Optimal nonlinear taxation: a simpler approach

Aart Gerritsen[1]

Abstract

The typical method of solving for the optimal nonlinear tax schedule relies on deriving optimal incentive-compatible allocations. While this "primal approach" is mathematically rigorous, it lacks intuitive appeal. This paper considers a different method that relies on directly solving for the optimal tax system. This "dual approach" is much closer to actual tax policy as it centers around the welfare effects of tax reforms. I show that it can easily incorporate preference heterogeneity, as well as individual utility misoptimization. Beyond solving for the optimal tax system, the dual approach allows one to obtain insights into the welfare effects of small nonlinear tax reforms outside the optimum.

Keywords Optimal taxation · Dual approach · Preference heterogeneity · Individual misoptimization · Tax reforms

JEL classification H21 · H23 · H24

1 Introduction

Generations of economists have struggled with the question of the optimal degree of tax progressivity. In its modern form, this question was first posed by Vickrey (1945), who stated that a full characterization of the optimum "produces a completely unwieldy expression," leading him to the conclusion that "the problem resists any facile solution." Indeed, it took another quarter of a century until Mirrlees (1971, 1976) offered a first solution to the problem. The solution was obtained by way of an indirect approach: he first solved for the optimal *allocation*, subject to resource and incentive compatibility constraints, and only then determined the tax system that would implement this allocation. Ever since, this has been the dominant

✉ Aart Gerritsen
  agerritsen@ese.eur.nl
  https://sites.google.com/site/aartgerritsen/

1   Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands

approach in the literature whenever it concerns nonlinear taxation (e.g., Stiglitz, 1982; Tuomala, 1990; Diamond, 1998).

The advantage of this indirect approach is its mathematical rigor. The problem of finding the optimal allocation conveniently lends itself to the toolbox of optimal control theory, yielding a mathematically well-defined procedure for solving it. But this solution procedure also harbors the main disadvantage of this indirect approach, namely the lack of intuition involved with the derivation of the optimal tax schedule. In reality, government does not exercise any direct control over individuals' allocations—how much they work and consume of every good in the economy. Instead, it controls the tax system. Interpreting the problem of optimal taxation as choosing the most preferred incentive-compatible allocation may well-alienate the applied world of tax policy, as well as students, from the academic discipline of tax design. In the worst case, it could lead policy makers to disregard academic insights, and academics to focus too much on technical issues that might be of limited practical relevance. In short, it could reduce the practical impact of an academic field whose *raison d'être* is its potential for practical impact.[1]

A more intuitive way of solving for optimal taxes is by directly considering the social welfare effects of changes in *taxes* rather than allocations. For optimal linear taxes, this has always been the dominant solution procedure (e.g., Diamond and Mirrlees, 1971; Sheshinski, 1972; Diamond, 1975; Dixit and Sandmo, 1977). The likely reason for this is that a linear tax can be captured by a single parameter, which allows for straightforward optimization techniques. The same techniques cannot directly be applied to solve for the optimal nonlinear tax schedule, as the object to be optimized is a function rather than a parameter. Some recent contributions have circumvented this problem heuristically (e.g., Saez, 2001; 2002; Piketty and Saez, 2013; Jacquet et al., 2013). They consider a small perturbation of the tax schedule and heuristically—i.e., verbally—deduce the social-welfare effects of this perturbation. Equating these social-welfare effects to zero solves for the optimum. To prove that their heuristic is valid, they subsequently show that their results correspond to results obtained by solving for the optimal incentive-compatible allocation. This last step is necessary as it may be unclear whether the heuristic derivation picked up on all the relevant welfare effects.

In what follows, I use the term "*primal approach*" to refer to the indirect approach of first solving for the optimal *allocation*. I use the term "*dual approach*" to describe the method of directly solving for optimal *taxes*.[2] I show how one can apply the dual approach to determine the optimal nonlinear income tax without relying on a verbal derivation of social-welfare effects. By doing so, I combine the intuitive appeal of the dual approach with the mathematical rigor of the primal approach. All that is

---

[1] It might not be a coincidence that the study by Saez (2001), in which he eschews the indirect approach of first solving for the optimal allocation, not only yielded a new relevant application of optimal tax theory (i.e., the optimal top tax rate), but also seems to have ushered in an era of renewed practical relevance of optimal tax theory.

[2] The terminology is typically used in macroeconomics to refer to optimization problems in which the control variables are taken to be the allocation (primal approach) or prices (dual approach). See, for example, Chari and Kehoe (1999).

needed is a minor adjustment to the definition of the tax schedule, which makes it amenable to simple optimization techniques.

The key to this adjustment is to recognize that a person's tax burden can change for two different reasons: due to a change in his taxable income and due to a reform of the tax schedule. Thus, instead of defining a nonlinear tax as $T(z)$, with z a person's taxable income, I define it as $T(z, \kappa) \equiv \mathcal{T}(z) + \kappa \tau(z)$. Here, κ is an arbitrary parameter and $\tau(z)$ is the schedule of any nonlinear tax reform one might want to consider. Writing social welfare in terms of $T(z, \kappa)$, one can deduce the marginal welfare effects of a reform by simply taking the derivative with respect to the parameter κ, and substituting for the specific reform of interest $\tau(z)$. Expressions for the optimal nonlinear tax schedule are derived by optimizing over κ for any possible function $\tau(z)$. In other words, at the optimum, social welfare is unaffected by any possible nonlinear reform of the tax schedule.

Beyond its intuitive appeal, a second advantage of the dual approach is that it allows for a large degree of flexibility regarding individual behavior. More specifically, I show that it is straightforward to account for heterogeneity not just in individuals' income, but also in their responsiveness to tax reforms. Doing so, I replicate findings by Jacquet and Lehmann (2021) who apply the primal approach to show that standard optimal tax formulas are adjusted by using income-conditional average elasticities. Moreover, the dual approach can easily incorporate individual behavior that is not based on utility maximization. Utility maximization might not be an appropriate behavioral framework when individuals form mistaken beliefs about the shape of their budget curve or about the functional form of their own utility function. In that case, optimal tax formulas include a corrective term, prescribing higher marginal taxes for individuals who work "too much" and lower marginal taxes for individuals who work "too little."[3] The importance of such corrective term crucially depends on misoptimizers' responsiveness to tax reforms.

Finally, I show how the dual approach can be applied to determine the welfare effects of tax reforms outside the tax optimum. Contrary to the primal approach, which deals with variations in allocations rather than tax schedules, the dual approach is ideally suited to study small nonlinear reforms of a given tax schedule. This is likely to be of more relevance to actual tax policy than a characterization of the optimum. Moreover, determining the desirability of a reform may be empirically less demanding than determining the optimal tax schedule. The reason for this is that the former depends in part on the responsiveness of taxable income at the actual tax system, whereas the latter depends on the responsiveness at the optimal tax system. While we typically cannot be certain about either of the two, it is arguably less problematic to use available elasticity estimates as measures of the responsiveness of taxable income at the actual tax system than as measures of the responsiveness in the optimum.

---

[3] A similar idea is put forward by Seade (1980), Blomquist and Micheletto (2006), and Kanbur et al. (2006) on the basis of the primal approach and one-dimensional heterogeneity, and within a context of a non-welfarist social planner; also see Gerritsen (2016), Farhi and Gabaix (2020), and Lockwood (2020).

The contribution of this paper is mostly methodological and pedagogical in nature. The optimal-tax results are themselves not novel. However, they are typically derived in ways that are either mathematically daunting or verbal and therefore mathematically imprecise. The aim of this paper is to show the reader how known results on optimal taxation can be derived in a fairly simple but precise way. The hope is that this will contribute to a deeper understanding of these results among a broader audience.

Beyond the above-mentioned references, this paper relates to a number of earlier studies. To the best of my knowledge, Christiansen (1981, 1984) was the first to parameterize the nonlinear tax schedule to make it amenable to the analysis of tax reforms. His focus is on the evaluation of public projects and commodity taxation, however, and he does not consider a full characterization of the optimal nonlinear income tax—which is the focus of this study. More recently, Golosov et al. (2014) formalize the dual approach to optimal nonlinear income taxation in a dynamic model by applying Gateaux differentials with respect to the tax schedule; Hendren (2020) uses the dual approach to derive implicit welfare weights; and Spiritus et al. (2022) employ the dual approach to derive optimal taxes when households earn multiple incomes and differ across multiple dimensions. Finally, this paper also relates to earlier contributions that identify desirable tax reforms within any given non-optimal tax system (e.g., Tirole and Guesnerie, 1981; Weymark, 1981; Guesnerie, 1995; Bierbrauer et al., 2022).

Section 2 introduces the parameterization of the tax schedule, and Sect. 3 shows how this helps in deriving the welfare effects of any nonlinear tax reform. Section 4 derives expressions for optimal tax rates using the dual approach, allowing for preference heterogeneity and individuals who do not maximize their utility. Section 5 illustrates how the dual approach can be usefully applied to obtain insights into more limited tax reforms outside the optimum. Section 6 discusses the broader applicability of the dual approach and I wrap up with some concluding remarks.

2 Taxes, revenue, and social welfare

2.1 Parameterization of the tax schedule

I assume that individuals in the economy constitute a continuum $\mathcal{I}$ of unit mass, and that an individual $i \in \mathcal{I}$ earns taxable income z^i. I furthermore assume that $\{z^i : i \in \mathcal{I}\}$ is a closed interval so that it is integrable over the population $\mathcal{I}$, and denote the cumulative distribution function of taxable income by $H(z)$ and its density by $h(z)$. A person's income tax is denoted by T^i and depends on his taxable income. As such, the tax can be affected by both a change in income and a reform of the tax schedule. I formalize this by writing the income tax as the following function of gross income and a parameter κ:

$$T^i \equiv T(z^i, \kappa) = \mathcal{T}(z^i) + \kappa \tau(z^i), \tag{1}$$

which is assumed to be twice differentiable in z^i. This parameterization of the tax function is central to this paper, as it ensures that optimal tax rules can be derived with simple optimization techniques—similar to the case of optimal linear taxation. I refer to κ as the *reform parameter*, and to $\tau(z^i)$ as the *reform function* or simply the reform. The reform parameter takes on an arbitrary value and the reform function depends on whatever reform of the tax schedule one would like to study. The function $\mathcal{T}(z^i)$ is determined to ensure that $T(z^i, \kappa)$ gives the actual tax schedule around which a reform is evaluated. A marginal reform of the income tax can be studied by considering a change $d\kappa$. For a given taxable income z, such reform increases the tax burden by $\tau(z)d\kappa$. As I allow the reform function to depend on z, I can analyze any nonlinear marginal reform of the tax schedule.

2.2 The nature of behavioral responses to taxation

Studies on optimal taxation typically introduce a structural model of individual decision making that determines equilibrium levels of income z^i. However, while behavioral responses to taxation are crucial determinants of optimal taxes, they could be captured by reduced-form elasticities without the need for a structural model. I therefore do not impose any specific model of individual behavior. Nevertheless, writing behavioral responses in terms of elasticities *does* require me to specify how changes in tax rates *may* affect individual income. To that end, I assume that z^i is differentiable in κ. In other words, I rule out that marginal changes in the tax schedule lead to discrete changes in individuals' taxable income. In the typical model of utility-maximizing individuals, this implies that individuals' indifference curves are tangent to the budget curve at exactly one point and that there is no extensive margin.[4] I moreover assume that the derivative of z^i is integrable over the population $\mathcal{I}$.

2.3 The impact of a tax reform on individual taxes

The effect of a reform on an individual's tax burden is obtained by taking the total derivative of Eq. (1):

$$\frac{dT^i}{d\kappa} = \tau(z^i) + T_z^i \cdot \frac{dz^i}{d\kappa}, \tag{2}$$

where a subscript denotes a partial derivative, such that $T_z^i \equiv \partial T(z^i, \kappa)/\partial z^i$ gives the marginal tax rate of an individual with income z^i. An individual's tax burden may be affected both directly by the reform of the tax schedule (first term) and indirectly by an income response to the reform (second term).

---

[4] Jacquet and Lehmann (2021) identify sufficient conditions on structural parameters for this to hold within the context of a multidimensionally heterogeneous population. While it is possible to allow for discrete changes in income, it does complicate derivations and raises the question of exactly *which* discrete changes to consider. In Sect. 6, I discuss how the dual approach can easily be adapted to account for discrete changes along the participation margin.

The same general point can be made for the change in the individual's marginal tax rate, obtained by taking the partial derivative of Eq. (1) with respect to z, and subsequently taking the total derivative with respect to κ:

$$\frac{\mathrm{d}T_z^i}{\mathrm{d}\kappa} = \tau_z(z^i) + T_{zz}^i \cdot \frac{\mathrm{d}z^i}{\mathrm{d}\kappa}. \tag{3}$$

The first term illustrates that the reform raises the marginal tax rate at income level z^i by $\tau_z(z^i)\mathrm{d}\kappa$. A reform-induced change in individual i's taxable income further alters his marginal tax rate as long as the tax schedule is locally nonlinear ($T_{zz}^i \neq 0$). This latter effect is illustrated by the second term in Eq. (3).

2.4 Government revenue

Government revenue equals the integral of all individuals' taxes and is given by:

$$\mathcal{R} \equiv \int_{\mathcal{I}} T(z^i, \kappa) \mathrm{d}i. \tag{4}$$

I do not here concern myself with the expenditure side of the government, but as usual it is straightforward to allow for expenditures on public goods or on some exogenous spending requirement (cf. Christiansen, 1981). The effect on government revenue of a tax reform is obtained by taking the derivative of Eq. (4):

$$\frac{\mathrm{d}\mathcal{R}}{\mathrm{d}\kappa} = \int_{\mathcal{I}} \left(\tau(z^i) + T_z^i \cdot \frac{\mathrm{d}z^i}{\mathrm{d}\kappa} \right) \mathrm{d}i, \tag{5}$$

which equals the integral of Eq. (2). The revenue effects of a tax reform can be decomposed into a mechanical effect and a behavioral effect on the tax base. The mechanical effect indicates that the reform raises an amount $\tau(z^i)\mathrm{d}\kappa$ of resources from every individual $i \in \mathcal{I}$. But a tax reform also tends to affect individuals' taxable income, changing tax revenue by $T_z^i \mathrm{d}z^i$.

2.5 Individual utility

A benevolent social planner cares not only about revenue, but also about the utility of its citizens. Utility in this context refers to an individual's actually experienced utility.[5] The utility of individual i is assumed to be an individual-specific function of the own before- and after-tax income, denoted by $u^i(z^i - T^i, z^i)$.[6] For a given gross income, higher net income allows the individual to consume more and thus tends to

---

[5] Kahneman et al. (1997) distinguish between decision utility and experienced utility. The former is whatever rationalizes individual behavior; the latter is his experienced well-being. Also see Hausman (2011) for a thorough discussion on various concepts of utility.

[6] I abstract from interdependent preferences in which utility of one person depends on the actions of others—as would be the case if an individual were to care about relative as well as absolute income.

raise his utility. For a given net income, higher gross income implies that the individual needs to exert more effort in earning income and thus tends to lower his utility. As the income tax is itself a function of gross income and the reform parameter, we can write utility as the following function:

$$U^i \equiv U^i(z^i, \kappa) = u^i(z^i - T(z^i, \kappa), z^i). \tag{6}$$

As with taxable income, I assume that utility and its derivatives are integrable over the population $\mathcal{I}$.

Individuals may or may not choose their income levels to maximize utility.[7] I define ω^i as the "behavioral wedge" between the marginal rate of substitution and the marginal rate of transformation between gross and net income:

$$\omega^i \equiv \frac{-U_z^i}{u_c^i} = \left(\frac{-u_z^i}{u_c^i} - (1 - T_z^i) \right), \tag{7}$$

where $u_c^i \equiv \partial u^i / \partial (z^i - T^i)$. The behavioral wedge measures the degree to which an individual works "too much" because of behavioral biases. The marginal rate of substitution $(-u_z^i / u_c^i)$ measures the consumption-equivalent utility *loss* of raising gross income. The marginal rate of transformation $(1 - T_z^i)$ measures the consumption-equivalent utility *gain* of raising gross income. If individual i maximizes his utility, both terms cancel out and the behavioral wedge equals $\omega^i = 0$. However, the behavioral wedge will be positive $(\omega^i > 0)$ or negative $(\omega^i > 0)$ if individual i works "too much" or "too little." To be precise, individual i's income would have corresponded to utility maximization had his marginal tax rate (T_z^i) been ω^i percentage points lower.

2.6 Social welfare

I assume a welfarist social objective, such that social welfare can be written as a (weighted) integral of individual utility:

$$\mathcal{W} \equiv \int_{\mathcal{I}} \gamma^i U^i \mathrm{d}i, \tag{8}$$

where γ^i is an individual-specific weight that determines the importance of individual i's utility within the social objective. In the special case of a utilitarian social objective, $\gamma^i = \gamma$ for all i. The effect of a tax reform on the social objective is

---

[7] There are numerous reasons why taxable income might not be chosen to maximize utility. For example, individuals might have mistaken beliefs about the shape of their budget curve (e.g., Chetty et al. 2009; Liebman and Zeckhauser, 2004) or about their own utility function (e.g., Loewenstein et al. 2003). Another reason might be that the tax base is not fully under control of the individual. For example, in Piketty et al. (2014) and Rothschild and Scheuer (2016) the tax base is partly the result of third party bargaining or rent-seeking efforts, whereas in Stantcheva (2014) and Bastani et al. (2015), employers decide—sometimes against the interest of individual employees—on the number of working hours.

obtained by taking the derivative of Eq. (8) with respect to κ. Doing so, while substituting for Eq. (7), yields:

$$\frac{\mathrm{d}\mathcal{W}}{\mathrm{d}\kappa} = -\int_{\mathcal{I}} \gamma^i u_c^i \left(\tau(z^i) + \omega^i \cdot \frac{\mathrm{d}z^i}{\mathrm{d}\kappa} \right) \mathrm{d}i. \tag{9}$$

As with government revenue, a reform's effect on social welfare can be decomposed into a mechanical effect and a behavioral effect. The first term within brackets, representing the mechanical effect, reflects the direct social welfare loss from reducing individuals' net income by $\tau(z^i)\mathrm{d}\kappa$. The second term within brackets represents the reform's behavioral effect on social welfare. If the reform causes individuals to increase their gross income ($\mathrm{d}z^i/\mathrm{d}\kappa > 0$), it reduces social welfare if their income is already chosen too high ($\omega^i > 0$) and raises social welfare if their income is chosen too low ($\omega^i < 0$). The opposite holds if the reform causes individuals to reduce their gross income ($\mathrm{d}z^i/\mathrm{d}\kappa < 0$). Naturally, if individuals choose their tax base to maximize utility ($\omega^i = 0$), a reform only affects social welfare through its mechanical effect.

3 Net social welfare effect of a tax reform

3.1 …in terms of the behavioral policy response

The net social welfare effects of a tax reform account for its impact on both the social welfare function and government revenue. Denote the social marginal value of public resources by λ. Furthermore, denote the social welfare weight of individual i by $g^i \equiv \gamma^i u_c^i / \lambda$. It equals the social marginal value of individual consumption in terms of public resources. I can now formulate the following Proposition.

Proposition 1 *The marginal net social welfare effect of a nonlinear reform $\tau(\cdot)$ is given by*:

$$\frac{\mathrm{d}\mathcal{W}/\lambda}{\mathrm{d}\kappa} + \frac{\mathrm{d}\mathcal{R}}{\mathrm{d}\kappa} = \int_{\mathcal{I}} \left((1 - g^i)\tau(z^i) + (T_z^i - g^i\omega^i)\frac{\mathrm{d}z^i}{\mathrm{d}\kappa} \right) \mathrm{d}i. \tag{10}$$

Proof The left-hand side is the definition of the net social welfare effect of a tax reform. The right-hand side is obtained by substituting Eqs. (5) and (9). $\qquad\square$

Equation (10) decomposes the welfare impact of any reform into a redistributional and a behavioral effect. The first term within large brackets gives the redistributional gain of the tax reform. The government redistributes $\tau(z^i)\mathrm{d}\kappa$ resources away from every individual i, and toward the government budget. The net gain per unit of redistributed resources equals the revenue gain minus the utility loss $(1 - g^i)$. The second term within large brackets gives the behavioral effect of the reform. To the extent that a reform raises individual i's gross income ($\mathrm{d}z^i/\mathrm{d}\kappa > 0$), the behavioral effect raises government revenue proportional to the marginal tax rate (T_z^i) and

lowers utility proportional to the behavioral wedge (ω^i). The behavioral wedge is weighted by the social welfare weight (g^i) to obtain the social value of the behavioral utility effect.

Proposition 1 writes the net social welfare effect in terms of the reform's effect on taxable income $dz^i/d\kappa$—what Hendren (2016) dubs the policy response or policy elasticity. For most reforms, however, we typically lack empirical estimates of the policy response to that specific reform. This makes it impossible to directly calibrate Eq. (10). It is therefore useful to rewrite the net social welfare effect in terms of elasticities that we *do* often measure.

3.2 ...in terms of net-of-tax rate elasticities

Although we typically lack evidence on the policy response to any given reform, we do have estimates of how individuals' taxable income responds to changes in their net-of-tax rates, as well as how it responds to changes in disposable income (e.g., Blundell and MaCurdy, 1999; Saez et al., 2012; Chetty, 2012). Writing the welfare effects of a tax reform in terms of these "known" elasticities requires us to impose more structure on the behavioral responses to taxation. In particular, it necessitates the assumption that an individual's income *only* responds to changes in the "own" marginal tax rate (T^i_z) and tax burden (T^i).[8] In that case, we can define the policy response of taxable income as:

$$\frac{dz^i}{d\kappa} \equiv -\frac{z^i}{1-T^i_z}\left(e^i_c \tau_z(z^i) + \eta^i \cdot \frac{\tau(z^i)}{z^i} \right), \tag{11}$$

where e^i_c and η^i measure the responsiveness of individual i's tax base to, respectively, changes in marginal tax rates and changes in the tax burden.

More specifically, the compensated elasticity e^i_c is defined as the percent change in taxable income (z^i) due to a reform that raises the marginal net-of-tax rate ($1 - T^i_z$) by one percent, while leaving the tax burden unchanged. To see this, set $\tau(z^i) = 0$ and rearrange Eq. (11) to find:

$$e^i_c \equiv -\left. \frac{dz^i}{\tau_z(z^i)d\kappa} \frac{1-T^i_z}{z^i} \right|_{\tau(z^i)=0}. \tag{12}$$

Notice that the exogenous change in the net of tax rate is given by $-\tau_z(z^i)d\kappa$, so that e^i_c indeed measures the compensated net-of-tax rate elasticity of taxable income.

---

[8] Thus, I abstract from the possibility that individual i's taxable income is affected by the marginal or average tax rates of any individual j with $z^j \neq z^i$. By implication, I also abstract from the possibility that individual i's taxable income is affected by changes in taxable income of any individual $j \neq i$. See Farhi and Gabaix (2020) for an example in which individuals' tax bases are potentially responsive to what they call "ambient" or "non-own" tax rates.

The income effect η^i measures the change in disposable income due to a reform that lowers the tax burden by one unit, while leaving marginal taxes unchanged. To see this, set $\tau_z(z^i) = 0$ and rearrange Eq. (11) to find:

$$\eta^i \equiv - \left. \frac{(1 - T_z^i)\mathrm{d}z^i}{\tau(z^i)\mathrm{d}\kappa} \right|_{\tau_z(z^i)=0}. \tag{13}$$

Notice that the exogenous change in the tax burden is given by $\tau(z^i)\mathrm{d}\kappa$, so that η^i indeed measures the effect of an exogenous reduction in the tax burden.

The compensated elasticity and the income effect are defined as relative changes in income along the actual budget curve—as in Jacquet et al. (2013)—and not as changes along a linearized 'virtual' budget line—as in Saez (2001). That is, e_c^i and η^i take into account that changes in taxable income affect an individual's marginal tax rate, which in turn affects taxable income, and so on. The advantage of defining behavioral effects as moves along the actual budget curve is that it allows me to later on express the optimal tax schedule in terms of these elasticities and characteristics of the actual income distribution, rather than a virtual income distribution.[9]

Decomposing the policy response in terms of elasticities with respect to the "own" tax rates yields the following Corollary.

Corollary 1 *If income responds only to changes in "own" tax rates, the marginal net social welfare effect of a nonlinear reform $\tau(\cdot)$ is given by*:

$$\frac{\mathrm{d}\mathcal{W}/\lambda}{\mathrm{d}\kappa} + \frac{\mathrm{d}\mathcal{R}}{\mathrm{d}\kappa} = \int_{\mathcal{I}} \left((1 - g^i)\tau(z^i) - \frac{T_z^i - g^i\omega^i}{1 - T_z^i} \cdot (\eta^i\tau(z^i) + z^i e_c^i \tau_z(z^i)) \right)\mathrm{d}i. \tag{14}$$

Proof Substitute Eq. (11) into Eq. (10) to obtain Eq. (14). □

A tax reform can be seen to have three effects on social welfare, illustrated in expression (14). The first term gives the redistributional gain as discussed above. The second

---

term gives the behavioral income effects of a tax reform. As long as $\eta^i < 0$, an increase in individual i's tax burden ($\tau(z^i) > 0$) leads him to increase his income. This leads to an increase in tax revenue if the tax rate is positive ($T_z^i > 0$) and a decrease in utility if the behavioral wedge is positive ($\omega^i > 0$) or an increase in utility if the behavioral wedge is negative ($\omega^i < 0$). The third term gives the behavioral substitution effects of a tax reform. An increase in the marginal income tax ($\tau_z(z^i) > 0$) leads to a reduction in taxable income as long as $e_c^i > 0$. This reduction leads to tax revenue losses (if $T_z^i > 0$) and to utility gains (if $\omega^i > 0$) or losses (if $\omega^i < 0$).

Corollary 1 and Eq. (14) are central to the analysis of the rest of this paper. It determines both optimal taxes and the desirability of limited reforms outside the optimum. To see this, notice that taxes can only be set optimally if the marginal net social welfare effect of any reform is nil. Thus, the optimal tax schedule is determined by equating expression (14) to zero for any possible nonlinear tax reform $\tau(\cdot)$. Indeed, the next section sheds more light on the optimal tax schedule by considering two specific reforms for which the net marginal social welfare gains are set to zero. Furthermore, expression (14) also plays a central role when considering limited tax reforms outside the optimum. Such a tax reform is desirable if and only if expression (14) is positive for that specific reform $\tau(\cdot)$. In Sect. 5, I further elaborate on this.

4 Optimal taxation

4.1 Reform 1: A uniform tax increase

Taxes are set optimally if no reform of the tax schedule can raise net social welfare. A full characterization of optimal tax rates can thus be obtained by equating expression (14) to zero for all possible reforms $\tau(\cdot)$. To obtain more insight into what constitutes an optimal tax schedule, I focus here on two specific reforms. The first reform raises the tax burden uniformly across individuals by $d\kappa$, such that $\tau(z^i) = 1$ and $\tau_z(z^i) = 0$ for all i. This reform is illustrated in Fig. 1, panel (a). Substituting the reform function into expression (14), while equating it to zero, yields:

$$\int_{\mathcal{I}} \left(1 - g^i - \frac{T_z^i - g^i \omega^i}{1 - T_z^i} \cdot \eta^i \right) di = 0. \tag{15}$$

As the reform leaves all marginal tax rates unchanged, it does not generate any substitution effect, and only affects social welfare through its redistribution from taxpayers to government and through behavioral income effects.

To further interpret Eq. (15), it is useful to introduce a term to denote the social marginal value of individual i's private resources in terms of public resources. This term is given by:

$$\alpha^i \equiv g^i + \frac{T_z^i - g^i \omega^i}{1 - T_z^i} \cdot \eta^i. \tag{16}$$

Fig. 1 Three reform functions $\tau(z)$. Panel **a** illustrates reform 1, a uniform tax increase. Panel **b** illustrates reform 2, an increase in the marginal tax rate around one income level z^*. Panel **c** illustrates reform 3, an increase in a bracket's tax rate between z^a and z^b

(a) Reform 1

$\tau(z)$

(b) Reform 2

$\tau(z)$

z^* $z^* + \mathrm{d}z$

(c) Reform 3

$\tau(z)$

$z^b - z^a$

z^a z^b
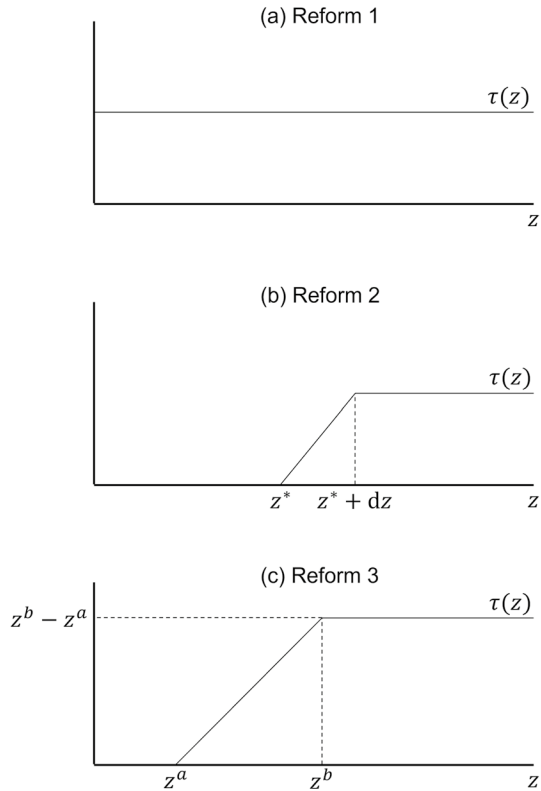
Denoted in terms of public resources, a marginal unit increase in individual i's income yields additional social utility of consumption equal to g^i. On top of that, it induces an income effect on taxable income, causing a revenue effect equal to $T_z^i \eta^i / (1 - T_z^i)$, and a further social utility effect equal to $-g^i \omega^i \eta^i / (1 - T_z^i)$. Taken together, α^i indicates how many resources government is willing to give up in order to provide individual i with an additional unit of income.[10] The pattern of α^i determines the social willingness to redistribute between any pair of individuals, i.e., the social planner values redistribution of resource from individual i to individual j if $\alpha^i < \alpha^j$. I can now formulate the following proposition.

Proposition 2 *In the tax optimum, the average social marginal value of private resources must equal the social marginal value of public resources*:

$$\int_{\mathcal{I}} \alpha^i \mathrm{d}i = 1. \tag{17}$$

---

[10] It also corresponds to what Diamond (1975) called the social marginal utility of individual income, divided by the social marginal value of public resources λ.

Proof Substitute Eq. (16) into (15) and rearrange to obtain Eq. (17). □

Proposition 2 implies that in the optimum, a marginal transfer of resources from everyone in the private sector to the public sector does not affect net social welfare. This simple optimality condition has important consequences for public policy. As documented by Jacobs (2018), it implies that the marginal cost of public funds—defined as the inverse of the left-hand side of Eq. (17)—equals one in the tax optimum.[11] As a result, evaluations of public projects should not inflate the financing costs of these projects simply because of the existence of distortive taxes. Since a nonlinear tax schedule implies that government has access to nondistortive taxes—as illustrated by the reform I consider here—distortions are irrelevant for the marginal financing costs of a project. This validates standard cost-benefit analyses (cf., Christiansen, 1981).

4.2 Reform 2: Raising the marginal tax rate

The second reform I consider raises the marginal tax rate over an infinitesimal interval above some income level z^*, such that $\tau_z(z^i) > 0$ for $z^* \leq z^i < z^* + \mathrm{d}z$. As a result, the tax burden remains constant for everyone below this interval, but is raised for everyone above the interval. That is, $\tau(z^i) = 0$ for all $z^i < z^*$, and $\tau(z^i) > 0$ for $z^i \geq z^* + \mathrm{d}z$. This reform is illustrated in Fig. 1, panel (b). Substituting the reform function into the social welfare effects as given by Eq. (14), equating it to zero, and letting the interval $\mathrm{d}z$ go to zero, yields:

$$\int_{\mathcal{I}:z^i>z^*} \left(1 - g^i - \frac{T_z^i - g^i\omega^i}{1 - T_z^i} \cdot \eta^i\right)\mathrm{d}i = \lim_{\mathrm{d}z \to 0} \int_{\mathcal{I}:z^* \leq z^i < z^*+\mathrm{d}z} \left(\frac{T_z^i - g^i\omega^i}{1 - T_z^i} \cdot z^i e_c^i\right)\frac{\mathrm{d}i}{\mathrm{d}z}. \tag{18}$$

To derive this equation, I substituted for $\tau_z(z^*) = \lim_{\mathrm{d}z \to 0} \tau(z^* + \mathrm{d}z)/\mathrm{d}z$, which follows from the definition of the derivative.

4.2.1 ...when individuals maximize utility

To clarify the implications of Eq. (18) for optimal income taxes, I first concentrate on the special case in which individuals maximize utility, such that $\omega^i = 0$ for all i. The number of individuals that face an increase in marginal taxes can be written as $\lim_{\mathrm{d}z \to 0}(H(z^* + \mathrm{d}z) - H(z^*)) = \lim_{\mathrm{d}z \to 0} h(z^*)\mathrm{d}z$. This allows me to write the average compensated elasticity of individuals with income level z^* as:

$$\bar{e}_c^* \equiv \lim_{\mathrm{d}z \to 0} \int_{\mathcal{I}:z^* \leq z^i < z^*+\mathrm{d}z} e_c^i \frac{\mathrm{d}i}{h(z^*)\mathrm{d}z} \tag{19}$$

---

[11] Outside the optimum, the marginal cost of public funds may be either greater or less than one, depending on one's distributional preferences and the tax instrument in question—see Jacobs (2018).

Moreover, I define the average social marginal value of private resources of individuals who earn more than z as $\bar{\alpha}_{z^i>z} \equiv \int_{\mathcal{I}:z^i>z} \alpha^i \mathrm{d}i / \int_{\mathcal{I}:z^i>z} \mathrm{d}i$. With the help of these definitions, I can now formulate the following Proposition.

Proposition 3 *In the tax optimum with utility-maximizing individuals, the marginal tax rate at income level* z^*, *denoted by* $T_z^* \equiv T_z(z^*, \kappa)$, *must satisfy the following condition*:

$$\frac{T_z^*}{1 - T_z^*} = \frac{1}{\bar{e}_c^*} \cdot \frac{1 - H(z^*)}{z^* h(z^*)} \cdot \left(1 - \bar{\alpha}_{z^i>z^*}\right). \tag{20}$$

Proof Substituting $\omega^i = 0$, Eq. (19), and the definition of $\bar{\alpha}_{z^i>z}$ into Eq. (18) and rearranging yields Eq. (20). □

Equation (20) presents the canonical formula for the optimal marginal tax rate if people differ in their elasticity of taxable income. Variations of this result appear in Diamond (1998), Saez (2001), Piketty and Saez (2013), and Jacquet and Lehmann (2021). Those earlier contributions elaborately discuss the intuition behind the optimal tax formula.

The optimal marginal tax rate at income level z^* depends on four terms. It is decreasing in the average elasticity of the tax base $\bar{e}_c^*$ as this raises the distortive costs of the marginal tax rate. It is decreasing in the income concentration $z^* h(z^*)$ as this raises the amount of income that is distorted by the marginal tax rate. It is increasing in the share of people with higher income $1 - H(z^*)$ as this increases the amount of redistribution caused by the marginal tax rate. And, finally, it is decreasing in the marginal value of private resources of the relatively rich $\bar{\alpha}_{z^i>z^*}$ as this lowers the value of redistribution.

4.2.2 ...when individuals do not maximize utility

Now, consider the general case in which individuals do not necessarily choose their tax base to maximize utility, so that ω^i might be nonzero. Before deriving the optimal tax formula, it is useful to define the income-conditional covariance between two variables as $\mathrm{cov}[x^i, y^i] \equiv \overline{x^i y^i} - \bar{x}^i \bar{y}^i$, where an overline indicates average values *conditional on labor income* z^i. This definition allows us to rewrite part of the right-hand side of Eq. (18) as:

$$\begin{aligned}
\lim_{\mathrm{d}z \to 0} \int_{\mathcal{I}:z^* \leq z^i < z^*+\mathrm{d}z} g^i \omega^i e_c^i \frac{\mathrm{d}i}{\mathrm{d}z} &= \overline{g^* \omega^* e_c^*} h(z^*) \\
&= \left(\overline{g^* \omega^* e_c^*} - \overline{g^* \omega^*} \bar{e}_c^* + \overline{g^* \omega^*} \bar{e}_c^* - \bar{g}^* \bar{\omega}^* \bar{e}_c^* + \bar{g}^* \bar{\omega}^* \bar{e}_c^*\right) h(z^*) \\
&= \left(\bar{g}^* \bar{\omega}^* + \mathrm{cov}[g^*, \omega^*] + \mathrm{cov}\left[g^* \omega^*, \frac{e_c^*}{\bar{e}_c^*}\right]\right) \bar{e}_c^* h(z^*).
\end{aligned} \tag{21}$$

The first line follows from the recognition that $\int_{\mathcal{I}:z^* \leq z^i < z^*+\mathrm{d}z} x^i \mathrm{d}i$ measures the sum of x^i over all individuals that earn around z^*. Thus, division by the number of individuals $h(z^*)\mathrm{d}z$ yields the income-conditional average $\bar{x}^*$. The second line follows from

adding and subtracting identical terms. The third line follows from the definition of the income-conditional covariance. Substituting this back into Eq. (18) yields the following Proposition.

Proposition 4 *In the tax optimum with individuals who might not maximize their own utility, the marginal tax rate at income level z^* must satisfy the following condition:*

$$\frac{T_z^* - \bar{g}^*\bar{\omega}^* - \text{cov}[g^*, \omega^*] - \text{cov}\left[g^*\omega^*, \frac{e_c^*}{\bar{e}_c^*}\right]}{1 - T_z^*} = \frac{1}{\bar{e}_c^*} \cdot \frac{1 - H(z^*)}{z^* h(z^*)} \cdot \left(1 - \bar{\alpha}_{z^i > z^*}\right).$$

(22)

Proof Substituting Eqs. (19) and (21), the definitions of the income distribution, and the average social value of private resources into Eq. (18), and rearranging yields Eq. (22). ☐

Equation (22) gives the canonical formula for the optimal marginal tax rate with behavioral biases. Elaborate discussions of the intuition behind this result are provided by Kanbur et al. (2006), Gerritsen (2016), and Farhi and Gabaix (2020). Compared to the results in Eq. (20), there are three new terms on the left-hand side of the optimal tax formula. These terms indicate how the marginal tax rate at z^* should be adjusted to "correct" the behavior of individuals with income z^*.

The first novel term ($\bar{g}^*\bar{\omega}^*$) shows that marginal tax rates should be higher if individuals work too much ($\bar{\omega}^* > 0$) and government cares about their utility ($\bar{g}^* > 0$). The second novel term ($\text{cov}[g^*, \omega^*]$) shows that taxes should be even higher if government cares more about the utility of people that overwork relatively much (if g^i is increasing in ω^i for individuals with income $z^i = z^*$).

Finally, the third term ($\text{cov}[g^*\omega^*, e_c^*/\bar{e}_c^*]$) shows that the corrective argument for taxation depends on the tax responsiveness of biased individuals. In particular, it could be that individuals with larger deviations from utility maximization are less responsive to changes in tax rates.[12] This would imply that the degree of misoptimization is negatively correlated with behavioral elasticities ($\text{cov}[g^*\omega^*, e_c^*/\bar{e}_c^*] < 0$) if individuals with income z^* mistakenly earn too much on average. Conversely, this correlation would be positive ($\text{cov}[g^*\omega^*, e_c^*/\bar{e}_c^*] > 0$) if they earn too little on average. This would mean that the covariance of the third novel term in Eq. (22) takes on the opposite sign of the first novel term ($\bar{g}^*\bar{\omega}^*$). The corrective argument for taxation becomes weaker as a result, bringing optimal tax rates closer to the ones obtained with utility-maximizing individuals.

---

[12] Chetty et al. (2014) make an argument to this effect within the context of subsidies for retirement savings, see also Chetty (2015).

5 The desirability of limited reforms

5.1 Reform 3: Raising a bracket's tax rate

Contrary to much of the literature on optimal taxation, actual tax policy is typically concerned with some limited tax reform rather than a search for the best possible tax system. Moreover, the actual tax system might be far from optimal so that the reform should be evaluated outside the tax optimum. The primal approach is ill-equipped to deal with these issues, as it is concerned with the effects of changes in allocations rather than changes in taxes. The dual approach, on the other hand, is ideally situated to deal with issues of actual tax policy. To see this, note that as long as small changes in tax rates lead to only small behavioral changes in income, the welfare effects identified in Eq. (14) are valid for any small reform $\tau(z)$ and for any optimal or suboptimal initial allocation.[13]

To show how the dual approach can directly generate insights for actual tax policy, I consider a reform that is part of a policy maker's or politician's typical range of policy options: a tax rate increase for a specific tax bracket.[14] Rather than focusing on the optimal level of the tax rate, I simply determine whether raising the rate is desirable or not, and how this depends on features of the actual, possibly suboptimal, tax system. For simplicity, I disregard income effects on the tax base ($\eta^i = 0$) and suboptimal behavior ($\omega^i = 0$).[15] Consider a tax bracket that applies to gross income between z^a and z^b. A tax reform that raises this bracket's tax rate by $d\kappa$ can be modeled as $\tau(z) = 0$ for $z < z^a$, $\tau(z) = (z - z^a)$ for $z \in [z^a, z^b]$, and $\tau(z) = (z^b - z^a)$ for $z > z^b$. This indeed implies that $\tau_z(z) = 1$ for $z \in [z^a, z^b]$ and $\tau_z(z) = 0$ otherwise. The reform is illustrated in Fig. 1, panel (c). Corollary 1 establishes that this reform raises net social welfare if and only if expression (14) is strictly positive. Substituting the reform into expression (14), we thus get the following desirability condition for increasing the bracket's tax rate:

$$\int_{\mathcal{I}:z^i \in [z^a, z^b]} (z^i - z^a)(1 - g^i)\mathrm{d}i + \int_{\mathcal{I}:z^i > z^b} (z^b - z^a)(1 - g^i)\mathrm{d}i$$
$$> \int_{z^a}^{z^b} \frac{T_z}{1 - T_z} \cdot \bar{e}_c \cdot z^i h(z^i) \cdot \mathrm{d}z^i, \tag{23}$$

---

[13] If there are kinks in the actual tax system, then there may be bunching among tax payers—which would violate the assumption that small changes in tax rates lead to only small behavioral changes in income. Thus, the latter assumption does impose restrictions on the status quo tax system. At the same time, actual bunching at kink points is rarely observed for individuals who are not self-employed (e.g., Saez, 2010). Thus, the assumption that small changes in tax rates only lead to small behavioral changes in income may still be reasonable—even if there are kinks in the tax schedule.

[14] See Bierbrauer et al. (2022) for an application of the dual approach to a reform of the earned income tax credit.

[15] Saez et al. (2012) note that there is little compelling evidence on significant income effects on taxable income.

where I substituted for the income density on the right-hand side. The left-hand-side of Eq. (23) represents the redistributional benefits of the reform. It gives the difference between the social marginal value of public resources and the social marginal value of private resources for every mechanical unit of tax revenue raised from individuals within the bracket (first integral) and from individuals above the bracket (second integral). Thus, an individual i within the bracket sees his tax burden increase by $(z^i - z^a)\mathrm{d}\kappa$, whereas the tax burden of an individual i above the tax bracket increases by $(z^b - z^a)\mathrm{d}\kappa$. The total redistributional benefits of the reform generally depend on welfare weights g^i, which ultimately makes desirability a matter of political judgment.[16]

Whereas the redistributional benefits of the reform crucially depend on political values, we can say more about the distortionary costs of the reform, given by the right-hand side of Eq. (23). As usual, these costs are increasing with the responsiveness of the tax base, as measured by the compensated elasticity, the marginal tax wedges within the bracket, and the amount of income that falls within the bracket. Notice, however, that the distortionary costs do not simply equal the product of these three factors' averages. As can be seen from Eq. (23), it also matters how these factors are correlated. This issue is sidestepped by almost every study that measures the distortionary costs of raising the tax rate within a certain income interval. That is, the literature typically assumes that both the marginal tax rates and the elasticity are constant over the interval of interest. In that case, the marginal distortionary costs indeed reduce to the product of the elasticity, the tax wedge, and the amount of income within the interval.[17]

In reality, tax schedules tend to be highly nonlinear, causing this approach to yield biased estimates of the marginal distortionary costs of taxation. Nonlinearities in actual tax schedules stem from means-tested welfare programs such as an earned income tax credit, rental support, or child benefits, as well as different tax brackets. As means-tested benefits are being phased out with income, marginal taxes tend to decline with income from relatively high rates in the phase-out interval to relatively low rates in the *phased-out* interval. The same income range is typically associated with increasing income concentrations. Equation (23) then tells us that the distortionary costs of a bracket's tax rate are lower if this bracket overlaps with the phase-out (and phased-out) income interval of means-tested welfare programs.

---

[16] The only exception is if Eq. (23) is strictly violated even if $g^i = 0$ for all affected individuals. In that case, it is beneficial to *lower* the tax rate even if government does not care about the individuals who receive the tax cut. In other words, it would indicate that the status quo is a Pareto inefficient tax system with tax rates beyond the top of the Laffer curve. Also see Bierbrauer et al. (2022), who apply the dual approach to derive a test for the Pareto efficiency of the income tax schedule.

[17] See, for example, Kleven and Kreiner (2006) for a prominent study that provides estimates of tax distortions for 10 different income intervals; a recent study by Blomquist and Simula (2019), who estimate the marginal deadweight loss of increasing the marginal income tax across the entire population, do properly account for the nonlinearity of the tax schedule.

6 Broader applicability of the dual approach

The focus of this paper has been on illustrating how the dual approach can be applied to solve for optimal nonlinear income taxes. I show this within a standard context with individuals that only make one intensive-margin decision on the size of their tax base—while allowing for heterogeneous preferences and individual utility misoptimization. However, the dual approach is versatile enough to be much more broadly applicable. In what follows, I therefore illustrate how the above analysis can be adjusted to take into account various nonlinear reforms outside the optimum, multiple intensive decision margins, a participation margin, and multiple tax bases that are subject to separate nonlinear tax schedules.

Nonlinear reforms outside the optimum—The third reform in the previous section just looked at one specific tax reform that might be relevant for actual policy making. That reform was essentially linear—raising the proportional tax rate of a specific bracket—though evaluated within the context of an actual nonlinear schedule of effective marginal tax rates. However, the dual approach can be readily applied to more complicated nonlinear reforms that play a role in actual policy discussions. For example, one could analyze different types of phase-out schedules for the EITC or other welfare programs, or changes to a quadratic tax schedule.[18] Is it better to phase out the EITC at a linear rate—raising effective marginal tax rates by the same amount across the phase-out range—or at an increasing or decreasing rate? Introducing an increasing phase-out rate within the range $[z^a, z^b]$ could be modeled with a specific reform function $\tau(z)$ with $\tau_z(z) > 0$ and increasing over the phase-out range. Conversely, a decreasing phase-out rate could be modeled with a reform function that has $\tau_z(z) > 0$ and decreasing over the phase-out range. As before, substituting these reforms into Eq. (14) allows one to readily evaluate the welfare consequences of either phase-out function for any arbitrary initial tax schedule.

Multiple intensive margins— It is straightforward to allow individuals to make more decisions than only the one that determines their tax base. As long as these decisions are unobservable to the tax authority, and therefore untaxed, the analysis remains unchanged in the case of utility-maximizing individuals. Then, even if a tax reform affects individual behavior on these additional decision margins, this does not affect their utility (because of individual utility maximization), nor does it affect government revenue (because the additional decisions are untaxed).

This convenient conclusion no longer holds if individuals do not perfectly maximize utility when making these additional decisions. To see this, notice that the term ω^i enters Eq. (14) as a welfare effect of the tax reform. With multiple decision margins, similar terms for every decision margin would enter Eq. (14), thereby yielding multiple corrective reasons for marginal taxes. As a simple example, imagine that individuals perfectly maximize utility when deciding on their (taxed) labor income, but mistakenly consume too much and save too little of their earned income. Then if future consumption is complementary with

---

[18] An example of a country with quadratic tax schedule is Germany, where the income of most households fall in tax brackets with linearly increasing marginal tax rates.

leisure, higher labor income taxes would be helpful in correcting individuals' savings decision even though there is no need for a labor-supply correction.

Participation margin—The analysis can further be adapted to allow for a participation margin. For simplicity, I only consider the standard case in which individuals with the same income have the same intensive-margin elasticities, and in which individuals maximize their utility. The latter assumption ensures that a small tax reform only mechanically affects individuals' utility due to changes in tax burdens, but not through behavioral changes. As a result, a reform of the marginal income tax affects individuals' utility in essentially the same way as in the case without a participation margin. I can therefore focus attention on how adding a participation margin affects a reform's effect on government revenue.

For this, I refine the definition of z^i as the "notional tax base," i.e., the tax base individual i would choose if he decides to participate. His actual tax base when deciding not to participate equals 0. I furthermore introduce a parameter $\pi^i(\kappa)$ that indicates the share of labor market participants among individuals with notional income z^i. The government budget can then be rewritten as:

$$\mathcal{B} = \int_{\mathcal{I}} \left(\pi^i(\kappa) T(z^i, \kappa) + (1 - \pi^i(\kappa)) T(0, \kappa) \right) \mathrm{d}i, \tag{24}$$

which gives the integral over participants' and non-participants' tax burdens. Taking derivatives, the effect of a marginal tax reform on government revenue can be seen to equal:

$$\frac{\mathrm{d}\mathcal{B}}{\mathrm{d}\kappa} = \int_{\mathcal{I}} \left(\pi^i(\kappa) \left(\tau(z^i) + T_z^i \frac{\mathrm{d}z^i}{\mathrm{d}\kappa} \right) + (1 - \pi^i(\kappa)) \tau(0) + \left(T^i - T^0 \right) \frac{\mathrm{d}\pi^i}{\mathrm{d}\kappa} \right) \mathrm{d}i, \tag{25}$$

with $T^0 \equiv T(0, \kappa)$. Thus, the reform yields mechanical revenue changes for both participants and non-participants, an intensive behavioral effect on the tax base ($\mathrm{d}z^i/\mathrm{d}\kappa$), and an extensive behavioral effect on the tax base ($\mathrm{d}\pi^i/\mathrm{d}\kappa$). The latter behavioral response would typically be unaffected by changes in marginal taxes, but responsive to changes in average tax rates. As a result, the total welfare effect of an increase in the marginal tax rate at z^* now includes the reduced government revenue due to lower participation rates among individuals whose notional income exceeds z^*. This additional cost of taxation should be taken into account in the optimum and tends to reduce optimal marginal tax rates.

Multiple tax bases—The dual approach can also be fruitfully employed to study the desirability of other types of government policy in combination with a nonlinear tax schedule. For linear commodity taxation and public good provision, this has previously been illustrated by Christiansen (1981, 1984). But one can also deal with multiple nonlinear tax schedules as in the case of labor-income and capital-income taxes (e.g., Gerritsen et al., 2022). For example, let T^z denote a nonlinear labor-income tax with tax base z, and T^y a nonlinear capital income tax with tax base y. Similar to the analysis above, both nonlinear taxes can be parameterized as $T^z(z, \kappa^z)$ and $T^y(z, \kappa^y)$ to allow for straightforward welfare analysis of any nonlinear reform of either tax.

7 Conclusion

This paper develops a method to solve for the optimal nonlinear income tax based on the dual approach. The procedure is intuitive and remains close in spirit to actual tax policy. It moreover relies on optimization techniques that are well-known to any undergraduate student of economics, which should make it easier to convey key results to policy makers and students, as well as other academic scholars. I showed that the approach can be applied to not only obtain well-known results in a more intuitive way, but also to solve for optimal nonlinear taxes when individuals have heterogeneous preferences and when they do not perfectly maximize their utility. It moreover allows one to gain new insights into the welfare effects of limited tax reforms outside the optimum, something for which the primal approach is especially ill-suited. I furthermore sketched how the dual approach can be applied to deal with nonlinear tax reforms outside the optimum, and with multiple decision margins, a participation margin, and multiple nonlinear tax bases.

Appendix

Actual and virtual behavioral responses to taxation

The elasticities that are used in most studies on optimal taxation represent behavioral responses to taxation that would occur in the hypothetical case in which an individual's actual nonlinear budget curve were to be replaced by a linear budget line. This "virtual" budget line is defined so that it is tangent to the actual nonlinear budget curve at the point of the individual's actual income-consumption decision. The virtual budget line can be written as $c^i = (1 - T_z^i)z^i + V^i$, with V^i termed virtual income and the marginal tax rate T_z^i assumed invariant to z^i. The income-consumption point that an individual chooses on this budget line depends on its intercept and slope, and thus on the marginal tax rate and virtual income. We can therefore write $z^i = \tilde{z}^i(V^i, T_z^i)$. The virtual uncompensated elasticity gives the relative change in labor income along the virtual budget line due to a relative increase in the marginal net-of-tax rate and for a given virtual income. It is given by:

$$\tilde{e}_u^i \equiv \frac{1 - T_z^i}{z^i} \frac{\partial \tilde{z}(V^i, T_z^i)}{-\partial T_z^i}. \tag{26}$$

Intuitively, it gives the behavioral response that results from rotating the virtual budget line counter-clockwise around its intercept. Since the budget line rotates around its intercept, the uncompensated elasticity represents both a substitution and an income effect. The virtual income effect is given by:

$$\tilde{\eta}^i \equiv (1 - T_z^i)\frac{\partial \tilde{z}(V^i, T_z^i)}{\partial V^i}, \tag{27}$$

which represents the behavioral response that results from an upward shift of the budget line. Finally, the virtual compensated elasticity, like the uncompensated one, gives the relative change in labor income along the virtual budget line in response to a relative increase in the marginal net-of-tax rate. This time, however, virtual income is simultaneously decreased to ensure that the budget line passes through the initial equilibrium. The Slutsky equation implies that the virtual compensated elasticity is given by:

$$\tilde{e}_c^i \equiv \tilde{e}_u^i - \tilde{\eta}^i = \frac{1 - T_z^i}{z^i} \frac{\partial \tilde{z}(V^i, T_z^i)}{-\partial T_z^i} - (1 - T_z^i) \frac{\partial \tilde{z}(V^i, T_z^i)}{\partial V}. \tag{28}$$

Intuitively, it gives the behavioral response that results from rotating the virtual budget line counter-clockwise around (z^i, c^i).

While these virtual behavioral effects are widely used in the literature, they should not be confused with the actual behavioral effects of tax policy as defined in the main text of the paper. As long as the budget curve is locally nonlinear, the behavioral responses that are suggested by virtual elasticities are simply not feasible since only the initial point on the virtual budget line corresponds with the actual budget line. Indeed, Blomquist and Simula (2019) show that confusing the two concepts could lead to significant biases in marginal dead-weight loss estimates. The reason why virtual elasticities are nevertheless so often used is that many popular utility functions feature constant virtual elasticities though not necessarily constant actual elasticities. Thus, conditional on those utility functions being close enough to representing true preferences, they lend themselves more easily to empirical estimation. And, as we show below, once virtual elasticities are estimated, it is straightforward to retrieve the actual elasticities.

So how do the virtual behavioral effects relate to the actual behavioral effects? First note that the actual budget curve is given by $c^i = z^i - T(z^i, \kappa)$. This implies that we can rewrite virtual income as a function of z^i and κ as $V^i = V(z^i, \kappa) = z^i T_z(z^i, \kappa) - T(z^i, \kappa)$. Substituting for this and $T_z^i = T_z(z^i, \kappa)$ into the labor income function yields $z^i = \tilde{z}^i(V^i(z^i, \kappa), T_z^i(z^i, \kappa)) = \tilde{z}^i(z^i T_z(z^i, \kappa) - T(z^i, \kappa), T_z(z^i, \kappa))$. Taking total derivatives with respect to z^i and κ, and rearranging, yields:

$$\left(1 + T_{zz}^i \left(\frac{\partial \tilde{z}^i}{-\partial T_z^i} - z^i \frac{\partial \tilde{z}^i}{\partial V^i}\right)\right) \frac{dz^i}{d\kappa} = -\tau(z^i) \frac{\partial \tilde{z}^i}{\partial V^i} - \tau_z(z^i) \left(\frac{\partial \tilde{z}^i}{-\partial T_z^i} - z^i \frac{\partial \tilde{z}^i}{\partial V^i}\right). \tag{29}$$

Substituting for the virtual behavioral elasticities from Eqs. (26)–(28) yields:

$$\left(1 + \frac{T_{zz}^i z^i}{1 - T_z^i} \tilde{e}_c^i\right) \frac{dz^i}{d\kappa} = -\frac{z^i}{1 - T_z^i} \left(\frac{\tau(z^i)}{z^i} \tilde{\eta}^i + \tau_z(z^i) \tilde{e}_c^i\right). \tag{30}$$

Now, set $\tau(z^i) = 0$ and substitute for the definition of the actual compensated elasticity e_c^i, from Eq. (12), to obtain:

$$\left(1 + \frac{T_{zz}^i z^i}{1 - T_z^i} \tilde{e}_c^i\right) e_c^i = \tilde{e}_c^i. \tag{31}$$

Similarly, set $\tau_z(z^i) = 0$ and substitute for the definition of the actual income effect η^i, from Eq. (13), to obtain:

$$\left(1 + \frac{T_{zz}^i z^i}{1 - T_z^i} \tilde{e}_c^i\right) \eta^i = \tilde{\eta}^i. \tag{32}$$

This proves the statement in footnote 9: with knowledge on the tax schedule, e_c^i and η^i can easily be derived from $\tilde{e}_c^i$ and $\tilde{\eta}^i$ and vice versa.

Empirical estimation of actual and virtual elasticities

So how could one empirically estimate either set of elasticities? First consider the estimation of virtual elasticities, which is also discussed in Gruber and Saez (2002). Note that taking the total derivative of $\tilde{z}(V^i, T_z^i)$ and substituting for the virtual elasticities, I can write:

$$\frac{dz^i}{z^i} = -\tilde{e}_c^i \frac{dT_z^i}{1 - T_z^i} + \tilde{\eta}^i \frac{dV^i - z^i dT_z^i}{z^i(1 - T_z^i)}. \tag{33}$$

Provided that the virtual elasticities are constants, one could substitute (yearly) differences in individuals' income and marginal tax rates for the infinitesimal changes dz^i and dT_z^i. Moreover, notice that $dV^i - z^i dT_z^i = -\tau_z(z^i)d\kappa$, for which one could substitute policy-induced changes in the tax burden for a given labor income.

However, one cannot simply estimate Eq. (33) by regressing changes in income on changes in marginal tax rates and tax burdens. The reason is that the change in marginal tax rates mechanically depends on labor income due to nonlinearities in the tax schedule. Simple estimation of Eq. (33) would therefore lead to an endogeneity bias. To see this clearly, note that the change in marginal tax rates is given by:

$$dT_z^i = T_{zz}^i dz + \tau_z(z^i)d\kappa. \tag{34}$$

From this, we can see that exogenous policy variation in the marginal tax rate, $\tau_z(z^i)d\kappa$, would be an ideal instrument for $dT_z^i/(1 - T_z^i)$. And indeed, such policy variation is typically used for empirical estimation of $\tilde{e}_c^i$—see again Gruber and Saez (2002) for an example. Thus, with exogenous policy variation in marginal tax rates as an instrument and constant virtual elasticities, one could estimate Eq. (33) to obtain unbiased estimates of these elasticities.

Having obtained unbiased estimates of $\tilde{e}_c^i$ and $\tilde{\eta}^i$, one could obtain values of the actual elasticities by use of Eqs. (31)–(32). Alternatively, one could substitute for Eq. (34) and $dV^i - z^i dT_z^i = -\tau_z(z^i)d\kappa$ into Eq. (33) and rearrange to obtain the reduced form regression equation:

$$\frac{\mathrm{d}z^i}{z^i} = -e_c^i \frac{\tau_z(z^i)\mathrm{d}\kappa}{1 - T_z^i} + \eta^i \frac{-\tau_z(z^i)\mathrm{d}\kappa}{z^i(1 - T_z^i)}. \tag{35}$$

Thus, one could obtain unbiased estimates of the actual elasticities by directly regressing changes in income on policy-induced variation in marginal and absolute taxes. However, since actual elasticities are likely to depend on the curvature of the tax schedule, it is difficult to make the case for constant actual elasticities—making direct estimation of Eq. (35) problematic.

References

Bastani, S., Blumkin, T., & Micheletto, L. (2015). Optimal wage redistribution in the presence of adverse selection in the labor market. *Journal of Public Economics, 131*, 41–57.

Bierbrauer, F. J., Boyer, P. C. & Hansen, E. (2022). Pareto-improving tax reforms and the earned income tax credit. *Econometrica* forthcoming.

Blomquist, S., & Micheletto, L. (2006). Optimal redistributive taxation when government's and agents' preferences differ. *Journal of Public Economics, 90*(6), 1215–1233.

Blomquist, S., & Simula, L. (2019). Marginal deadweight loss when the income tax is nonlinear. *Journal of Econometrics, 211*(1), 47–60.

Blundell, R., & MaCurdy, T. (1999). Labor supply: A review of alternative approaches. In O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Vol. 3, pp. 1559–1695). Amsterdam: Elsevier.

Chari, V. V., & Kehoe, P. J. (1999). Optimal fiscal and monetary policy. In B. John (Ed.), *Handbook of macroeconomics* (Vol. 1, pp. 1671–1745). Taylor and Michael Woodford. Amsterdam: Elsevier.

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

Chetty, R. (2015). Behavioral economics and public policy: A pragmatic perspective. *American Economic Review, 105*(5), 1–33.

Chetty, R., Friedman, J. N., Leth-Petersen, S., Nielsen, T. H., & Olsen, T. (2014). Active vs. passive decisions and crowd-out in retirement savings accounts: Evidence from Denmark. *Quarterly Journal of Economics, 129*(3), 1141–1219.

Chetty, R., Looney, A., & Kroft, K. (2009). Salience and taxation: Theory and evidence. *American Economic Review, 99*(4), 1145–1177.

Christiansen, V. (1981). Evaluation of public projects under optimal taxation. *Review of Economic Studies, 48*(3), 447–457.

Christiansen, V. (1984). Which commodity taxes should supplement the income tax? *Journal of Public Economics, 24*(2), 195–220.

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

Diamond, P. A. (1998). Optimal income taxation: An example with a U-shaped pattern of optimal marginal tax rates. *American Economic Review, 88*(1), 83–95.

Diamond, P. A., & Mirrlees, J. A. (1971). Optimal taxation and public production II: Tax rules. *American Economic Review, 61*(3), 261–278.

Dixit, A., & Sandmo, A. (1977). Some simplified formulae for optimal income taxation. *Scandinavian Journal of Economics, 79*(4), 417–423.

Farhi, E., & Gabaix, X. (2020). Optimal taxation with behavioral agents. *American Economic Review, 110*(1), 298–336.

Gerritsen, A. (2016). Optimal taxation when people do not maximize well-being. *Journal of Public Economics, 144*, 122–139.

Gerritsen, A., Jacobs, B., Spiritus, K., & Rusu, A., (2022). Optimal taxation of capital income with heterogeneous rates of return. Mimeo.

Golosov, M., Tsyvinski, A., & Werquin, N., (2014). A variational approach to the analysis of tax systems. NBER Working Paper No. 20780.

Gruber, J., & Saez, E. (2002). The elasticity of taxable income: Evidence and implications. *Journal of Public Economics, 84*(1), 1–32.

Guesnerie, R. (1995). *A contribution to the pure theory of taxation*. Cambridge University Press.

Hausman, D. M. (2011). *Preference, value, choice, and welfare*. New York: Cambridge University Press.

Hendren, N. (2016). The policy elasticity. *Tax Policy and the Economy, 30*(1), 51–89.

Hendren, N. (2020). Measuring economic efficiency using inverse-optimum weights. *Journal of Public Economics, 187*, 104198.

Jacobs, B. (2018). The marginal cost of public funds is one at the optimal tax system. *International Tax and Public Finance, 25*, 883–912.

Jacquet, L., & Lehmann, E. (2021). Optimal income taxation with compositional effects. *Journal of the European Economic Association, 19*(2), 1299–1341.

Jacquet, L., Lehmann, E., & Van der Linden, B. (2013). Optimal redistributive taxation with both extensive and intensive responses. *Journal of Economic Theory, 148*(5), 1770–1805.

Kahneman, D., Wakker, P. P., & Sarin, R. (1997). Back to Bentham? Explorations of experienced utility. *Quarterly Journal of Economics, 112*(2), 375–405.

Kanbur, R., Pirttilä, J., & Tuomala, M. (2006). Non-welfarist optimal taxation and behavioural public economics. *Journal of Economic Surveys, 20*(5), 849–868.

Kleven, H. J., & Kreiner, C. T. (2006). The marginal cost of public funds: Hours of work versus labor force participation. *Journal of Public Economics, 90*(10), 1955–1973.

Liebman, J. B., & Zeckhauser, R. J. (2004). Schmeduling. Mimeo.

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

Loewenstein, G., O'Donoghue, T., & Rabin, M. (2003). Projection bias in predicting future utility. *Quarterly Journal of Economics, 118*(4), 1209–1248.

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

Mirrlees, J. A. (1976). Optimal tax theory: A synthesis. *Journal of Public Economics, 6*(4), 327–358.

Piketty, T., & Saez, E. (2013). Optimal labor income taxation. In: A. J. Auerbach, R. Chetty, M. Feldstein, and E. Saez (Ed.) *Handbook of public economics*, (pp. 391–474). vol. 5, Amsterdam: Elsevier.

Piketty, T., Saez, E., & Stantcheva, S. (2014). Optimal taxation of top labor incomes: A tale of three elasticities. *American Economic Journal: Economic Policy, 6*(1), 230–271.

Rothschild, C., & Scheuer, F. (2016). Optimal taxation with rent-seeking. *Review of Economic Studies, 83*(3), 1225–1262.

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

Saez, E. (2002). Optimal income transfer programs: Intensive versus extensive labor supply responses. *Quarterly Journal of Economics, 117*(3), 1039–1073.

Saez, E. (2010). Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy, 2*(3), 180–212.

Saez, E., Slemrod, J., & Giertz, S. H. (2012). The elasticity of taxable income with respect to marginal tax rates: A critical review. *Journal of Economic Literature, 50*(1), 3–50.

Seade, J. (1980). Optimal non-linear policies for non-utilitarian motives. In D. A. Collard, R. Lecomber, & M. Slater (Eds.), *Income distribution: The limits to redistribution*. Bristol: Scientechnica.

Sheshinski, E. (1972). The optimal linear income-tax. *Review of Economic Studies, 39*(3), 297–302.

Spiritus, K., Lehmann, E., Renes, S., & Zoutman, F. T. (2022). Optimal taxation with multiple incomes and types. Tinbergen Institute Discussion Paper 2022-006/VI.

Stantcheva, S. (2014). Optimal income taxation with adverse selection in the labour market. *Review of Economic Studies, 81*(3), 1296–1329.

Stiglitz, J. E. (1982). Self-selection and Pareto efficient taxation. *Journal of Public Economics, 17*(2), 213–240.

Tirole, J., & Guesnerie, R. (1981). Tax reform from the gradient projection viewpoint. *Journal of Public Economics, 15*(3), 275–293.

Tuomala, M. (1990). *Optimal income tax and redistribution*. Oxford: Clarendon Press.

Vickrey, W. (1945). Measuring marginal utility by reactions to risk. *Econometrica, 13*(4), 319–333.

Weymark, J. A. (1981). Undominated directions of tax reform. *Journal of Public Economics, 16*(3), 343–369.