010), and the proof technique, is distinct from ours because time preferences, and thus preferences for period-2 consumption, are assumed homogeneous.

$\mathcal{A} = (c^*(\theta), s^*(\theta), z^*(\theta))_\theta$ as follows:

$$T_s(s^*(\theta)) := \int_{\vartheta=\theta_{min}}^{\theta} \left(U_s'(\vartheta)/U_c'(\vartheta) - 1\right) s^{*\prime}(\vartheta) d\vartheta, \tag{225}$$

$$T_z(z^*(\theta)) := z^*(\theta_{min}) - s^*(\theta_{min}) - c^*(\theta_{min}) + \int_{\vartheta=\theta_{min}}^{\theta} \left(U_z'(\vartheta)/U_c'(\vartheta) + 1\right) s^{*\prime}(\vartheta) d\vartheta. \tag{226}$$

where θ_{min} 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 (223) and (224). By Lemma A3, 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:

$$\tau_s(z^*(\theta)) := U_s'(\theta)/U_c'(\theta) - 1, \tag{227}$$

$$\begin{aligned} T_z(z^*(\theta)) := & z^*(\theta_{min}) - s^*(\theta_{min}) - c^*(\theta_{min}) \\ & + \int_{\vartheta=\theta_{min}}^{\theta} \left(U_z'(\vartheta)/U_c'(\vartheta) + 1\right) s^{*\prime}(\vartheta) d\vartheta - s^*(z) \cdot \left(\tau_s(z) - \tau_s(z^*(\theta_{min}))\right). \end{aligned} \tag{228}$$

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 A6. *Suppose that an allocation satisfies the conditions in Theorem A1. Under the SN tax system defined by Equations* (225) *and* (226)*, each individual's assigned choice of savings and earnings is a local optimum if the following conditions hold at each point in the allocation:*

$$\mathcal{S}_c' \geq 0,\ \mathcal{S}_z' \geq 0,\ \mathcal{S}_\theta' \geq 0 \tag{229}$$

and

$$\mathcal{Z}_c' \leq 0,\ \mathcal{Z}_s' \geq 0,\ \mathcal{Z}_\theta' \geq 0. \tag{230}$$

Under the LED tax system defined by Equations (227) *and* (228)*, each individual'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:*

$$\mathcal{S}_\theta' \geq 0,\ \mathcal{S}_\theta' \leq \frac{z^{*\prime}(\theta)}{s^{*\prime}(\theta)} \mathcal{Z}_\theta',\ \mathcal{S}_\theta' \leq s^{*\prime}(\theta) \left(\mathcal{S} \cdot \mathcal{S}_c' - \mathcal{S}_s'\right). \tag{231}$$

The sufficiency conditions (229) and (230) 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 (227). 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

(228). 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 A7. *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 (225) and (226), and to the candidate LED tax system defined in equations (227) and (228). Because these candidate tax systems are defined in terms of individuals' preferences and optimal allocations, we can then express conditions 1 and 2 fully in terms of individuals' preferences and optimal allocations.

D.C Proof of Lemma A2 (Monotonicity with Preference Heterogeneity)

We show by contradiction that the extended Spence-Mirrlees condition (222) implies that type $\theta_2 > \theta_1$ chooses earnings $z(\theta_2) > z(\theta_1)$. Note that, by Assumption A1, $c(\theta)$, $s(\theta)$, and $z(\theta)$ are smooth functions of θ in the optimal incentive-compatible allocation, and that by Assumption 1 utility U is twice continuously differentiable.

Assume (without loss of generality) that there is an open set $(\theta_1, \theta_2) \in \Theta$ where $z(\theta)$ is decreasing with θ such that $z(\theta_2) < z(\theta_1)$. Then,

$$U(c(\theta_2), s(\theta_2), z(\theta_2); \theta_2) - U(c(\theta_1), s(\theta_1), z(\theta_1); \theta_2)$$

$$= \int_{\theta=\theta_1}^{\theta_2} \left[\frac{dU(c(\theta), s(\theta), z(\theta); \theta_2)}{d\theta} \right] d\theta \tag{232}$$

$$= \int_{\theta=\theta_1}^{\theta_2} U'_c(c(\theta), s(\theta), z(\theta); \theta_2) \Big[c'(\theta) + \mathcal{S}(c(\theta), s(\theta), z(\theta); \theta_2) s'(\theta) + \mathcal{Z}(c(\theta), s(\theta), z(\theta); \theta_2) z'(\theta) \Big] d\theta \tag{233}$$

Now, for each $\theta \in (\theta_1, \theta_2)$ the first-order condition implied by incentive compatibility implies that, at point $(c(\theta), s(\theta), z(\theta))$,

$$U'_c(c(\theta), s(\theta), z(\theta); \theta) c'(\theta) + U'_s(c(\theta), s(\theta), z(\theta); \theta) s'(\theta) + U'_z(c(\theta), s(\theta), z(\theta); \theta) z'(\theta) = 0$$

$$\iff c'(\theta) + \mathcal{S}(c(\theta), s(\theta), z(\theta); \theta) s'(\theta) + \mathcal{Z}(c(\theta), s(\theta), z(\theta); \theta) z'(\theta) = 0. \tag{234}$$

When $z'(\theta) \neq 0$, the extended Spence-Mirrlees condition states that for any θ',

$$\mathcal{S}'_\theta\left(c(\theta), s(\theta), z(\theta); \theta'\right) \frac{s'(\theta)}{z'(\theta)} + \mathcal{Z}'_\theta\left(c(\theta), s(\theta), z(\theta); \theta'\right) > 0$$
$$\iff \mathcal{S}'_\theta\left(c(\theta), s(\theta), z(\theta); \theta'\right) s'(\theta) + \mathcal{Z}'_\theta\left(c(\theta), s(\theta), z(\theta); \theta'\right) z'(\theta) < 0 \tag{235}$$

where the last inequality is reversed because $z'(\theta) < 0$ for $\theta \in (\theta_1, \theta_2)$. When $z'(\theta) = 0$ on some open interval I, part (ii) of Assumption A1 implies that $s'(\theta)$ is constant on that open interval as well, and thus that

$$\mathcal{S}'_\theta\left(c(\theta), s(\theta), z(\theta); \theta'\right) s'(\theta) + \mathcal{Z}'_\theta\left(c(\theta), s(\theta), z(\theta); \theta'\right) z'(\theta) = 0$$

for all $\theta \in I$ and all $\theta' \in \Theta$. This implies that for $\theta_2 > \theta$,

$$c'(\theta) + \mathcal{S}\left(c(\theta), s(\theta), z(\theta); \theta_2\right) s'(\theta) + \mathcal{Z}\left(c(\theta), s(\theta), z(\theta); \theta_2\right) z'(\theta) \leq 0 \tag{236}$$

with equality if $z'(\theta') = 0$ for all $\theta' \in (\theta, \theta_2)$. Since $U'_c > 0$, this means that the integral (233) is weakly negative, and thus that

$$U\left(c(\theta_2), s(\theta_2), z(\theta_2); \theta_2\right) \leq U\left(c(\theta_1), s(\theta_1), z(\theta_1); \theta_2\right). \tag{237}$$

This is a contradiction with the fact that type θ_2 (strictly) prefers its allocation $(c(\theta_2), s(\theta_2), z(\theta_2))$ in the optimal incentive-compatible allocation, which concludes the proof.

D.D Proof of Theorem A1 (Implementation with a Smooth Tax System)

In the appendix, we adopt the notation that individual's allocations 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 A3 and Lemma A4, which we derive at the end of this subsection. Throughout, we adopt $p \equiv 1$ to economize on notations.

With unidimensional heterogeneity, type θ belongs to the compact space $\Theta = [\theta_{min}, \theta_{max}]$. Moreover, there is always a mapping $s^*(z)$ that denotes the savings level associated with earnings level $z = z^*(\theta)$ at the optimal incentive-compatible allocation. We consider without loss of generality the case in which $s(z)$ is strictly increasing; the proof can be adapted to cases with non-monotonic $s(z)$.

Let $z_{max} := z^*(\theta_{max})$ and $z_{min} := z^*(\theta_{min})$ denote the maximal and minimal, respectively, earnings levels in the allocation. Let $s_{max} := \max_z s^*(z)$ and $s_{min} := \min_z s^*(z)$ denote the 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 individuals' 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 the individuals' optimization problems 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. $T_s(s_{min}) = 0$ and $T_z(z_{min}) = z^*(\theta_{min}) - c^*(\theta_{min}) - s^*(\theta_{min})$
2. $\forall z \in [\, z_{min}\, ;\, z_{max}\,], T'_z(z) = \mathcal{Z}(c^*(\theta_z), s^*(\theta_z), z^*(\theta_z); \theta_z) + 1$ with θ_z such that $z = z^*(\theta_z)$
3. $\forall s \in [\, s_{min}\, ;\, s_{max}\,], T'_s(s) = \mathcal{S}(c^*(\theta_s), s^*(\theta_s), z^*(\theta_s); \theta_s) - 1$ with θ_s such that $s = s^*(\theta_s)$

4. $$\mathcal{T}(s,z,k)=\begin{cases} T_s(s)+T_z(z)+k(s-s^*(z))^2 & \text{if } z_{min}\le z\le z_{max}, \\ & s_{min}\le s\le s_{max} \\ T_s(s_{min})+T_z(z)+k(s-s^*(z))^2+T_s'(s_{min})(s-s_{min}) & \text{if } z_{min}\le z\le z_{max}, s<s_{min} \\ T_s(s_{max})+T_z(z)+k(s-s^*(z))^2+T_s'(s_{max})(s-s_{max}) & \text{if } z_{min}\le z\le z_{max}, s>s_{max} \\ T_s(s)+T_z(z_{min})+k(s-s_{min})^2+k(z-z_{min})^2 & \text{if } z<z_{min}, s_{min}\le s\le s_{max} \\ \qquad +T_z'(z_{min})(z-z_{min}) & \\ T_s(s_{min})+T_z(z_{min})+k(s-s_{min})^2+k(z-z_{min})^2 & \text{if } z<z_{min}, s<s_{min} \\ \qquad +T_z'(z_{min})(z-z_{min})+T_s'(s_{min})(s-s_{min}) & \\ T_s(s_{max})+T_z(z_{min})+k(s-s_{min})^2+k(z-z_{min})^2 & \text{if } z<z_{min}, s>s_{max} \\ \qquad +T_z'(z_{min})(z-z_{min})+T_s'(s_{max})(s-s_{max}) & \\ T_s(s)+T_z(z_{max})+k(s-s_{max})^2+k(z-z_{max})^2 & \text{if } z>z_{max}, s_{min}\le s\le s_{max} \\ \qquad +T_z'(z_{max})(z-z_{max}) & \\ T_s(s_{max})+T_z(z_{max})+k(s-s_{max})^2+k(z-z_{max})^2 & \text{if } z>z_{max}, s>s_{max} \\ \qquad +T_z'(z_{max})(z-z_{max})+T_s'(s_{max})(s-s_{max}) & \\ T_s(s_{min})+T_z(z_{max})+k(s-s_{max})^2+k(z-z_{max})^2 & \text{if } z>z_{max}, s<s_{min} \\ \qquad +T_z'(z_{max})(z-z_{max})+T_s'(s_{min})(s-s_{min}) & \end{cases}$$

Assumptions 1 and A1 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 type 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 type θ's 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 A3 (see below).

Now, defining

$$V(s,z;\theta,k):=U(z-s-\mathcal{T}(s,z;k),s,z;\theta), \tag{238}$$

the first-order conditions for type θ's 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 A4 (see below), second-order conditions at $(s^*(\theta), z^*(\theta))$ are

$$V_{ss}''=\frac{U_z'}{s^{*\prime}(z^*)}\mathcal{S}_c'-\frac{U_c'}{s^{*\prime}(z^*)}\mathcal{S}_z'-\frac{U_c'}{s^{*\prime}(\theta)}\mathcal{S}_\theta'+\frac{U_c'}{s^{*\prime}(z^*)}\mathcal{T}_{sz}''\le 0 \tag{239}$$

$$V_{zz}''=U_s's^{*\prime}(z^*)\mathcal{Z}_c'-U_c's^{*\prime}(z^*)\mathcal{Z}_s'-\frac{U_c'}{z^{*\prime}(\theta)}\mathcal{Z}_\theta'+U_c's^{*\prime}(z^*)\mathcal{T}_{sz}''\le 0 \tag{240}$$

$$\begin{aligned}(V_{sz}'')^2-V_{ss}''V_{zz}''=\frac{U_c'}{s^{*\prime}(\theta)}\Bigg[&(U_z'\mathcal{S}_c'-U_c'\mathcal{S}_z')\,\mathcal{Z}_\theta'+\left(U_s'\mathcal{Z}_c'-U_c'\mathcal{Z}_s'-U_c'\frac{\mathcal{Z}_\theta'}{s^{*\prime}(\theta)}\right)s^{*\prime}(z^*)\mathcal{S}_\theta'\\ &+(\mathcal{Z}_\theta'+s^{*\prime}(z^*)\mathcal{S}_\theta')\,U_c'\mathcal{T}_{sz}''\Bigg]\le 0\end{aligned} \tag{241}$$

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

This shows that for a given type 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 type θ's 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 $[z_{min}, z_{max}]$ guarantees that for any $\varepsilon > 0$, there exists k_1, such that for all $k \geq k_1$ and for all types θ, individuals' earnings choices belong to the compact set $[z_{min} - \varepsilon, z_{max} + \varepsilon]$. Indeed, starting from a given earnings level $z > z_{max} + \varepsilon$, 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 $\mathcal{T}'_z = 2k(z - z_{max}) + T'_z(z_{max})$. 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 $z \leq z_{max} + \varepsilon$ for sufficiently large k. Symmetrically, we can show that all individuals choose earnings $z \geq z_{min} - \varepsilon$ for sufficiently large k.

Second, since earnings shape individuals' disposable incomes, the fact that earnings belong to a compact set for sufficiently large k implies that individuals' allocation choices also belong to a compact set. Indeed, for sufficiently large k, individuals' allocation choices must belong to the set of (c, s, z) such that $c \geq 0$, $s \geq 0$, $z \in [z_{min} - \varepsilon, z_{max} + \varepsilon]$, 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 $[\theta_{min}, \theta_{max}]$, we know that the limit type $\hat{\theta}$ must belong to $[\theta_{min}, \theta_{max}]$. Now, from the previous paragraph, individuals' earnings choices have to be arbitrarily close to $[z_{min}, z_{max}]$ as the penalty grows. This implies that the limit $\hat{z}$ must belong to $[z_{min}, z_{max}]$.

Next, we establish that the limit $\hat{s}$ must be such that $\hat{s} = s^*(\hat{z})$. First fix an earnings level $z \in [z_{min}, z_{max}]$. Then, 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 $[s_{min}, s_{max}]$, 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 $k \geq k_2$, type $\hat{\theta}$ chooses savings $s \in [s^*(z) - \varepsilon, s^*(z) + \varepsilon]$ for a fixed z.[62] Since type $\hat{\theta}$'s 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 $\hat{\theta} \in [\theta_{min}, \theta_{max}]$, $\hat{z} \in [z_{min}, z_{max}]$, 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 type $\hat{\theta}$. Hence, $(\hat{c}, \hat{s}, \hat{z}) = (c^*(\hat{\theta}), s^*(\hat{\theta}), z^*(\hat{\theta}))$.

[62] 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 parametrized by k can be made arbitrarily large. As a result, individuals' first-order conditions can never hold for sufficiently large k, while we can similarly exclude corner solutions for sufficiently large k.

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 $k^\dagger$ 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 individuals 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 types 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 $\mathcal{T}_{k^\dagger}$, type $\theta_{j^\dagger}$'s 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 type $\theta_{j^\dagger}$'s maximum under the tax system $\mathcal{T}_{k^\dagger}$. This contradicts the fact established above that type $\theta_{j^\dagger}$ 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 A3. *A smooth tax system $\mathcal{T}$ satisfies type-specific feasibility over the compact type space $[\,\theta_{min}\,;\,\theta_{max}\,]$ if it satisfies the following conditions:*

1. $\mathcal{T}(s^*(\theta_{min}), z^*(\theta_{min})) = z^*(\theta_{min}) - c^*(\theta_{min}) - s^*(\theta_{min})$
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

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

Since $\mathcal{T}(s^*(\theta_{min}), z^*(\theta_{min})) = T^*_\theta(\theta_{min})$, 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 A4. *Consider a smooth tax system $\mathcal{T}$ satisfying the conditions in Lemma A3 and define*

$$V(s, z; \theta) := U(z - s - \mathcal{T}(s, z), s, z; \theta). \tag{242}$$

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

$$V''_{ss} = \frac{U'_z}{s^{*\prime}(z^*)}\mathcal{S}'_c - \frac{U'_c}{s^{*\prime}(z^*)}\mathcal{S}'_z - \frac{U'_c}{s^{*\prime}(\theta)}\mathcal{S}'_\theta + \frac{U'_c}{s^{*\prime}(z^*)}\mathcal{T}''_{sz} \tag{243}$$

$$V''_{zz} = U'_s s^{*\prime}(z^*)\mathcal{Z}'_c - U'_c s^{*\prime}(z^*)\mathcal{Z}'_s - \frac{U'_c}{z^{*\prime}(\theta)}\mathcal{Z}'_\theta + U'_c s^{*\prime}(z^*)\mathcal{T}''_{sz} \tag{244}$$

$$(V''_{sz})^2 - V''_{ss}V''_{zz} = \frac{U'_c}{s^{*\prime}(\theta)}\Bigg[(U'_z\mathcal{S}'_c - U'_c\mathcal{S}'_z)\,\mathcal{Z}'_\theta + \left(U'_s\mathcal{Z}'_c - U'_c\mathcal{Z}'_s - U'_c\frac{\mathcal{Z}'_\theta}{s^{*\prime}(\theta)}\right)s^{*\prime}(z^*)\mathcal{S}'_\theta + \left(\mathcal{Z}'_\theta + s^{*\prime}(z^*)\mathcal{S}'_\theta\right)U'_c\mathcal{T}''_{sz}\Bigg] \tag{245}$$

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

Proof. The first-order derivatives are

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

The second-order derivatives are

$$V''_{ss}(s,z;\theta) = -\mathcal{T}''_{ss}U'_c - (1+\mathcal{T}'_s)\left(-(1+\mathcal{T}'_s)U''_{cc} + U''_{cs}\right) - (1+\mathcal{T}'_s)U''_{cs} + U''_{ss} \tag{246}$$
$$V''_{zz}(s,z;\theta) = -\mathcal{T}''_{zz}U'_c + (1-\mathcal{T}'_z)\left((1-\mathcal{T}'_z)U''_{cc} + U''_{cz}\right) + (1-\mathcal{T}'_z)U''_{cz} + U''_{zz}. \tag{247}$$

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

$$\mathcal{T}''_{ss} = \mathcal{S}'_c\frac{1-\mathcal{T}'_z}{s^{*\prime}(z^*)} - \mathcal{S}'_c(1+\mathcal{T}'_s) + \mathcal{S}'_s + \frac{\mathcal{S}'_z}{s^{*\prime}(z^*)} + \frac{\mathcal{S}'_\theta}{s^{*\prime}(\theta)} - \frac{\mathcal{T}''_{sz}}{s^{*\prime}(z^*)}. \tag{248}$$

Moreover, we differentiate the definition $\mathcal{S} := \frac{U'_s}{U'_c}$ to express the derivative of $\mathcal{S}$ with respect to c as

$$\begin{aligned}\mathcal{S}'_c(c^*(\theta),s^*(\theta),z^*(\theta);\theta) &= \frac{U'_cU''_{sc} - U'_sU''_{cc}}{(U'_c)^2}\\ &= \frac{1}{U'_c}\left(-\frac{U'_s}{U'_c}U''_{cc} + U''_{sc}\right)\\ &= \frac{1}{U'_c}\left(-(1+\mathcal{T}'_s)U''_{cc} + U''_{sc}\right)\end{aligned} \tag{249}$$

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

$$\begin{aligned}\mathcal{S}'_s(c^*(\theta),s^*(\theta),z^*(\theta);\theta) &= \frac{U'_cU''_{ss} - U'_sU''_{cs}}{(U'_c)^2}\\ &= \frac{1}{U'_c}\left(-\frac{U'_s}{U'_c}U''_{cs} + U''_{ss}\right)\\ &= \frac{1}{U'_c}\left(-(1+\mathcal{T}'_s)U''_{cs} + U''_{ss}\right).\end{aligned} \tag{250}$$

Substituting equations (248), (249) and (250) into (246), we have

$$\begin{aligned}V''_{ss}\left(s^*\left(\theta\right),z^*\left(\theta\right);\theta\right) &= -U'_c\cdot\left(\mathcal{S}'_c\frac{1-\mathcal{T}'_z}{s^{*\prime}(z)}-\mathcal{S}'_c(1+\mathcal{T}'_s)+\mathcal{S}'_s+\frac{\mathcal{S}'_z}{s^{*\prime}(z)}+\frac{\mathcal{S}'_\theta}{s^{*\prime}(\theta)}-\frac{\mathcal{T}''_{sz}}{s^{*\prime}(z)}\right)-(1+\mathcal{T}'_s)U'_s\mathcal{S}'_c+U'_c\mathcal{S}'_s\\ &= -U'_c\cdot\left(\frac{1-T'_z}{s^{*\prime}(z)}\mathcal{S}'_c+\frac{1}{s^{*\prime}(z)}\mathcal{S}'_z+\frac{1}{s^{*\prime}(\theta)}\mathcal{S}'_\theta-\frac{\mathcal{T}''_{sz}}{s^{*\prime}(z)}\right)\\ &= \frac{U'_z}{s^{*\prime}(z)}\mathcal{S}'_c-\frac{U'_c}{s^{*\prime}(z^*)}\mathcal{S}'_z-\frac{U'_c}{s^{*\prime}(\theta)}\mathcal{S}'_\theta+\frac{U'_c}{s^{*\prime}(z^*)}\mathcal{T}''_{sz}\end{aligned}\tag{251}$$

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

$$\mathcal{T}''_{zz}=\mathcal{Z}'_c\left(1-\mathcal{T}'_z\right)-\mathcal{Z}'_c\left(1+\mathcal{T}'_s\right)s^{*\prime}(z^*)+\mathcal{Z}'_s s^{*\prime}(z^*)+\mathcal{Z}'_z+\frac{\mathcal{Z}'_\theta}{z^{*\prime}(\theta)}-\mathcal{T}''_{sz}s^{*\prime}(z^*).\tag{252}$$

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

$$V''_{zz}\left(s^*\left(\theta\right),z^*\left(\theta\right);\theta\right)=U'_s s^{*\prime}(z^*)\mathcal{Z}'_c-U'_c s^{*\prime}(z^*)\mathcal{Z}'_s-U'_c\frac{\mathcal{Z}'_\theta}{z^{*\prime}(\theta)}+U'_c\mathcal{T}''_{sz}s^{*\prime}(z^*).\tag{253}$$

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

$$\frac{d}{d\theta}V'_s(s^*(\theta),z^*(\theta);\theta)=V''_{ss}s^{*\prime}(\theta)+V''_{sz}z^{*\prime}(\theta)+V''_{s\theta}=0\tag{254}$$

which we can rearrange as

$$-V''_{sz}=V''_{ss}s^{*\prime}(z^*)+\frac{V''_{s\theta}}{z^{*\prime}(\theta)}.\tag{255}$$

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

$$-V''_{sz}=\frac{V''_{zz}}{s^{*\prime}(z^*)}+\frac{V''_{z\theta}}{s^{*\prime}(\theta)}.\tag{256}$$

Writing $\left(V''_{sz}\right)^2$ as the product of the right-hand sides of equations (255) and (256) yields

$$\begin{aligned}\left(V''_{sz}\right)^2 &= \left(V''_{ss}s^{*\prime}(z)+\frac{V''_{s\theta}}{z^{*\prime}(\theta)}\right)\left(\frac{V''_{zz}}{s^{*\prime}(z)}+\frac{V''_{z\theta}}{s^{*\prime}(\theta)}\right)\\ &= V''_{ss}V''_{zz}+\frac{1}{z^{*\prime}(\theta)}V''_{ss}V''_{z\theta}+\frac{1}{s^{*\prime}(\theta)}V''_{zz}V''_{s\theta}+\frac{1}{s^{*\prime}(\theta)z^{*\prime}(\theta)}V''_{s\theta}V''_{z\theta}.\end{aligned}\tag{257}$$

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 $(c^*(\theta), s^*(\theta), z^*(\theta))$ we have

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

to obtain

$$V''_{s\theta}(s^*(\theta), z^*(\theta); \theta) = U'_c \mathcal{S}'_\theta \tag{258}$$
$$V''_{z\theta}(s^*(\theta), z^*(\theta); \theta) = U'_c \mathcal{Z}'_\theta. \tag{259}$$

Substituting these into equation (257) and rearranging, we have

$$\left(V''_{sz}\right)^2 - V''_{ss}V''_{zz} = \frac{1}{z^{*\prime}(\theta)}V''_{ss}U'_c\mathcal{Z}'_\theta + \frac{1}{s^{*\prime}(\theta)}V''_{zz}U'_c\mathcal{S}'_\theta + \frac{1}{s^{*\prime}(\theta)z^{*\prime}(\theta)}\left(U'_c\right)^2\mathcal{S}'_\theta\mathcal{Z}'_\theta. \tag{260}$$

Expanding V''_{ss} from equation (251), and V''_{zz} from equation (253) 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. □

D.E Proof of Proposition A6 & A7 (Implementation with Simple Tax Systems)

D.E.1 Proof of Proposition A6

SN tax system. The sufficient conditions for local optimality under the candidate SN tax system follow directly from Lemma A4 which computes individuals' second-order conditions (SOCs) at the optimal incentive-compatible allocation under a general tax system $\mathcal{T}(s,z)$. Indeed, individuals' SOCs are satisfied if equations (243), (244), and (245) 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 (229) and (230) 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 individuals' 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 A4 which computes individuals' SOCs at the optimal incentive-compatible allocation. We consider the requirements $V''_{ss} < 0$, $V''_{zz} < 0$, and $V''_{ss}V''_{zz} > \left(V''_{sz}\right)^2$ 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 (246) 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 ($U''_{cc} \le 0$ and $U''_{ss} \le 0$) 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 (227) and (228) 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 (227) 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 (253), we have

$$\begin{aligned}V_{zz}''\left(s^*\left(\theta\right),z^*\left(\theta\right);\theta\right) &= U_s's^{*\prime}(z^*)\mathcal{Z}_c'-U_c's^{*\prime}(z^*)\mathcal{Z}_s'-U_c'\frac{\mathcal{Z}_\theta'}{z^{*\prime}(\theta)} \\ &\quad + U_c's^{*\prime}(z^*)\left[\mathcal{S}_c'\cdot\left(1-\mathcal{T}_z'\right)-\mathcal{S}_c'\cdot\left(1+\mathcal{T}_s'\right)s^{*\prime}(z^*)+\mathcal{S}_s'\cdot s^{*\prime}(z^*)+\mathcal{S}_z'+\frac{\mathcal{S}_\theta'}{z^{*\prime}(\theta)}\right]. \qquad (261)\end{aligned}$$

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 (261), and noting that $\mathcal{Z}_s'=\mathcal{S}_z'=0$ by separability, yields

$$V_{zz}''(s^*(\theta),z^*(\theta);\theta) = \left(s^{*\prime}(z^*)\right)^2\left[U_c'\mathcal{S}_s'-U_s'\mathcal{S}_c'\cdot\right]-\frac{U_c'}{z^{*\prime}(\theta)}\left[\mathcal{Z}_\theta'-s^{*\prime}(z^*)\mathcal{S}_\theta'\right]. \qquad (262)$$

Again employing separability, we have

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

implying that the first term on the right-hand side of equation (262) is negative. The condition $\mathcal{Z}_\theta'-s^{*\prime}(z^*)\mathcal{S}_\theta'\geq 0$ from (231) in the Proposition then implies equation (262) (and thus V_{zz}'') is negative.

Third, to show $V_{ss}''V_{zz}''>\left(V_{sz}''\right)^2$, we proceed from equation (245) in Lemma A4:

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

Recognizing that the last bracket term is exactly the expression for V_{zz}'' given in Lemma A4, 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

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

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 (231) in the Proposition imply that both terms on the right-hand side are negative, implying that all second-order conditions hold.

D.E.2 Proof of Proposition A7

We begin with a more general statement, and then derive Proposition A7 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 A5. *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*

$$\begin{aligned}&U'_c\Big(c^*(\theta), s^*(\theta), z^*(\theta);\theta\Big)\sigma_c\Big(s^*(\theta), z^*(\theta)\Big) + U'_z\Big(c^*(\theta), s^*(\theta), z;\theta\Big)\\ &< U'_c\Big(c\left(z^*(\theta),\theta'\right), s\left(z^*(\theta),\theta'\right), z^*(\theta);\theta\Big)\sigma_c\Big(s\left(z^*(\theta),\theta'\right), z^*(\theta)\Big)\\ &+ U'_z\Big(c\left(z^*(\theta),\theta'\right), s\left(z^*(\theta),\theta'\right), z^*(\theta);\theta\Big)\end{aligned} \tag{263}$$

and

$$\begin{aligned}&U'_s\Big(c^*(\theta), s^*(\theta), z^*(\theta);\theta\Big)\sigma_c\Big(s^*(\theta), z^*(\theta)\Big) + U'_z\Big(c^*(\theta), s^*(\theta), z;\theta\Big)\Big)\\ &< U'_s\Big(c\left(z^*(\theta),\theta'\right), s\left(z^*(\theta),\theta'\right), z^*(\theta);\theta\Big)\sigma_s\Big(s\left(z^*(\theta),\theta'\right), z^*(\theta)\Big)\\ &+ U'_z\Big(c\left(z^*(\theta),\theta'\right), s\left(z^*(\theta),\theta'\right), z^*(\theta);\theta\end{aligned} \tag{264}$$

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, type 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 that individuals' local optima are global optima, we want to show that for any given type 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 type θ^*'s 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

$$\begin{aligned}&\frac{d}{dz}U\left(c(z,\theta^*),s(z,\theta^*),z;\theta^*\right)\\&=U_s'\left(c(z,\theta^*),s(z,\theta^*),z;\theta^*\right)\sigma_s\left(s(z,\theta^*),z\right)+U_z'\left(c(z,\theta^*),s(z,\theta^*),z;\theta^*\right)\end{aligned}\tag{265}$$

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

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

$$\begin{aligned}&U\left(c(z,\theta^*),s(z,\theta^*),z;\theta^*\right)-U\left(c(z^*(\theta^*),\theta^*),s(z^*(\theta^*),\theta^*),z^*(\theta^*);\theta^*\right)\\&=\int_{x=z^*(\theta^*)}^{x=z}\Big[\min\left\{U_c'\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)\sigma_c\left(s(x,\theta^*),x\right),U_s'\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)\sigma_s\left(s(x,\theta^*),x\right)\right\}\\&\qquad+U_z'\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)\Big]dx\end{aligned}\tag{266}$$

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

The second step is to show that under our assumptions, type θ^*'s utility change in equation (266) is negative. To do so, let θ_x be the type that chooses earnings x. Then, by definition, type θ_x's 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

$$\begin{aligned}&\max\left\{U_c'\left(c^*(\theta_x),s^*(\theta_x),x;\theta_x\right)\sigma_c\left(s^*(\theta_x),x\right),U_s'\left(c^*(\theta_x),s^*(\theta_x),x;\theta_x\right)\sigma_s\left(s^*(\theta_x),x\right)\right\}\\&+U_z'\left(c^*(\theta_x),s^*(\theta_x),x;\theta_x\right)=0.\end{aligned}\tag{267}$$

Now, by condition 2, we either have[63]

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

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

$$\begin{aligned}&\max\left\{U_c'\left(c^*(\theta_x),s^*(\theta_x),x;\theta_x\right)\sigma_c\left(s^*(\theta_x),x\right),U_s'\left(c^*(\theta_x),s^*(\theta_x),x;\theta_x\right)\sigma_s\left(s^*(\theta_x),x\right)\right\}\\&+U_z'\left(c^*(\theta_x),s^*(\theta_x),x;\theta_x\right)\\&\geq\min\left\{U_c'\left(c(x,\theta^*),s(x,\theta^*),x;\theta_x\right)\sigma_c\left(s(x,\theta^*),x\right),U_s'\left(c(x,\theta^*),s(x,\theta^*),x;\theta_x\right)\sigma_s\left(s(x,\theta^*),x\right)\right\}\\&+U_z'\left(c(x,\theta^*),s(x,\theta^*),x;\theta_x\right).\end{aligned}\tag{268}$$

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

$$\begin{aligned}&U_c'\left(c(x,\theta^*),s(x,\theta^*),x;\theta_x\right)\sigma_c\left(s(x,\theta^*),x\right)+U_z'\left(c(x,\theta^*),s(x,\theta^*),x;\theta_x\right)\leq0\\&\iff\frac{U_z'\left(c(x,\theta^*),s(x,\theta^*),x;\theta_x\right)}{U_c'\left(c(x,\theta^*),s(x,\theta^*),x;\theta_x\right)}\leq-\sigma_c\left(s(x,\theta^*),x\right)\end{aligned}$$

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

or

$$\begin{aligned}&U_s'\left(c(x,\theta^*),s(x,\theta^*),x;\theta_x\right)\sigma_s\left(s(x,\theta^*),x\right)+U_z'\left(c(x,\theta^*),s(x,\theta^*),x;\theta_x\right)\le 0\\ \iff\ &\frac{U_z'\left(c(x,\theta^*),s(x,\theta^*),x;\theta_x\right)}{U_s'\left(c(x,\theta^*),s(x,\theta^*),x;\theta_x\right)}\le -\sigma_c\left(s(x,\theta^*),x\right).\end{aligned}$$

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[64]

$$\frac{U_z'\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)}{U_c'\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)}\le -\sigma_c\left(s(x,\theta^*),x\right)$$

or

$$\frac{U_z'\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)}{U_s'\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)}\le -\sigma_c\left(s(x,\theta^*),x\right)$$

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

$$\begin{aligned}&\min\left\{U_c'\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)\sigma_c\left(s(x,\theta^*),x\right),U_s'\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)\sigma_s\left(s(x,\theta^*),x\right)\right\}\\ &+U_z'\left(c(x,\theta^*),s(x,\theta^*),x;\theta^*\right)\le 0.\end{aligned}\tag{269}$$

As a result, the right hand-side of equation (266) is an integral of negative terms, which shows that

$$U\left(c(z,\theta^*),s(z,\theta^*),z;\theta^*\right)-U\left(c^*(\theta^*),s^*(\theta^*),z^*(\theta^*);\theta^*\right)\le 0.\tag{270}$$

The case with $z < z^*(\theta^*)$ follows identically, proving Lemma A5. □

Proof of Proposition A7

We now derive Proposition A7 as a consequence of Lemma A5 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 $U_{cc}'' \le 0$ and $U_{cs}'' \ge 0$ to obtain

$$U_c'\left(c^*(\theta),s^*(\theta),z^*(\theta);\theta\right)\sigma_c\left(s^*(\theta),z^*(\theta)\right)\ge U_c'\left(c,s,z^*(\theta);\theta\right)\sigma_c\left(s,z^*(\theta)\right).$$

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

$$\begin{aligned}&U_c'\left(c^*(\theta),s^*(\theta),z^*(\theta);\theta\right)\sigma_c\left(s^*(\theta),z^*(\theta)\right)+U_z'\left(c^*(\theta),s^*(\theta),z^*(\theta);\theta\right)\\ &\ge U_c'\left(c,s,z^*(\theta);\theta\right)\sigma_c\left(s,z^*(\theta)\right)+U_z'\left(c,s,z^*(\theta);\theta\right).\end{aligned}$$

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

$$\begin{aligned}&\frac{d}{ds}\left[\frac{U_s'(z-T_z\left(z\right)-s-T_s\left(s\right),s,z^*(\theta);\theta)}{1+T_s'(s)}\right]\\ &=-U_{cs}''+\frac{1}{(1+T_s'(s))}\left[U_{ss}''-U_s'\frac{T_{ss}''(s)}{1+T_s'(s)}\right].\end{aligned}$$

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

[64] 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.

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 independent of s, we obtain
$$\begin{aligned} &U_s'\left(c^*(\theta), s^*(\theta), z^*(\theta); \theta\right) \sigma_s\left(s^*(\theta), z^*(\theta)\right) + U_z'\left(c^*(\theta), s^*(\theta), z^*(\theta); \theta\right) \\ &\geq U_s'\left(c, s, z^*(\theta); \theta\right) \sigma_s\left(s, z^*(\theta)\right) + U_z'\left(c, s, z^*(\theta); \theta\right). \end{aligned}$$

LED systems. First, consider a type θ' choosing earnings $z = z^*(\theta) > z^*(\theta')$. We have
$$\begin{aligned} &\frac{d}{ds}\left[U_c'\left(z - s - \tau_s\left(z^*(\theta)\right)s - T_z\left(z^*(\theta)\right), s, z^*(\theta); \theta\right)\left(1 - T_z'\left(z^*(\theta)\right) - \tau_s'\left(z^*(\theta)\right)s\right)\right] \\ &= U_{cs}''\left(1 - T_z'\left(z^*(\theta)\right) - \tau_s'\left(z^*(\theta)\right)s\right) - U_{cc}''\left(1 + \tau_s(z^*(\theta))\right)\left(1 - T_z'\left(z^*(\theta)\right) - \tau_s'\left(z^*(\theta)\right)s\right) - U_c'\tau_s'\left(z^*(\theta)\right). \end{aligned}$$

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 *(2a)* of the proposition implies that the remaining terms are negative such that
$$U_c'\left(z - s - \tau_s(z^*(\theta))s - T(z^*(\theta)), s, z^*(\theta); \theta\right) \sigma_c\left(s, z^*(\theta)\right)$$
is increasing in s for $s < s^*(\theta)$, where $\sigma_c\left(s, z^*(\theta)\right) = 1 - T_z'(z) - \tau_s'(z)s$. As a result,
$$\begin{aligned} &U_c'\left(c^*(\theta), s^*(\theta), z^*(\theta); \theta\right) \sigma_c\left(s^*(\theta), z^*(\theta)\right) \\ &\geq U_c'\left(c(z^*(\theta), \theta'), s(z^*(\theta), \theta'), z^*(\theta); \theta\right) \sigma_c\left(s(z^*(\theta), \theta'), z^*(\theta)\right) \end{aligned}$$
and thus relying on the fact that $U_{cz}'' = 0$ and $U_{sz}'' = 0$, we have
$$\begin{aligned} &U_c'\left(c^*(\theta), s^*(\theta), z^*(\theta); \theta\right) \sigma_c\left(s^*(\theta), z^*(\theta)\right) + U_z'\left(c^*(\theta), s^*(\theta), z^*(\theta); \theta\right) \\ &\geq U_c'\left(c(z^*(\theta), \theta'), s(z^*(\theta), \theta'), z^*(\theta); \theta\right) \sigma_c\left(s(z^*(\theta), \theta'), z^*(\theta)\right) + U_z'\left(c(z^*(\theta), \theta'), s(z^*(\theta), \theta'), z^*(\theta); \theta\right). \end{aligned}$$

Second, consider a type θ' choosing $z = z^*(\theta) < z^*(\theta')$. We have
$$\begin{aligned} &\frac{d}{ds}\left[U_s'\left(z - s - \tau_s(z^*(\theta))s - T(z^*(\theta)), s, z^*(\theta); \theta\right) \frac{1 - T_z'\left(z^*(\theta)\right) - \tau_s'\left(z^*(\theta)\right)s}{1 + \tau_s(z)}\right] \\ &= -U_{cs}''\left(1 - T_z'\left(z^*(\theta)\right) - \tau_s'\left(z^*(\theta)\right)s\right) + U_{ss}'' \frac{1 - T_z'\left(z^*(\theta)\right) - \tau_s'\left(z^*(\theta)\right)s}{1 + \tau_s(z)} + U_s' \frac{\tau_s'\left(z^*(\theta)\right)}{1 + \tau_s(z)}. \end{aligned}$$

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 $z = z^*(\theta) < z^*(\theta')$ has a desired savings level $s(z^*(\theta), \theta') > s^*(\theta)$. Hence, condition *(2b)* 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
$$\begin{aligned} &U_s'\left(c^*(\theta), s^*(\theta), z^*(\theta); \theta\right) \sigma_c\left(s^*(\theta), z^*(\theta)\right) \\ &\geq U_s'\left(c(z^*(\theta), \theta'), s(z^*(\theta), \theta'), z^*(\theta); \theta\right) \sigma_c\left(s(z^*(\theta), \theta'), z^*(\theta)\right) \end{aligned}$$
and thus, relying on the fact that $U_{cz}'' = 0$ and $U_{sz}'' = 0$, we have
$$\begin{aligned} &U_s'\left(c^*(\theta), s^*(\theta), z^*(\theta); \theta\right) \sigma_c\left(s^*(\theta), z^*(\theta)\right) + U_z'\left(c^*(\theta), s^*(\theta), z^*(\theta); \theta\right) \\ &\geq U_s'\left(c(z^*(\theta), \theta'), s(z^*(\theta), \theta'), z^*(\theta); \theta\right) \sigma_c\left(s(z^*(\theta), \theta'), z^*(\theta)\right) + U_z'\left(c(z^*(\theta), \theta'), s(z^*(\theta), \theta'), z^*(\theta); \theta\right). \end{aligned}$$

E Details on the Empirical Application

This appendix describes the details underlying the numerical results presented in Section VII. In Section E.A, we describe how we calibrate a baseline two-period, unidimensional model of the U.S. economy, which we use to compute the simple savings tax schedules that are consistent with the prevailing income tax, i.e., that satisfy the Pareto efficiency formulas in Proposition 3. These are reported in Figure 3. In Section E.B, we describe how we extend this exercise to calibrate the optimal simple savings tax systems in the presence of multidimensional heterogeneity as in Proposition 4, assuming that redistributive preferences and other sufficient statistics are the same as in the baseline calibration. In Section E.C, we describe how we instead extend the baseline exercise to allow for heterogeneous rates of return, with an efficiency-based rationale for taxing those with access to high returns, as in Proposition 6. Results for these extensions are reported in Figure 4. Throughout this exercise, we make two assumptions for tractability: We assume that preferences are weakly separable as described in Proposition 2, so that the income effect on savings, $\eta_{s|z}(z)$, can be identified from $s'_{inc}(z)$, and we assume that income effects on labor supply are negligible ($\eta_z(z) \approx 0$) .

For comparability with the literature on wealth taxation, we express all savings tax rates in terms of "period-2" taxes on gross savings, so that a marginal savings tax rate of 0.1 indicates that if an individual's total wealth at retirement increases by \$1, then they must pay an additional \$0.10 in taxes when they retire.[65]

The LaTeX source code underlying this document—which can be viewed in the accompanying replication files—uses equation labels that match those in the Matlab simulation code.

E.A Baseline Calibration with Unidimensional Heterogeneity

We first calibrate a simplified version of the U.S. economy with unidimensional heterogeneity. This calibration has two periods, with the first period corresponding to working life and the second to retirement. To accord with the age groups reported in Piketty et al. (2018) (henceforth PSZ), which we use to calibrate our model, we interpret the working life period to correspond to people between ages 25 and 65 (average age: 45) with the retirement period beginning at age 65. We therefore assume that these periods are separated by 20 years, with savings accruing returns at a risk-free annual rate of return of 3.8% per year between period 1 and period 2 (see Fagereng et al. (2020), Table 3) so that the total (pre-tax) return on savings during working life is

$$1 + r = 1.038^{20} = 2.1. \tag{271}$$

E.A.1 Joint Distribution of Earnings and Savings, and the Status Quo Income Tax

We calibrate the joint distribution of earnings and savings using the Distributional National Accounts micro-files of PSZ from 2019. 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, and we partition the top percentile into the top 0.01%, the top 0.1% (excluding the top 0.01%), and the rest of the top percentile. We treat the age groups in each percentile as the young and middle-aged versions of a common earning type in steady state. Our measure of annualized earnings during work life z at the n-th percentile is constructed by averaging *plinc* 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. For households with one member above 65 years old, we keep only 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 that would correspond to $z = 0$ in our model.

Figure A1 displays average savings (*hweal*) among 45- to 65-year-olds within each percentile of the distribution. For married households, we take household wealth to be the average wealth of its members. This profile exhibits a convex pattern, consistent with $s'(z)$ rising across incomes. For the purposes of our two-period model, we are interested in the profile of savings at the time of retirement—age 65—rather than averaged across everyone ages 45 to

[65] Notationally, we write this translation as in Appendix B.E, with s_1 and s_g denoting gross savings before taxes, measured in period-1 and period-2 dollars, respectively, and $T_2(s_g, z)$ denoting the savings tax function in period 2. Appendix B.E demonstrates that the simplicity structure of a tax system (SL, SN, and LED) is preserved when translating between $\mathcal{T}(s, z)$ and $\{T_1(z), T_2(s_g, z)\}$. In the accompanying code replication files, all savings taxes are computed in terms of $\mathcal{T}(s, z)$, but marginal tax rates are converted into $\frac{\partial T_2(s_g,z)}{\partial s_g}$ when plotted in figures.

65. Although we lack finer information about age in the PSZ data, we can use evidence from the Federal Reserve Bank of New York's Survey of Consumer Expectations (from which we also draw estimates of $s'_{inc}(z)$ below) to understand how the wealth profile evolves. Using SCE data from 2015 to 2019, we compute average net wealth within each of the 11 reported income bins among respondents aged 45 to 65. The resulting schedule is plotted by the solid points in Figure A2; reassuringly, this profile is very similar to the one constructed from the PSZ data, although its sample size is much smaller and the measure of income is binned more coarsely. To understand how this profile evolves as respondents age, we recompute the profile using just respondents aged 55 to 65; this schedule is plotted by the hollow points in A2. The evolution suggests that savings retain a similar profile across income, but increases by about 10% from the first schedule (average age: 55) to the second (average age: 60). We therefore project that wealth would increase by a further 10% from age 60 to 65—the age of retirement in our calibrated model. Thus we construct our projected measure of total wealth at retirement by multiplying the PSZ profile by 1.2. We normalize this to the our measure of gross retirement savings per year worked, which we denote s_g in the notation of Appendix B.E, by dividing this total wealth by the number of work years, which we assume is 40, corresponding to work life lasting from age 25 to 65. This procedure pins down the cross-sectional variation in gross savings $s'_g(z)$.

We convert this discrete distribution of labor income and savings into a smooth density with 1000 gridpoints to ensure a smooth marginal tax function that converges to a fixed point when we iterate using the first-order conditions from our propositions. Because income percentile points are close together (in levels) at low incomes and very far apart at high incomes, it is useful to construct our income grid with equally log-spaced gridpoints in z. And because the observed (discretized) $h_z(z)$ is very small at some incomes, it is useful to construct our grid by smooth-fitting log-density (rather than density) to $\ln z$, to avoid producing negative smoothed values of $h_z(z)$. To do this, we define $H_{lnz}(\ln z) := H(z)$, with $h_{lnz}(\ln z) = \frac{d}{d\ln z}H_{lnz}(\ln z)$. We then smooth-fit the discretized $h_{lnz}(\ln z)$ to $\ln z$ from the PSZ data using the smoothing spline fit in Matlab, with a smoothing parameter of 0.9 and the scale normalization setting set to "on." We similarly fit log savings to log income. 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 (log) outliers 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 negative outliers. Our baseline uses $k = \$20,000$. Having obtained a smooth fit of $h_{lnz}(\ln z)$ to $\ln z$, we translate into the implied income density $h_z(z)$ using the identity

$$h(z) = \frac{d}{dz}H(z) = \frac{d}{dz}H_{lnz}(\ln z) = \frac{1}{z}\cdot\frac{d}{d\ln z}H_{lnz}(\ln z) = \frac{1}{z}\cdot h_{lnz}(\ln z). \tag{272}$$

We then rescale this resulting density so that $\int_z h(z)dz = 1$.

We construct the status quo income tax function by comparing gross income to the PSZ measure *diinc* ("extended disposable income") of post-tax income $z - T_1(z)$. We use the median value within each pre-tax income percentile to mute the effect of outliers from other income sources, and we construct a smoothed profile of disposable income y by fitting log *diinc* to log *plinc*, with the same setting described above. In the DINA files, total disposable income *diinc* exceeds total labor income *plinc*, reflecting non-labor factors of production in the economy and the taxes on them. For internal consistency, we apply a lump-sum adjustment so that total y and z are equal, although our results are not sensitive to this adjustment. We then calibrate the smooth marginal income tax rate schedule as $1 - \frac{dy}{dz}$. We treat Social Security as a fixed amount of forced savings, which are added to net-of-tax disposable savings to arrive at our total measure of net savings s. [66]

E.A.2 Status Quo Savings Tax Rates in the United States

We are interested in comparing our results to the profile of status quo effective tax rates on savings in the U.S. Constructing such a schedule presents several difficulties, however. There are many different types of taxes which apply to savings in the U.S., including capital gains taxes (which differ depending on the length of asset ownership), ordinary income taxes, and property taxes. Moreover, effective tax rates depend on assumptions about incidence, about which there is substantial disagreement.

We use a simple approach to construct an approximation of the U.S. savings tax based on the composition of

[66] The amount is computed as follows, using the SSA Fact Sheet: Retired workers receive on average $1,514 per month from social security, which is $12 \times 1,514 = \$18,168$ annually. Through the lens of our two-period model, these benefits are received over an average retirement length of 20 years, and stem from contributions paid over 40 working years. We therefore approximate this as forced savings at the time of retirement of $9000 per working year.

savings portfolios across the income distribution. Bricker et al. (2019) use the Survey of Consumer Finances to construct a decomposition of saving types by asset ownership percentile; we summarize the analogous decomposition by income percentile in Figure A3 below. We then construct a savings tax rate at each income level based on the asset-weighted average of the tax rates that apply to each asset class.

For comparison to our results, the savings tax rate of interest is the distortion between work-life consumption and savings. Therefore savings which are subject to labor income taxes but no further taxes, such as a Roth IRA, should be understood as being subject to zero savings tax. We similarly classify traditional IRAs and pension plans as being subject to zero taxes, since they are also subject only to ordinary income taxes. We therefore treat assets in the "Financial (retirement)" category as subject to zero savings tax. We assume "Financial (transaction)" assets, which include checking and savings accounts, represent liquidity needs and similarly do not count toward taxed savings. 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%. Therefore we view only the dotted-outline asset classes "Financial (market)" and "Nonfinancial (business)" as subject to savings taxes, in the form of capital gains. We do not know what share of these holdings represent gains, as opposed to the original contributions. To be conservative, we treat the entire asset classes as though they were subject to capital gains taxes at the time of retirement.

We treat this savings tax rate profile as a schedule of *average* tax rates on one's savings portfolio at each point in the income distribution. We smooth this schedule of average rates using the spline fit procedure described above, and apply that average tax rate to the calibrated level of gross savings at each point in the income distribution to reach a calibrated schedule of total savings taxes paid. We then compute the schedule of marginal rates that would give rise to that nonlinear profile of average tax rates; this schedule is plotted as the "U.S. Status quo" savings tax, e.g., in Figure 3.

E.A.3 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 net-of-tax savings at the time of retirement to changes in the amount of pre-tax earnings z. 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. (2021). Fagereng et al. (2021) estimate marginal propensities to consume (MPC) across the earnings distribution using information on lottery prizes linked with administrative data in Norway. They find that individuals' consumption peaks during the year in which the prize is won, before gradually reverting to their previous consumption level. Over a 5-year horizon, they estimate winners consume close to 90% of the tax-exempt lottery prize; see the "consumption" panels in Fagereng et al. (2021) Figure 2. This 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 2.)

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 recognizing that individuals' budget constraint is $s_1(z) = z - T_1(z) - c(z)$, we get

$$\frac{\partial s_1(z)}{\partial (z - T_1(z))} = 1 - \frac{\partial c(z)}{\partial (z - T_1(z))} = 0.1.$$

The identity $s = (s_1 - T_s(s))(1+r)$ implies that $\frac{\partial s}{\partial s_1} = \frac{1}{\frac{1}{1+r} + T'_s(s)}$, and thus that the local causal effect of *pre-tax* income z on *net* savings s satisfies

$$s'_{inc}(z) = \frac{\partial s_1(z)}{\partial(z - T_1(z))} \cdot \frac{\partial s}{\partial s_1} \cdot \frac{\partial(z - T_1(z))}{\partial z}$$
$$= 0.1 \cdot \frac{1 - T'_1(z)}{\frac{1}{1+r} + T'_s(s(z))}. \tag{273}$$

We can then use our calibrated U.S. tax schedule to obtain a profile of $s'_{inc}(z)$, under the key assumption that U.S. households have similar MPCs as Norwegian households. This profile is plotted in Figure 2.

U.S. estimates based on survey evidence. We additionally compute marginal propensities to save out of additional income in the U.S. based on two sources of survey evidence. The first is the Survey of Consumer Expectations, a monthly longitudinal survey conducted by the Federal Reserve Bank of New York since 2013.[67] In 2015, the survey's module on household spending added the following question about how a household would allocate an increase in income: "Suppose next year you were to find your household with 10% more income than you currently expect. What would you do with the extra income?" Answers include "Save or invest all of it," "Spend or donate all of it," and "Use all of it to pay down debts," as well as combinations of those three categories, in which case respondents were asked to specify the percentage they would allocate to each. We use these responses to compute the share of additional income that respondents would save rather than spend, treating debt reduction as a form of savings. We compute the MPS for each of the 15,540 respondents who answered this question from the date of its introduction until January 2020, and we plot the average within each of the eleven income bins reported in the survey in Figure A4, panel (a), splitting respondents into below- and above-median wealth within each income group. The average value is about 0.8, with a profile that is quite flat across incomes, and there is very little variation with wealth conditional on income. To construct a smooth profile of MPS across the range of incomes in our simulated economy, we linearly fit the profile of average MPS from Figure A4a on log income,

$$\widehat{MPS} = \beta_0 + \beta_1 \ln z, \tag{274}$$

which we then project to the full income distribution from the PSZ data. (Because the MPS profile is very flat, results are similar if we instead use a constant value across all incomes.)

An advantage of the Survey of Consumer Expectations, relative to the Fagereng et al. (2021) data from Norwegian lottery winners, is that respondents are based in the U.S., consistent with the setting of our policy simulations. However, like in Fagereng et al. (2021), the MPS identified by the Survey of Consumer Expectations represents a propensity to save out of a windfall increase in net-of-tax income, rather than labor earnings.[68] Moreover, the change might be interpreted as a one-time shock to income—akin to lottery winnings—rather than a persistent change. For the purposes constructing a measure of $s'_{inc}(z)$, we would ideally measure the MPS out of a *persistent* change in *labor earnings*, which could be different. To explore this possibility, we conducted a probability-based survey of the American population in the spring of 2021, which asked 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

[67] See `https://www.newyorkfed.org/microeconomics/sce` for more detail about the survey.

[68] The text of the SCE question states the percentages of an income change allocated to savings, spending, and debt reduction must sum to 100%, indicating that the income change is net-of-tax.

Answers to this question provide information about individuals' reported marginal propensity to consume (MPC) and marginal propensity to save (MPS) out of a small and *persistent* change in *earned* income. Our survey sample consisted of $1{,}703$ respondents who reported an average marginal propensity to save of 0.60 in the year of the raise.[69] We also requested information on household income in the survey, so we can observe marginal propensity to save across earnings levels. The results are plotted in Figure A4, panel (b). The results are remarkably consistent with the findings from the Survey of Consumer Expectations. Marginal propensities to save appear quite stable across income levels. The profile in panel (a) can be converted from the MPS out of net-of-tax income to an MPS out of pre-tax income by multiplying by the keep rate $1 - T'_z(z)$, reducing the MPS by around 30%, which suggests that the levels measured in the two surveys are similar as well. As with the SCE data, we linearly fit the MPS to log income, then project to the full range of incomes in our model economy.

Since both the SCE and our AmeriSpeak survey ask about consumption and spending within each year, we interpret the values reported in Figure A4 as estimates of short-run responses. Fagereng et al. (2021) show that positive income shocks are followed by consumption responses that can last up to 5 years. We use their impulse-response profile to convert these 1-year MPS estimates into a 5-year MPS, which we interpret as a total effect on savings before returns. To do so, we use the fact that they report a 1-year MPC of 0.52 and a 5-year MPC of 0.90, for a ratio of

$$\frac{1-0.90}{1-0.52} = 0.21. \tag{275}$$

We therefore compute our long run MPS by multiplying the short-run MPS estimates by this ratio. We then multiply the resulting long-run MPS in panel (a) by our schedule of calibrated marginal income tax rates $1-T'_1(z)$, as described in Appendix E.A above. (Because our survey question asked about a change in *pre-tax* income, we do not need to perform this conversion for panel (b).) Finally, we divide by

$$\frac{1}{1+r} + T'_s(s(z)) \tag{276}$$

to translate the long-run MPS out of annualized pre-tax labor income into reach our measure of $s'_{inc}(z)$. The resulting profiles are plotted in Figure 2. We use the profile produced by our AmeriSpeak surve as the baseline measure of $s'_{inc}(z)$ for our simulations, but given the similarity of the profiles, our main results are not sensative to this choice.

Comparison to Golosov et al. (2013). Golosov et al. (2013) also study preference heterogeneity, providing a useful point of comparison. In their baseline calibration, they assume individuals' preferences are Constant-Relative-Risk-Aversion

$$U(c,s,l) = \frac{\alpha(w)}{1+\alpha(w)} \ln c + \frac{1}{1+\alpha(w)} \ln s - \frac{1}{\sigma}(l)^{\sigma},$$

where l is the labor supply of an individual with hourly wage w such that earnings are given by $z = wl$. The risk-aversion parameter is set to $\gamma = 1$, the isoelastic disulity from labor effort is such that $\sigma = 3$, and the taste parameter is given by

$$\alpha(w) = 1.0526\,(w)^{-0.0036}.$$

In other words, the taste parameter 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). This means that this taste parameter is almost constant with income around an average of $\overline{\alpha} = 1.042$.

To illustrate how little across-income heterogeneity this implies, we compute the schedules of $s'(z)$ and $s'_{inc}(z)$ implied by their calibration. Individuals' savings choices follow from maximizing $U(c, s, \frac{z}{w})$ subject to the budget constraint $c \leq z - \frac{1}{(1+r)} s - \mathcal{T}(s,z)$. This implies

$$s = \frac{z - \mathcal{T}(s,z)}{1/(1+r) + \alpha\left(1/(1+r) + \mathcal{T}'_s\right)}$$

such that, neglecting the (potential) curvature of the tax function $\mathcal{T}'' \approx 0$, we can decompose the variation of savings

[69] This average is computed using the sample weights provided AmeriSpeak; the unweighted average is 0.59.

s across earnings z as

$$\underbrace{\frac{ds}{dz}}_{s'(z)} = \underbrace{\frac{1-\mathcal{T}_z'}{1/(1+r)+\alpha\left(1/(1+r)+\mathcal{T}_s'\right)+\mathcal{T}_s'}}_{s'_{inc}(z)} + \underbrace{\frac{-\left(1/(1+r)+\mathcal{T}_s'\right)}{1/(1+r)+\alpha\left(1/(1+r)+\mathcal{T}_s'\right)+\mathcal{T}_s'}\frac{d\alpha}{dz}s}_{\text{local preference heterogeneity}}.$$

To obtain an approximation of $s'(z)-s'_{inc}(z)$ in their setting, we use the fact that Golosov et al. (2013) report in their simulation results that individuals with an annual income $z = \$100,000$ have an hourly wage $w = \$40$ while those with an annual income $z = \$150,000$ have an hourly wage $w = \$62.5$. We can thus approximate $\frac{d\alpha}{dz} = \frac{\alpha(62.5)-\alpha(40)}{150,000-100,000} = \frac{1.0370-1.0387}{50,000} = -34 * 10^{-9}$. For an approximate computation, we assume for $\mathcal{T}_z'$ a linear income tax rate $\tau_z = 0.3$, , and for $\mathcal{T}_s'$ we assume a linear income tax rate $\tau_s = 0.01$ which we show below (see equation (278)) to be consistent with a linear tax of 4% on capital gains (the approximate average in Figure A3b).

This gives a constant $s'_{inc} = \frac{1-0.3}{1/2.1+1.042*(1/2.1+0.01)} = 0.71$, which is much higher than our estimate. Leveraging the fact that s'_{inc} is constant, we can also infer that at an annual income of \$125,000, the annual amount of savings available for consumption in period 2 (including compounded interest) is approximately equal to $s = s'_{inc} * \$125,000 = 0.71*125,000 = \$88,750$. Thus, $s'(z)-s'_{inc}(z) = \frac{1/2.1+0.02}{1/2.1+1.042*(1/2.1+0.02)+0.02}*(34*10^{-9})*88,750 = 0.0015$.[70]

These values for $s'(z)$ and s'_{inc} imply that in the calibration of Golosov et al. (2013), preference heterogeneity is substantially smaller than our estimate of across-income heterogeneity, as it only explains $\frac{0.0015}{0.71+0.0015} = 0.2\%$ of the variation in savings between individuals earning \$100,000 annually and those earning \$150,000.

E.A.4 Savings elasticity

For purposes of calibration, we assume that the income-conditional compensated elasticity of savings is constant across earnings, $\zeta^c_{s|z}(z) = \bar{\zeta}^c_{s|z}$. We follow Golosov et al. (2013) in drawing on the literature estimating the intertemporal elasticity of substitution (IES), and reporting results for a range of values. To motivate these values, we describe here how we can translate from the IES to a compensated elasticity $\zeta^c_{s|z}$ in the case of a representative agent.

The IES is defined as the elasticity of the growth rate of consumption with respect to the net price of consumption. We assume consumption is smoothed during retirement, so that retirement consumption is proportional to the net stock of savings s, and thus the elasticity of the growth rate of consumption (with respect to a tax change) is the same as the elasticity of the ratio of s to work-life consumption c. We consider a change in the price of retirement consumption induced by a small reform to a SL system like the one described in Table 1 with a constant linear tax rate τ_s, in which case the net-of-tax price of retirement savings is $\frac{1+r}{1+(1+r)\tau_s}$. (This can be found using the relationship $(s_1-\tau_s s)(1+r) = s$ and solving for $\frac{ds}{ds_1} = -\frac{ds}{dc}$.) We can therefore write

$$\begin{aligned} IES &= \frac{d\ln(s/c)}{d\ln(\frac{1+r}{1+(1+r)\tau_s})} \\ &= -\frac{d\ln(s/c)}{d\ln(1+(1+r)\tau_s)} \\ &= -\frac{d\ln s}{d\ln(1+(1+r)\tau_s)} + \frac{d\ln c}{d\ln(1+(1+r)\tau_s)} \\ &= -\frac{d\ln s}{d\ln(1+(1+r)\tau_s)} + \frac{dc}{d\ln(1+(1+r)\tau_s)}\frac{1}{c} \\ &= -\frac{d\ln s}{d\ln(1+(1+r)\tau_s)} + \frac{ds}{d\ln(1+(1+r)\tau_s)}\frac{dc}{ds}\frac{1}{c}. \end{aligned}$$

[70] More specifically, we here postulate $s'(z)-s'_{inc}(z) \ll s'_{inc}(z)$ to infer $s(z) = s'_{inc}\cdot z$ and then compute $s'(z)-s'_{inc}(z)$. Since we obtain a value that verifies $s'(z)-s'_{inc}(z) \ll s'_{inc}(z)$, this proves that $s'(z)-s'_{inc}(z) \ll s'_{inc}(z)$. Put differently, if we were to assume that $s'(z)-s'_{inc}(z) \approx s'_{inc}(z)$, in which case $s(z) = 2s'_{inc}\cdot z$, we would still obtain that $s'(z)-s'_{inc}(z) \ll s'_{inc}$, showing that our conclusion is not implied by our assumption (i.e., our reasoning is not circular).

Substituting for $\frac{dc}{ds} = \frac{1+(1+r)\tau_s}{1+r}$, we then obtain

$$\begin{aligned}
IES &= -\frac{d\ln s}{d\ln(1+(1+r)\tau_s)} - \frac{d\ln s}{d\ln(1+(1+r)\tau_s)} \frac{1+(1+r)\tau_s}{1+r} \frac{s}{c} \\
&= -\left(1+\left(\frac{1+(1+r)\tau_s}{1+r}\right)\frac{s}{c}\right)\frac{d\ln s}{d\ln(1+(1+r)\tau_s)} \\
&= -\left(1+\left(\frac{1+(1+r)\tau_s}{1+r}\right)\frac{s}{c}\right)\frac{d\ln(1+\tau_s)}{d\ln(1+(1+r)\tau_s)}\frac{d\ln s}{d\ln(1+\tau_s)} \\
&= -\left(1+\left(\frac{1+(1+r)\tau_s}{1+r}\right)\frac{s}{c}\right)\left(\frac{d(1+(1+r)\tau_s)}{d\tau_s}\right)^{-1}\frac{1+(1+r)\tau_s}{1+\tau_s}\frac{d\ln s}{d\ln(1+\tau_s)} \\
&= -\left(1+\left(\frac{1+(1+r)\tau_s}{1+r}\right)\frac{s}{c}\right)\frac{1+(1+r)\tau_s}{(1+r)(1+\tau_s)}\frac{d\ln s}{d\ln(1+\tau_s)} \\
\implies \quad \frac{d\ln s}{d\ln(1+\tau_s)} &= -\frac{IES}{\left(1+\left(\frac{1+(1+r)\tau_s}{1+r}\right)\frac{s}{c}\right)\frac{1+(1+r)\tau_s}{(1+r)(1+\tau_s)}}. \qquad (277)
\end{aligned}$$

Using a value of $s/c = 0.54$ (the population average in our calibrated two-period economy), our assumed rate of return $R = 2.1$, and $\tau_s = 0.01$ (corresponding to a linear tax of 4% on capital gains, the approximate average in Figure A3b), we find[71]

$$\frac{d\ln s}{d\ln(1+\tau_s)} = -\frac{IES}{0.61}.$$

Treating this as the population estimate of $\frac{d\ln\bar{s}}{d\ln(1+\tau_s)}$, we can then compute the value of the elasticity $\bar{\zeta}^c_{s|z}$ that is consistent with this estimate. From the proof of the optimal SL tax system (see Appendix C.F.1, equation (123)), the response of aggregate savings $\bar{s}$ to a change in the separable linear tax rate τ_s (measured in period-1 dollars, as distinct from $\tau_{s,2}$) is:

$$\begin{aligned}
\frac{d\bar{s}}{d\tau_s} &= -\int_z \left\{ \frac{1}{1+\tau_s}\left(s(z)\bar{\zeta}^c_{s|z} + \eta_{s|z}(z)s(z)\right) + \frac{s'_{inc}(z)}{1-T'_z(z)}\left(z\,\zeta^c_z(z)s'_{inc}(z) + \eta_z(z)s(z)\right)\right\} dH_z(z) \\
\frac{d\bar{s}}{d\tau_s}\frac{1+\tau_s}{1} &= -\bar{\zeta}^c_{s|z}\bar{s} - \int_z\left\{\eta_{s|z}(z)s(z) + s'_{inc}(z)\frac{1+\tau_s}{1-T'_z(z)}\left(z\,\zeta^c_z(z)s'_{inc}(z) + \eta_z(z)s(z)\right)\right\} dH_z(z) \\
\underbrace{\frac{d\bar{s}}{d\tau_s}\frac{1+\tau_s}{\bar{s}}}_{\frac{d\ln\bar{s}}{d\ln(1+\tau_s)}} &= -\bar{\zeta}^c_{s|z} - \int_z\left\{\eta_{s|z}(z)\frac{s(z)}{\bar{s}} + \frac{s'_{inc}(z)}{\bar{s}}\frac{1+\tau_s}{1-T'_z(z)}\left(z\,\zeta^c_z(z)s'_{inc}(z) + s(z)\,\eta_z(z)\right)\right\} dH_z(z) \\
\bar{\zeta}^c_{s|z} &= -\frac{d\ln\bar{s}}{d\ln(1+\tau_s)} - \mathbb{E}\left[\eta_{s|z}(z)\frac{s(z)}{\bar{s}} + \frac{s'_{inc}(z)}{\bar{s}}\frac{1+\tau_s}{1-T'_z(z)}\left(z\,\zeta^c_z(z)s'_{inc}(z) + \eta_z(z)s(z)\right)\right]
\end{aligned}$$

This could be computed directly if we had an independent estimate of the income-conditional income effect $\eta_{s|z}$. We instead invoke our assumptions of weak separability and a separable tax system, implying $\eta_{s|z}(z) = s'_{inc}(z)\frac{1+T'_s(s(z))}{1-T'_z(z)}$

[71] A linear tax rate τ^{cg} on capital gains rs_1 leads to net savings $s = s_1(1+r(1-\tau^{cg}))$. Similarly, a period-1 linear tax τ_s on net savings s leads to net savings $s = (s_1 - \tau_s s)(1+r) \iff s = \frac{s_1(1+r)}{1+\tau_s(1+r)}$. As a result,

$$\begin{aligned}
s_1(1+r(1-\tau^{cg})) &= \frac{s_1(1+r)}{1+\tau_s(1+r)} \\
\iff 1+\tau_s(1+r) &= \frac{(1+r)}{1+r(1-\tau^{cg})} \\
\iff \tau_s &= \frac{1}{1+r(1-\tau^{cg})} - \frac{1}{1+r}. \qquad (278)
\end{aligned}$$

(see Proposition 2), and negligible income effects on earnings, to write

$$\begin{aligned}\bar{\zeta}^c_{s|z} &= -\frac{d\ln\bar{s}}{d\ln(1+\tau_s)} - \mathbb{E}\left[\frac{1+T'_s(z)}{1-T'_z(z)}\frac{s'_{inc}(z)}{\bar{s}}\left(s(z)+z\bar{\zeta}^c_z s'_{inc}(z)\right)\right] \\ &= -\frac{d\ln\bar{s}}{d\ln(1+\tau_s)} - \frac{1}{\bar{s}}\cdot\mathbb{E}\left[\frac{1+T'_s(z)}{1-T'_z(z)}s'_{inc}(z)\left(s(z)+z\bar{\zeta}^c_z s'_{inc}(z)\right)\right]. \qquad (279)\end{aligned}$$

In our calibration, the value of the second term is 0.37, suggesting a translation of $\bar{\zeta}^c_{s|z} \approx IES/0.61 - 0.37$. Thus a value of $IES = 1$, the baseline in Golosov et al. (2013), suggests an elasticity of $\bar{\zeta}^c_{s|z} = 1.3$. We use a baseline value of $\bar{\zeta}^c_{s|z} = 1$. IES values of 0.5 and 2 (the "low" and "high" values considered in Golosov et al. (2013)) suggest savings elasticities of $\bar{\zeta}^c_{s|z} = 0.4$ and $\bar{\zeta}^c_{s|z} = 2.9$. This is a wide range; values of savings elasticities below $\bar{\zeta}^c_{s|z} = 0.6$ in particular suggest that consistency with the status quo income tax requires a savings tax that is extreme or non-convergent.[72] We report results for alternative values of $\bar{\zeta}^c_{s|z} = 0.7$, $\bar{\zeta}^c_{s|z} = 2$, and $\bar{\zeta}^c_{s|z} = 3$.

E.A.5 Computing Pareto-efficient tax rates

Having calibrated the status quo savings and income distribution and parameters, we then solve for the schedule of Pareto-efficient savings tax rates reported in Figure 3 using a fixed-point iteration method. That is, we begin with a candidate suboptimal tax schedule (whether SL, SN, or LED) which we use to compute the right-hand sides of the desired Pareto efficiency equations (31), (32), and (33). We use the result to compute an updated tax schedule, which we then plug into the right-hand sides of those equations again, and we iterate until we reach convergence. We use an analogous procedure for the other solutions to follow.

Although the schedule of optimal marginal savings tax rates is identical under SN and LED for a given set of welfare weights (see equations (29) and (30) of Proposition 3), the schedule of Pareto-efficient tax rates for a given income tax $T_z(z)$ is slightly different, because the earnings-dependence of the LED savings tax itself contributes to the labor distortion wedge (see equations (32) vs. (33) of that Proposition).

E.B Simulations of Optimal Savings Taxes with Multidimensional sHeterogeneity

We now extend the above calibration to accommodate multidimensional heterogeneity, which we use to apply the formulas derived in Proposition 4. In the multidimensional setting, we do not have Pareto efficiency formulas like those for unidimensional setting, because in the presence of income-conditional savings heterogeneity, Pareto-improving reforms are not generally available. Therefore, we use the formulas in Proposition 4 to compute the *optimal* schedule of savings tax rates for each type of simple tax system. In order to isolate and illustrate the effects of multidimensional heterogeneity, we hold fixed the sufficient statistics used in the unidimensional setting. We also hold fixed the distributional preferences of the policy maker. The Pareto-efficiency computations above are equivalent to computing the optimal tax under "inverse optimum" welfare weights that would rationalize the status quo income tax. We compute these welfare weights explicitly, as described below, assuming that they vary with earnings, but not with savings conditional on earnings. We then use those inverse optimum weights for the optimal tax calculations.

E.B.1 Calibration Details and Assumptions

To extend our calibrated two-period model economy to a multidimensional setting, we retain the same discretized grid of incomes as in the unidimensional case, using the calibration described in Appendix E.A. At each income, we now allow for heterogeneous levels of savings, and we assume that within each earnings-savings cell, agents are homogeneous. Specifically, using the same measure of gross savings described in Appendix E.A, we now use a calibration with four different levels of savings at each level of income, each representing a quartile of the income-conditional savings distribution. Across the income distribution, we assume savings within each quartile are a constant

[72] Intuitively, as the savings elasticity becomes low, one's level of savings becomes a reliable signal of underlying ability, and more of the total redistribution in the tax system should be carried out through the savings tax, rather than the income tax. Thus for sufficiently low $\bar{\zeta}^c_{s|z}$, the status quo income tax cannot be Pareto efficient.

ratio of the income-conditional average level of saving.[73] These ratios are 15%, 39%, 69% and 280% of the income-conditional average savings level; they are calibrated to reflect the average ratios across percentiles 50 to 100 in the PSZ data. We calibrate these ratios excluding the bottom portion of the distribution because the average level of saving is very low in the bottom half, resulting in noisily measured ratios. To calibrate $s'_{inc}(s,z)$, which may vary with both income and (conditional on income) with savings, we use the data from SCE, described in Appendix E.A.3. Figure A4a shows that the marginal propensity to save is very similar among below- and above-median savers within each income bin; we therefore assume that s'_{inc} is constant conditional on s, with the same profile across incomes as in the unidimensional case, i.e., $s'_{inc}(s,z) = s'_{inc}(z)$. Also as in the unidimensional case, we assume that utility from savings and consumption is weakly separable from labor, so that $\eta_{s|z}(s,z) = s'_{inc}(z)\frac{1+\mathcal{T}'_s}{1-\mathcal{T}'_z}$.

E.B.2 Computing Welfare Weights via the Inverse Optimum Approach

As noted above, we compute the welfare weights that rationalize the status quo tax system, which we assume is approximately an SN system, and we assume that these welfare weights vary with earnings, but not with savings conditional on earnings.[74] As a result, we can write welfare weights $g(\theta)$ as a function of z only: $g(z)$. To ensure that we hold fixed distributional motives between these multidimensional results and the unidimensional results shown in Figure 3, we compute these inverse optimum weights under unidimensional heterogeneity, and we then impose the same weights when computing the optimal taxes with multidimensional heterogeneity. Employing these assumptions, and maintaining our assumption from the unidimensional case that income effects are negligible ($\eta_z \approx 0$), we can write our expression for income-augmented welfare weights from equation (17) as

$$\hat{g}(z) = g(z) + \left(\frac{\mathcal{T}'_s}{1+\mathcal{T}'_s}\right)\eta_{s|z}(z). \tag{280}$$

The inverse optimum computes the social marginal welfare weights (SMWW) consistent with existing tax policy (Bourguignon and Spadaro, 2012; Lockwood and Weinzierl, 2016). This exercise is typically performed using labor income taxes. Our setting presents a complication, as we have both a status quo income tax and savings tax, which need not produce a consistent set of weights. We compute weights assuming that the status quo schedule of earnings tax rates is optimal, for consistency with the Pareto efficiency formulas above. Since the status quo savings tax rates also appear in this calculation, we must choose whether to use the status quo rates, or the rates that would counterfactually be optimal. In practice, results are insensitive to this latter issue; for consistency with the "inverse optimum" motivation, we use the Pareto-efficient set of SN tax rates.

Under these assumptions, we can compute the inverse optimum social marginal welfare weights at each earnings z by inverting the optimal tax rate condition,

$$\frac{T'_z(z)}{1-T'_z(z)} = \frac{1}{\zeta^c_z(z)\,z}\frac{1}{h_z(z)}\int_{x\geq z}(1-\hat{g}(x))\,dH_z(x) - s'_{inc}(z)\frac{T'_s(s(z))}{1-T'_z(z)} \tag{281}$$

$$\iff \int_{x\geq z}(1-\hat{g}(x))\,dH_z(x) = \zeta^c_z(z)\,z\,h_z(z)\frac{T'_z(z)+s'_{inc}(z)T'_s(s(z))}{1-T'_z(z)}, \tag{282}$$

where the right-hand side term can be identified from the data. Differentiating with respect to z yields the expression we use to implement this computation numerically,

$$\hat{g}(z) = 1 + \frac{1}{h_z(z)}\cdot\frac{d}{dz}\left[\zeta^c_z(z)\,z\,h_z(z)\frac{T'_z(z)+s'_{inc}(z)T'_s(s(z))}{1-T'_z(z)}\right]. \tag{283}$$

Using the fact that augmented social marginal welfare weights are defined as

$$\hat{g}(z) := g(z) + T'_z(z)\frac{\eta_z(z)}{1-T'_z(z)} + T'_s(s(z))\left(\frac{\eta_{s|z}(z)}{1+T'_s(s(z))} + s'_{inc}(z)\frac{\eta_z(z)}{1-T'_z(z)}\right), \tag{284}$$

[73] Our results are robust to using a finer distribution of income-conditional savings, e.g., by doubling the number of income-conditional savings bins.

[74] Like Assumption 1 in Saez (2002), this amounts to assuming that government preferences for redistribution are neutral with respect to consumption pattterns over s.

and assuming preferences are weakly separable, such that by Proposition 2 we have $s'_{inc}(z) = \frac{1-T'_z(z)}{1+T'_s(s(z))}\eta_{s|z}(z)$, inverse optimum weights $g(z)$ are obtained from $\hat{g}(z)$ as follows:

$$g(z) = \hat{g}(z) - s'_{inc}(z)\left(\frac{T'_s(s(z))}{1-T'_z(z)}\right). \tag{285}$$

Figure A5 plots our estimated profile of inverse optimum weights. These weights are computed for the prevailing income tax under an SN policy, and thus for which the Pareto-efficient SN policy is optimal. These are approximately (but not exactly) the same as the weights for which the Pareto-efficient LED policy is optimal, which creates a different labor supply distortion. For consistency, we use SN-inverse-optimum weights for both sets of simulations.

E.B.3 Separable linear (SL) tax system

The optimal SL tax formula with multidimensional heterogeneity (Proposition 4) is

$$\begin{aligned}\frac{\tau_s}{1+\tau_s} =& \frac{1}{\bar{s}\zeta^c_{s|z}}\left\{\int_z\left[\left(\bar{s}'(z) - \overline{s'_{inc}}(z)\right)\int_{y\geq z}\mathbb{E}\left[1-\hat{g}\middle|y\right]dH_z(y)\right]dz\right. \\ & \left. -\int_z\left[Cov\left[\hat{g},s\middle|z\right] + Cov\left[FE_z, s'_{inc}\middle|z\right]\right]dH_z(z)\right\},\end{aligned} \tag{286}$$

where the first line is a direct extension of the optimal SL tax formula with unidimensional heterogeneity (using average sufficient statistics) and the second line represents new terms stemming from multidimensionality.

Under the aforementioned assumptions and negligible η_z, we expand $\hat{g}(s,z)$ to write

$$Cov\left[\hat{g},s\middle|z\right] = Cov\left[g(z) + \frac{\tau_s}{1-T'_z(z)}s'_{inc}(s,z), s\middle|z\right] = \frac{\tau_s}{1-T'_z(z)}Cov\left[s'_{inc}, s|z\right] \tag{287}$$

$$Cov\left[FE_z, s'_{inc}|z\right] = Cov\left[\frac{T'_z(z) + s'_{inc}(s,z)\tau_s}{1-T'_z(z)}z\zeta^c_z(s,z), s'_{inc}(s,z)\middle|z\right] = \frac{\tau_s}{1-T'_z(z)}z\zeta^c_z\mathbb{V}\left[s'_{inc}|z\right] \tag{288}$$

Under the assumption that $s'_{inc}(s,z)$ is constant conditional on income, both of these terms are zero and the optimal SL tax is simply

$$\frac{\tau_s}{1+\tau_s} = \frac{1}{\bar{s}\zeta^c_{s|z}}\left\{\int_z\left[\left(\bar{s}'(z) - \overline{s'_{inc}(z)}\right)\int_{y\geq z}\left(1 - g(y) - \frac{\tau_s}{1-T'_z(y)}s'_{inc}(y)\right)dH_z(y)\right]dz \right. \tag{289}$$

Thus if $s'_{inc}(s,z)$ is constant conditional on income, as in our calibration, the SL optimal tax is identical to the optimal tax in a unidimensional system with the same income-conditional average wealth.[75]

To compute the optimum, we iterate over the above formula until we find a fixed point value for the tax. Because allocations change as we iterate, we require an assumption about how welfare weights adjust in response. We assume that they remain proportional to the weights computed under our inverse optimum procedure, but rescaled in order to ensure that the income-effect-augmented weights $\hat{g}$ integrate to one, as must be the case at the optimum.[76]

[75] For numerical efficiency, in our simulation code we use the following rearrangement of the upper integral over $\hat{g}$:

$$\int_{y\geq z}(1-\hat{g}(y)dH_z(y) = \int_0^z \hat{g}(y)dH_z(y) - H_z(z). \tag{290}$$

[76] Specifically, letting $g^0(z)$ denote our baseline welfare weights, we set $g(z) = \kappa g^0(z)$, so in the SL case,

$$\kappa = \frac{1 - \int_z \frac{\tau_s}{1-T'_z(z)}s'_{inc}(z)\,dH_z(z)}{\int_z g^0(z)dH_z(z)}, \tag{291}$$

E.B.4 Separable nonlinear (SN) tax system

At any given savings level s^0, the optimal SN tax with multidimensional heterogeneity (Proposition 4) satisfies

$$\frac{T_s'(s^0)}{1+T_s'(s^0)} = \frac{1}{\mathbb{E}[s\zeta_{s|z}^c|s^0]}\left\{\frac{1}{h_s(s^0)}\int_{s\geq s^0}\mathbb{E}[1-\hat{g}|s]dH_s(s) - \mathbb{E}[FE_z\, s_{inc}'|s^0]\right\}.$$

Under the aforementioned assumptions, expanding $\hat{g}$ and FE_z gives

$$\begin{aligned}\frac{T_s'(s^0)}{1+T_s'(s^0)} =& \frac{1}{s^0\zeta_{s|z}^c}\left\{\frac{1}{h_s(s^0)}\int_{s\geq s^0}\mathbb{E}\left[1-g(z)-\frac{T_s'(s)}{1-T_z'(z)}s_{inc}'(z)\middle|s\right]dH_s(s)\right. \\ &\left. - \mathbb{E}\left[\frac{T_z'(z)+s_{inc}'(z)T_s'(s)}{1-T_z'(z)}z\zeta_z^c s_{inc}'(z)\middle|s=s^0\right]\right\}\end{aligned} \tag{294}$$

or equivalently, expressing this as a function of the savings density $h_s(s) = \int_z h(s,z)dz$,

$$\begin{aligned}\frac{T_s'(s^0)}{1+T_s'(s^0)} =& \frac{1}{s^0\zeta_{s|z}^c h_s(s^0)}\left\{\int_{s\geq s^0}\mathbb{E}\left[1-g(z)-\frac{T_s'(s)}{1-T_z'(z)}s_{inc}'(z)\middle|s\right]h_s(s)ds\right. \\ &\left. - h_s(s^0)\mathbb{E}\left[\frac{T_z'(z)+s_{inc}'(z)T_s'(s)}{1-T_z'(z)}z\zeta_z^c s_{inc}'(z)\middle|s=s^0\right]\right\}\end{aligned} \tag{295}$$

where the expectations operator denotes integration with respect to earnings conditional on savings.

For numerical 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 of 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 statistic $x(s,z)$ is $\mathbb{E}[x(s,z)|s] = \frac{\sum_{m=1}^M x(s_m,z)h_{s,m}(s)}{h_s(s)}$.

E.B.5 Linear earnings dependent (LED) tax system

The optimal LED tax formula with multidimensional heterogeneity (Proposition 4) is

$$\begin{aligned}LED:\ \frac{\tau_s(z^0)}{1+\tau_s(z^0)} =& \frac{1}{\mathbb{E}[s\zeta_{s|z}^c|z^0]}\left\{\left(\bar{s}'(z^0)-\overline{s_{inc}'}(z^0)\right)\frac{1}{h_z(z^0)}\int_{z\geq z^0}\mathbb{E}\left[1-\hat{g}\middle|z\right]dH_z(z)\right. \\ &\left. - Cov\left[\hat{g},s\middle|z^0\right] - Cov\left[FE_z, s_{inc}'\middle|z^0\right] + \frac{1}{h_z(z^0)}\frac{d}{dz^0}\left(Cov\left[FE_z,s\middle|z^0\right]\ h_z(z^0)\right)\right\}.\end{aligned} \tag{296}$$

where the first line is a direct extension of the optimal LED tax formula with unidimensional heterogeneity (using average sufficient statistics) and the second line contains new terms stemming from multidimensional heterogeneity.

In the LED case, noting $\mathcal{T}_s' = \tau_s(z)$ and $\mathcal{T}_z' = T_z'(z) + \tau_s'(z)s$, we have

$$\kappa = \frac{1-\int_z \mathbb{E}\left[\frac{\tau_s(z)}{1-T_z'(z)-\tau_s'(z)s}\middle|z\right]s_{inc}'(z)\,dH_z(z)}{\int_z g^0(z)dH_z(z)}, \tag{292}$$

and with $\mathcal{T}_s' = T_s'(s)$ and $\mathcal{T}_z' = T_z'(z)$ in the SN case we have

$$\kappa = \frac{1-\int_s T_s'(s)\int_z \frac{s_{inc}'(z)}{1-T_z'(z)}h(s,z)dzds}{\int_z g^0(z)dH_z(z)}. \tag{293}$$

Under the aforementioned assumptions, expanding $\hat{g}$ and FE_z, we have

$$Cov\left[\hat{g}, s\middle|z\right] = Cov\left[g(z) + \frac{\tau_s(z)}{1 - T_z'(z) - \tau_s'(z)s} s_{inc}', s\middle|z\right] = \tau_s(z) s_{inc}'(z) Cov\left[\frac{1}{1 - T_z'(z) - \tau_s'(z)s}, s|z\right] \quad (297)$$

$$Cov\left[FE_z, s_{inc}'\middle|z\right] = Cov\left[\frac{T_z'(z) + \tau_s'(z)s + s_{inc}'\tau_s(z)}{1 - T_z'(z) - \tau_s'(z)s} z\zeta_z^c, s_{inc}'\middle|z\right] = 0 \quad (298)$$

$$\frac{d}{dz}\left(Cov\left[FE_z, s\middle|z\right] h_z(z)\right) = \frac{d}{dz}\left(z\zeta_z^c \tau_s'(z) Cov\left[\frac{s}{1 - T_z'(z) - \tau_s'(z)s}, s|z\right] h_z(z)\right) \quad (299)$$

which gives

$$\begin{aligned} LED: \ \frac{\tau_s(z^0)}{1+\tau_s(z^0)} = & \frac{1}{\bar{s}(z^0)\zeta_{s|z}^c} \Bigg\{ \left(\bar{s}'(z^0) - s_{inc}'(z^0)\right) \frac{1}{h_z(z^0)} \int_{z \geq z^0} \mathbb{E}\left[1 - \hat{g}\middle|z\right] dH_z(z) \\ & - \tau_s(z) s_{inc}'(z) Cov\left[\frac{1}{1 - T_z'(z) - \tau_s'(z)s}, s|z\right] \\ & + \frac{1}{h_z(z^0)} \frac{d}{dz^0}\left(z^0 \zeta_z^c \tau_s'(z^0) Cov\left[\frac{s}{1 - T_z'(z^0) - \tau_s'(z^0)s}, s\middle|z^0\right] h_z(z)\right) \Bigg\}, \end{aligned} \quad (300)$$

where the second and third line drop out in the case of unidimensional heterogeneity, as both covariances are then zero conditional on income.

E.C Simulations of Optimal Savings Taxes with Heterogeneous Returns

For the extension to the case with efficiency arbitrage effects, considered in Section VI.B, we now compute the optimal savings tax rates using the formulas derived in Proposition 6, again using the same set of inverse optimum welfare weights derived above.

These results are reported in the bottom two panels of Figure 4, which display schedules of LED and SN savings tax rates computed under the assumption that (i) individuals with different income levels differ in their private rates of return, and that (ii) the savings tax is levied in period-2 dollars. We compute the tax schedules that satisfy the equations for the optimal tax conditions in Proposition 6. As in the case of multidimensional heterogeneity, we hold fixed the schedule of marginal social welfare weights $g(z)$ proportional to those which rationalize the status quo income tax in our baseline inverse optimum calculation. Building on the findings of Fagereng et al. (2020), we follow Gerritsen et al. (2020) in assuming that rates of return rise by 1.4% from the bottom to the top of the income distribution. We linearly interpolate this difference across income percentiles, centered on our 3.8% baseline rate of return.

Maintaining our assumptions of negligible labor supply income effects and weakly separable preferences, equation (201) simplifies to

$$\hat{g}(x) := g(x) + \frac{\lambda_2}{\lambda_1} \frac{T_2'(s)}{1 + pT_2'(s)} \eta_{s|z}(z) \quad (301)$$

for an SN system. To ensure that $\hat{g}(z)$ still integrates to one, the rescaling factor in equation (293) now becomes

$$\kappa = \frac{1 - \int_z \left(\frac{\lambda_2}{\lambda_1} \frac{T_2'(s)}{1+pT_2'(s)}\right) \eta_{s|z}(z) dH_z(z)}{\int_z g^0(z) dH_z(z)}. \quad (302)$$

Similarly, equation (202) simplifies to

$$\varphi(x) = -\left(T_1'(x) + s_{inc}'(x) \frac{\lambda_2}{\lambda_1} T_2'(s)\right) \left(\zeta_z^c(x) \frac{x}{1 - T_1'(x)}\right) \frac{\partial p}{\partial z}. \quad (303)$$

For an LED system we can replace $T_2'(s)$ with $\tau_s(z)$ in the previous formulas.

F Details of Tax Systems by Country

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 investments, 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 of simple tax systems we consider (SL, SN, or LED) 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 A1, 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 A1. 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 wealth tax is assessed on the net value of certain classes of assets and liabilities. In eight cantons, the tax is flat, with an exemption varying by canton; the other 18 cantons feature a progressive wealth tax schedule (Scheuer and Slemrod, 2021).
- **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.