Working papers of the Max Planck Institute for Tax Law and Public Finance Research Paper Series serve to disseminate the research results of work in progress prior to publication to encourage the exchange of ideas and academic debate. Inclusion of a paper in the Research Paper Series does not constitute publication and should not limit publication in any other venue. The preprints published by the Max Planck Institute for Tax Law and Public Finance represent the views of the respective author(s) and not of the Institute as a whole. Copyright remains with the author(s).
Optimal nonlinear taxation: 
the dual approach

Aart Gerritsen*

January 2016

The usual method of solving for an optimal nonlinear tax schedule is that of the primal approach – first solving for the optimal allocation, and subsequently determining which tax system decentralizes this allocation. While this method is mathematically rigorous, it lacks intuitive appeal. I propose a different method based on the dual approach – directly solving for the optimal tax system – which is equally rigorous, while being much closer in spirit to actual tax policy. I show that this approach can easily incorporate preference heterogeneity, as well as individual behavior that is not fully consistent with utility maximization. Over and above solving for the optimum, the dual approach allows one to obtain new insights into the welfare effects of small nonlinear tax reforms outside the optimum.

JEL: H21, H23, H24
Keywords: Optimal taxation, dual approach, preference heterogeneity, individual misoptimization, tax reforms

*Max Planck Institute for Tax Law and Public Finance, Department of Public Economics, Marstallplatz 1, 80539 Munich, Germany. Tel.: +49-89-24246-5256; Fax: +49-89-24246-5299; E-mail: aart.gerritsen@tax.mpg.de; Internet: https://sites.google.com/site/aartgerritsen/. I thank Robin Boadway, Pierre Boyer, Bas Jacobs, and Laurent Simula for helpful comments and discussions.
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 applying the primal approach: he first solved for the optimal allocation, subject to resource and incentive compatibility constraints, and then derived the tax system that would implement this allocation. Ever since, this has been the dominant approach in the literature whenever it concerns nonlinear taxation (e.g., Stiglitz, 1982; Tuomala, 1990; Diamond, 1998).

The advantage of applying the primal approach to solve for the optimal tax schedule 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 the primal 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 might be more than an innocuous abstraction; in the worst case, it alienates the applied world of tax policy, as well as students, from the academic discipline of tax design. This would be detrimental on several counts. It could lead practitioners to disregard academic insights, and academics to focus too much on ethereal issues instead of new insights that might be of more practical relevance. In short, it could reduce the practical impact of an academic field whose raison d’être is its potential for practical impact.\footnote{It might not be a coincidence that the study by Saez (2001), in which he eschews the primal approach, 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.}
A more intuitive way of solving for optimal taxes is by directly considering the social welfare effects of changes in taxes – i.e., to apply the dual approach. 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 by heuristically applying the dual approach (e.g., Saez, 2001, 2002; Piketty and Saez, 2013; Jacquet, Lehmann, and Van der Linden, 2013). They consider a small perturbation of the tax schedule and heuristically 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 on the basis of the primal approach. This last step is necessary as the heuristic lacks the mathematical rigor of the primal approach.

In this paper, I show how one can apply the dual approach to derive the optimal nonlinear income tax without relying on heuristics. By doing so, I combine the intuitive appeal of the dual approach with the mathematical rigor of the primal approach. All that is 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 \( T(z, \kappa) \equiv T(z) + \kappa \tau(z) \). Here, \( \kappa \) is an arbitrary parameter and \( \tau(z) \)

\(^2\)The primal and dual approaches should not be confused with the primal and dual forms of a constrained optimization problem. As is well known from duality theory, the primal form concerns the maximization of utility subject to a budget constraint, whereas the dual form minimizes expenditures subject to a utility constraint. For examples of the dual form of the optimal tax problem, see Boadway and Jacquet (2008) and Lehmann, Simula, and Trannoy (2014). This paper is concerned with the primal and dual approaches to social welfare optimization, which refer to the parameters over which to optimize. Thus, the primal approach refers to optimization with respect to the allocation, whereas the dual approach refers to optimization with respect to taxes.
is the schedule of any non-linear 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 \( \kappa \), and substituting for the specific reform of interest \( \tau(z) \). Expressions for the optimal nonlinear tax schedule are derived by optimizing over \( \kappa \) 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 (2015) who apply the primal approach to show that standard optimal tax formulae 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 formulae include a corrective term, prescribing higher marginal taxes for individuals who mistakenly work too much and lower marginal taxes for individuals who mistakenly 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. And, perhaps more important, determining the desirability of a reform is empirically much less demanding than determining the optimal tax schedule. The reason for this

\(^3\)A similar idea is put forward by Seade (1980), Blomquist and Micheletto (2006), and Kanbur, Pirtilä, and Tuomala (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 (2015) and Farhi and Gabaix (2015).
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 much 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. Applying the dual approach to the welfare analysis of tax reforms can yield important and novel results. For example, when considering raising a tax bracket’s statutory tax rate, I show that the income-weighted average of effective marginal tax rates – rather than a simple average – crucially determines the distortive costs of the reform. This finding contradicts the way in which empirical studies typically determine the marginal distortive costs of taxation.

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, Tsyvinski, and Werquin (2014) also formalize the dual approach to optimal nonlinear income taxation in a mathematically rigorous way. Contrary to the current study, they concentrate on a dynamic model in which individuals always maximize utility. Rather than directly parameterizing the nonlinear tax schedule, they rely on Gateaux derivatives with respect to the tax schedule to obtain the welfare effects of a nonlinear tax reform. Finally, as I use behavioral elasticities to avoid explicitly modeling individual behavior, the current study closely relates to the literature on sufficient statistics (e.g., Chetty, 2009).

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

2 The welfare effects of a tax reform

Studies in optimal taxation typically begin by introducing a model of individual decision making, and then continue by deriving the tax system that maximizes a social-welfare function within the context of that particular model. However, many of the insights from optimal taxation can be obtained without specifying an underlying model of individual behavior. This is not to say that individual behavior is irrelevant – indeed, the behavioral responses of individuals’ taxable income to a change in taxation are a crucial determinant of optimal taxes. But these responses can be captured directly by measurable elasticities that function as sufficient statistics, obviating the need to microfound behavior. I therefore proceed by directly considering the social planner’s optimization problem. I start by introducing the parametrization of the otherwise standard nonlinear income tax, which subsequently allows me to determine the effects of taxation on government revenue and social welfare.

2.1 Individual taxes and government revenue

I assume that individuals in the economy constitute a continuum \( I \) of unit mass, and that an individual \( i \in I \) earns taxable income \( z_i \). I furthermore assume that \( \{z^i : i \in I\} \) is a closed interval so that it is integrable over the population \( I \), 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 \( \kappa \):

\[
T_i \equiv T(z^i, \kappa) = T(z^i) + \kappa \tau(z^i),
\]

which is assumed to be twice differentiable in \( z^i \). I refer to \( \kappa \) 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 $T(z')$ is determined to ensure that $T(z', \kappa)$ gives the actual tax schedule around which a reform is analyzed. 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.

In the analysis below, I assume that $z^i$ is differentiable in $\kappa$. 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. I moreover assume that the derivative of $z^i$ is integrable over the population $\mathcal{I}$.\footnote{Jacquet and Lehmann (2015) identify sufficient conditions on structural parameters for this to hold within the context of a multidimensionally heterogeneous population. In Section 5, I discuss how the dual approach can easily take into account an extensive behavioral margin.} The effect of a reform on an individual’s tax burden is obtained by taking the total derivative of eq. (1):

$$
\frac{dT^i}{d\kappa} = \tau(z') + T^i_z \cdot \frac{dz^i}{d\kappa},
$$

where a subscript denotes a partial derivative, such that $T^i_z \equiv \partial T(z', \kappa)/\partial z^i$ gives the marginal tax rate of an individual with income $z^i$. Since a change in the tax schedule typically affects the tax base itself, an individual’s income tax is affected both directly by the reform of the tax schedule (first term) and indirectly by a change in income (second term). 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 $\kappa$:

$$
\frac{dT^i_z}{d\kappa} = \tau_z(z') + T^i_{zz} \cdot \frac{dz^i}{d\kappa}.
$$

The first term illustrates that the reform raises the marginal tax rate at income level $z'$ by $\tau_z(z')d\kappa$. A reform-induced change in individual $i$’s tax-
able 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).

The government’s budget equals the simple integral of all individuals’ income taxes and is given by:

$$\mathcal{B} \equiv \int_I T^i(z^i, \kappa) d\kappa.$$  

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):

$$\frac{d\mathcal{B}}{d\kappa} = \int_I \left( \tau(z^i) + T^i \cdot \frac{dz^i}{d\kappa} \right) d\kappa,$$

which is simply the integral of eq. (2). The government revenue effects of a tax reform can be decomposed into a mechanical effect and a behavioral effect on the tax base. The mechanical effect simply indicates that the reform raises an amount of resources $\tau(z^i)d\kappa$ from every individual $i \in I$. But a tax reform also tends to affect individuals’ taxable income, leading to a change in tax revenue of $T^i_i dz^i$.

### 2.2 Individual utility and social welfare

A benevolent social planner cares not only about revenue but also about the utility of its citizens. Utility in this context refers to the individual’s actually experienced utility, i.e., his actual well-being.\(^5\) The utility of individual $i$ is a function of his net-of-tax income and his gross-of-tax income, denoted by $u^i(z^i - T^i, z^i)$. For a given gross income, higher net income allows the individual to consume more and thus tends to 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

---

\(^5\)Kahneman, Wakker, and Sarin (1997) distinguish between decision utility and experienced utility. The former is whatever rationalizes individual behavior; the latter is his experienced well-being.
utility. As the income tax is itself a function of gross income and the reform parameter, we can write utility as the following function:

\[(6) \quad U^i \equiv U^i(z^i, \kappa) = u^i(z^i - T^i(z^i, \kappa), z^i).\]

As with taxable income, I assume that utility and its derivatives are integrable over the population \(I\).

For now, I remain agnostic about how individuals decide on their taxable income. They might or might not maximize their utility.\(^6\) I define \(\omega_i\) as the degree to which individual \(i\) mistakenly chooses too large a gross income. More specifically, it gives the marginal utility of reducing one’s gross income, normalized in terms of consumption:

\[(7) \quad \omega^i \equiv \frac{-U^i_z}{u^i_c} = \left(\frac{-u^i_z}{u^i_c} - (1 - T^i_z)\right),\]

where \(u^i_c\) refers to the partial derivative of utility with respect to net income. I allow \(\omega^i\) to vary across individuals, even if they earn the same income \(z^i\). If individual \(i\) maximizes his utility, his marginal rate of substitution of net income for gross income (first term within brackets) must equal his marginal net-of-tax rate (second term within brackets). In that case \(\omega^i = 0\). If an individual mistakenly earns too much, then \(\omega^i > 0\); and if he mistakenly earns too little, then \(\omega^i < 0\). More generally, individual \(i\)’s taxable income would have been his utility-maximizing income level, had his marginal net-of-tax rate been \(\omega^i\) percentage points higher.

I assume that the social objective is welfarist. This implies I can write the social welfare function as a (weighted) integral of all individuals’ utility:

\[(8) \quad \mathcal{W} \equiv \int_I \gamma^i U^i \, dI,\]

---

\(^6\)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, Looney, and Kroft, 2009; Lieberman and Zeckhauser, 2004) or about their own utility function (e.g., Loewenstein, O’Donoghue, and Rabin, 2003). Another reason might be that the tax base is not fully under control of the individual. For example, in Piketty, Saez, and Stantcheva (2014) and Rothschild and Scheuer (2014) the tax base is partly the result of third party bargaining or rent-seeking efforts. Naturally, in that case \(U^i(z, \kappa)\) cannot be a full characterization of the individual’s utility, as it should depend on, e.g., his bargaining effort, as well as gross income.
where $\gamma^i$ is an individual-specific weight that determines the importance of individual $i$’s utility within the social objective. In the special (though intuitively appealing) case of a utilitarian social objective, $\gamma^i = \gamma$ for all $i$. The effect of a tax reform on the social objective is obtained by taking the derivative of eq. (8) with respect to $\kappa$. Doing so, while substituting for eq. (7), yields:

$$
\frac{dW}{d\kappa} = - \int_I \gamma^i u^i_c \left( \tau(z^i) + \omega^i \cdot \frac{dz^i}{d\kappa} \right) \, di.
$$

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 $\tau(z^i)d\kappa$. The second term within brackets represents the reform’s behavioral effect on social welfare. If the reform causes individuals to increase their gross income ($dz^i/d\kappa > 0$), 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 ($dz^i/d\kappa < 0$). Naturally, if individuals choose their tax base to maximize utility ($\omega^i = 0$), a reform only affects social welfare through its mechanical effect.

### 2.3 Elasticities

Before I consider the net social-welfare effect of a tax reform, by aggregating its effects on government revenue and social welfare, it is useful to elaborate on how the tax base is affected by a tax reform. This of course depends on the nature of the reform – i.e., on how the reform affects the tax schedule. Specifically, it is typically observed that changes in marginal tax rates and changes in the absolute tax burden affect taxable income in different ways (cf., Blundell and MaCurdy, 1999; Saez, Slemrod, and Giertz, 2012). I capture this by decomposing the effects of a reform on taxable income into a substitution effect and an income effect.

Recall that a reform raises the marginal tax rate at income level $z$ by $\tau_z(z)d\kappa$, and raises the absolute tax burden by $\tau(z)d\kappa$. I characterize the
substitution effect of a reform on individual $i$’s taxable income by reference to his compensated net-of-tax rate elasticity of taxable income, $e^c_i$. It gives the relative change in his taxable income, $dz^i/z^i$, due to a relative change in his marginal net-of-tax rate, $-\tau(z^i)dz^i/(1-T_i^z)$, for a constant absolute tax burden, $\tau(z^i) = 0$. Hence, I can write:

$$e^c_i \equiv \left. \frac{1 - T_i^z}{z^i} \cdot \frac{dz^i}{-\tau(z^i)dz^i} \right|_{\tau(z^i)=0}.$$  

An increase in the marginal net-of-tax rate typically causes an increase in taxable income, such that $e^c_i > 0$. However, I allow $e^c_i$ to vary across individuals, even among those with the same taxable income.

Before characterizing a reform’s income effect, I first define the uncompensated net-of-tax rate elasticity of taxable income, $e^u_i$. The uncompensated elasticity also gives the relative change in taxable income due to a relative change in the marginal tax rate – but now for an equal increase in the average tax rate: $\tau(z^i)/z^i = \tau(z^i)$. Hence, I can write:

$$e^u_i \equiv \left. \frac{1 - T_i^z}{z^i} \cdot \frac{dz^i}{-\tau(z^i)dz^i} \right|_{\tau(z^i)/z^i=\tau(z^i)}.$$  

Notice that the uncompensated elasticity represents both substitution and income effects. Indeed, I obtain a measure of a reform’s income effect by subtracting the compensated elasticity from the uncompensated elasticity. This yields:

$$\eta^i \equiv e^u_i - e^c_i = \left. (1 - T_i^z) \cdot \frac{dz^i}{-\tau(z^i)dz^i} \right|_{\tau(z^i)=0}.$$  

Thus, $\eta^i$ measures the effect on taxable income of a reform that only raises the absolute tax burden but leaves the marginal tax rate unchanged. A lower tax burden normally causes a reduction in taxable income, such that $\eta^i < 0$. However, as with the compensated elasticity, I allow $\eta^i$ to vary across individuals.

Notice that the compensated elasticity and the income effect are defined as relative changes in income along the actual budget curve – as in Jacquet, Lehmann, and Van der Linden (2013) – and not as changes along
a linearized ‘virtual’ budget line – as in Saez (2001). That is, $e_i^c$ and $\eta_i^i$ 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.\footnote{In the Appendix, I show that the two different behavioral concepts are closely related. More specifically, if $\tilde{e}_i^c$ and $\tilde{\eta}_i^i$ denote the virtual compensated elasticity and income effect defined along a linearized virtual budget line, then we can write:

\[
\left(1 + \frac{T_i(z_i)z_i^i}{1 - T_i} \tilde{e}_i^c\right) e_i^c = \tilde{e}_i^c \\
\left(1 + \frac{T_i(z_i)z_i^i}{1 - T_i} \tilde{e}_i^c\right) \eta_i^i = \tilde{\eta}_i^i
\]

Thus, with knowledge of the tax schedule, one can easily derive one pair of behavioral effects from the other. In the Appendix, I furthermore show that either elasticity concept could be empirically estimated by use of exogenous policy variation in the tax system. Specifically, $\tilde{e}_i^c$ would follow from using policy variation as an instrument for marginal tax rates, whereas $e_i^c$ would follow from a reduced-form regression of income on the policy variation itself.

\footnote{To see this, notice that if $z^i$ is only affected by changes in the marginal tax rate ($\tau_i(z^i)dz^i$) and changes in the absolute tax burden ($\tau(z^i)dz^i$) the following must hold:

\[
\frac{dz^i}{d\kappa} = \frac{dz^i}{d\kappa} \bigg|_{\tau(z^i)=0} + \frac{dz^i}{d\kappa} \bigg|_{\tau(z^i)=0}.
\]

Substituting for eqs. (10) and (12) yields eq. (13).}

Armed with the above elasticity concepts, the effect of a tax reform on individual $i$’s taxable income can now be decomposed into income and substitution effects. Provided that a tax reform only affects a person’s gross income through its effects on his marginal tax rate and absolute tax burden, the definitions in eqs. (10) and (12) allow me to write:\footnote{To see this, notice that if $z^i$ is only affected by changes in the marginal tax rate ($\tau_i(z^i)dz^i$) and changes in the absolute tax burden ($\tau(z^i)dz^i$) the following must hold:

\[
\frac{dz^i}{d\kappa} = \frac{dz^i}{d\kappa} \bigg|_{\tau(z^i)=0} + \frac{dz^i}{d\kappa} \bigg|_{\tau(z^i)=0}.
\]

Substituting for eqs. (10) and (12) yields eq. (13).}

\[
\frac{dz^i}{d\kappa} = -\frac{z^i}{1 - T_i} \left( e_i^c \tau_i(z^i) + \eta_i^i \cdot \frac{\tau_i(z^i)}{z^i} \right).
\]

Eq. (13) is an accounting identity that allows one to differentiate changes in taxable income due to reforms of the marginal tax rate from changes in taxable income due to reforms of the average tax rate. A reform that raises the marginal tax rate ($\tau_i(z^i) > 0$) leads to a reduction in taxable income proportional to $e_i^c$. A reform that raises the average tax rate ($\tau_i(z^i)/z_i > 0$)
leads to an increase in taxable income proportional to $-\eta^i$.

2.4 Net social welfare effects of a reform

The net social welfare effects of a tax reform are obtained by aggregating its effects on social welfare and the government budget. For this, I denote the social marginal value of public resources by $\lambda$. Moreover, I denote the social marginal value of individual $i$’s consumption by $g^i \equiv \gamma^i u^i_e / \lambda$, where $g^i$ is expressed in terms of public resources. This allows me to formulate the following proposition.

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

$$
(14) \int_Z \left( (1 - g^i) \tau(z^i) - \frac{T^i_z - g^i \omega^i}{1 - T^i_z} \cdot \frac{\eta^i \tau(z^i)}{1 - T^i_z} \cdot z^i e^i_c \tau(z^i) \right) \, d_i.
$$

**Proof.** The net social welfare effect of a reform is given by $\frac{dW}{\lambda} + \frac{dG}{\lambda}$. Substituting for eqs. (5), (9), and (13) yields eq. (14). ■

A tax reform can be seen to have three effects on social welfare, illustrated in expression (14) by the three terms within large brackets. The first term gives the mechanical effects of a tax reform. The reform mechanically raises $\tau(z^i)$ in tax revenue from individuals with income $z^i$, simultaneously causing a social utility loss of $g^i \tau(z^i)$. 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 gross income. This leads to an increase in tax revenue as long as the marginal income tax is positive ($T^i_z > 0$). It also leads to a reduction in utility if individual $i$ is earning more than what is good for him ($\omega^i > 0$), or to an increase in utility if he is earning less than what is good for him ($\omega^i < 0$). The third term within large brackets 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 $e^i_c > 0$. This reduction leads to tax revenue losses (as long as $T^i_z > 0$) and to utility gains (if $\omega^i > 0$) or utility losses (if $\omega^i < 0$).
Proposition 1 and expression (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 Section 4, I further elaborate on this.

3 Optimal taxation

3.1 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 \( i \).

Substituting this into expression (14), while equating it to zero, yields:

\[
\int_{\mathcal{Z}} \left( 1 - g^i - \frac{T^i}{1 - T^i} \cdot \eta^i \right) \, dx = 0.
\]

As the reform leaves all marginal tax rates unchanged, it does not generate any substitution effect, and only affects social welfare through mechanical and 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:

\[
\alpha^i \equiv g^i + \frac{T^i}{1 - T^i} \cdot \omega^i \times \eta^i.
\]
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 $T^i_z\eta^i/(1-T^i_z)$, and a further social utility effect equal to $-g^i\omega^i\eta^i/(1-T^i_z)$. Taken together, $\alpha^i$ indicates how many resources government is willing to give up in order to provide individual $i$ with an additional unit of income.\footnote{It also corresponds to what Diamond (1975) called the social marginal utility of individual income, divided by the social marginal value of public resources $\lambda$.} The pattern of $\alpha^i$ 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:

\begin{equation}
\int_\mathcal{I} \alpha^i \, di = 1.
\end{equation}

**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 sweeping consequences for public policy. As documented by Jacobs (2013), it implies that the marginal cost of public funds – defined as the inverse of the left-hand side of eq. (17) – equals one. 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).

The same argument holds for revenue-generating public policy: distortive taxes could always be reduced with the revenue from a nondistortive reform as the one considered here. The existence of such distortive taxes should therefore be irrelevant for the valuation of a policy’s revenue.
gains. As a result, optimal environmental levies correspond to standard Pigouvian levies (Jacobs and de Mooij, 2015; Sandmo, 1975), using public debt to smooth the tax burden over time is not necessary (Werning, 2007), and a positive inflation tax cannot be justified on the basis of preexisting tax distortions (Da Costa and Werning, 2008). Generally, Proposition 2 implies that neither of the following two statements can serve as a valid justification of an adjustment to public policy: (i) the net financing costs of the policy lead to higher distortive taxes, or (ii) the net revenue gains of the policy can be used to reduce distortive taxes.10

3.2 Reform 2: Raising marginal income taxes

The second reform I consider raises the tax burden by $d\kappa$ for individuals who earn an income that is larger than some level $z^*$. Thus, $\tau(z^*) = 1$ for all individuals with $z^i > z^*$, and $\tau(z^i) = 0$ otherwise. As a result, only the marginal tax rate for individuals with income $z^*$ is raised ($\tau_z(z^*) > 0$). To determine the effect of this reform on the marginal tax rate, notice that the change in the tax burden at $z^*$ equals $\tau(z^*) d\kappa = 0$, whereas the change in tax burden at $z^* + dz$ equals $\tau(z^* + dz) d\kappa = d\kappa$. By the definition of the derivative, the change in the marginal tax rate at $z^*$ is given by $\tau_z(z^*) d\kappa \equiv \left( \frac{\tau(z^* + dz) - \tau(z^*)}{dz} \right) d\kappa = \frac{d\kappa}{dz}$. Substituting this into expression (14), while equating it to zero, yields:

$$\int_{I: z^i > z^*} \left( 1 - g^i - \frac{T^i}{1 - T^i} \cdot \omega^i \cdot \eta^i \right) \, di = \int_{I: z^i = z^*} \left( \frac{T^i}{1 - T^i} \cdot z^i \cdot e^i \right) \, dz.$$

The left-hand side simply gives the difference between the social marginal value of public resources and the social marginal value of private resources for the subset of the population that earns more than $z^*$. From eqs. (15) and (17), we know that this difference equals zero for the total population. Hence, if the social marginal utility of income is decreasing with income – e.g., because the social objective exhibits egalitarian preferences – the

\[\text{Note that I only refer to the financing costs or benefits of public policy. If the policy itself alleviates or exacerbates existing distortions (e.g., through relative complementarity with the tax base) or if it yields a degree of redistribution that is superior to a nonlinear tax schedule (e.g., if the policy’s effects on utility are correlated with individuals’ innate ability), the resulting welfare effects should naturally enter the cost-benefit analysis.}\]
left-hand side of eq. (18) is strictly positive for all but the lowest level of income $z^*$, and can be seen as the redistributive benefits of the reform.

The right-hand side gives the social marginal costs associated with distorting the tax base of individuals with gross income $z^*$. As long as the marginal tax rate and the compensated elasticity at that income level are positive, the tax base erosion due to the increase in the marginal tax rate diminishes government revenue and therefore total welfare. On top of this revenue effect, the reduction in the tax base leads to welfare gains if individuals with gross income $z^*$ tend to earn more than what is good for them ($\omega^i > 0$), and to welfare losses if they earn less than what is good for them ($\omega^i < 0$). The importance of this correction effect of marginal taxes is increasing with individual welfare weights $g^i$. Intuitively, the more a government cares about an individual, the more important it is to raise the individual’s utility by correcting his behavior.

3.2.1 ...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 perfectly choose the tax base to maximize their utility. In that case, $\omega^i = 0$ for all individuals $i \in I$. Note that one can write the cumulative distribution function of taxable income as $H(z) \equiv \int_{I: z_i \leq z} di$, which gives the proportion of individuals that have gross income equal to or below $z$. It follows that $dH(z) = h(z)dz = \int_{I: z^*_i = z} di$. This allows me to define the average compensated elasticity for individuals with income level $z^*$ as $\bar{e}_{c}^*$:

\begin{equation}
\bar{e}_{c}^* \equiv \frac{\int_{I: z^*_i = z^*} e^i_c di}{\int_{I: z^*_i = z^*} di} = \frac{\int_{I: z^*_i = z^*} e^i_c di}{h(z^*)dz^*}
\end{equation}

Moreover, I define the average social marginal value of private resources of individuals who earn more than $z$ as $\bar{\alpha}_{z^*_i > z} \equiv \int_{I: z^*_i > z} \alpha^i di / \int_{I: z^*_i > z} di$. 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:

\[
T^*_z = \frac{1}{1-T^*_z} \cdot \frac{1-H(z^*)}{z^* h(z^*)} \cdot \left(1 - \bar{\alpha}_{z^* > z^*}\right).
\]

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

Eq. (20) is virtually identical to the standard optimality condition in Saez (2001) or Piketty and Saez (2013) with two minor adjustments. First, contrary to the standard formulation, I defined elasticities as moves along the actual budget curve rather than moves along a hypothetical linearized budget line. This allows me to write eq. (20) in terms of the actual income density rather than a ‘virtual’ income density that would arise if individuals’ nonlinear tax schedule is replaced by a linearized tax. Second, the average elasticity in eq. (20) takes into account that behavioral responses to marginal tax changes might differ across individuals with income \(z^i\). Both adjustments can also be found in Jacquet and Lehmann (2015) who derive eq. (20) by means of the primal approach.

Proposition 3 shows that the optimal marginal tax rate at income level \(z^*\) crucially depends on three terms that indicate the responsiveness of the tax base at income level \(z^*\), the hazard rate of the income distribution at income level \(z^*\), and the redistributive effects of the marginal tax rate at income level \(z^*\). First, the optimal marginal tax for individuals with income \(z^*\) is decreasing in the average compensated elasticity of individuals with income \(z^*\). Intuitively, higher elasticities imply that the tax base at that income level is more responsive to marginal tax rates. As a result, the social marginal costs of tax-base erosion are larger, yielding lower marginal taxes in the optimum.

Second, the optimal marginal tax rate at \(z^*\) is decreasing in the density of taxable income at that income level, \(z^* h(z^*)\), and increasing in the share of individuals with a higher income, \(1 - H(z^*)\). Intuitively, the marginal income tax distorts the total tax base at income level \(z^*\). The larger this total tax base, and thus the larger \(z^* h(z^*)\), the larger the distortion caused by the marginal tax and the smaller the optimal tax rate. Furthermore, the marginal tax rate raises revenue from every individual with income above
The more people with income above $z^*$, and thus the larger $1 - H(z^*)$, the higher the amount of revenue raised and the larger the optimal tax rate.

Third and final, the optimal marginal tax rate at $z^*$ is decreasing in the average social marginal value of private resources in the hands of individuals with income above $z^*$. This can be seen from the bracketed term in eq. (20). It gives the social marginal gains minus costs of raising one unit of public resources from individuals who earn more than $z^*$. Expressed in terms of government revenue, the social marginal gains simply equal 1. The social marginal costs are given by the average social marginal value of private resources ($\alpha_i$) of individuals who earn more than $z^*$. Intuitively, the larger this bracketed third term, the more valuable are public resources compared to private resources in the hands of relatively rich individuals. And since the marginal tax at $z^*$ redistributes away from those individuals towards the government, the higher is the optimal marginal tax rate.

3.2.2 ...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 $\omega_i^j$ might be nonzero. Before deriving the optimal tax formula, it is useful to define the income-conditional covariance between two variables as $\chi(x_i^j, y_i^j) \equiv \bar{x}_i^j \bar{y}_i^j - \bar{x}_i^j \bar{y}_i^j$, where an overline indicates average values conditional on labor income $z_i^j$. This allows me to formulate 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:

$$
\frac{T_z^* - \bar{g}_w^* \bar{\omega}_w^* - \chi(g_w^*, \omega_w^*) - \chi\left(g_w^* \omega_w^*, \frac{\alpha_{z^*}}{\bar{z}_{\bar{c}_c}^*} \cdot \left(1 - \bar{\alpha}_{z^* > z^*}^* \right) \right)}{1 - T_z^*} = \frac{1}{\bar{c}_c^*} \cdot \frac{1 - H(z^*)}{z^* h(z^*)} \cdot \left(1 - \bar{\alpha}_{z^* > z^*}^* \right).
$$

Only in the special case that the social marginal value of consumption ($g_i^j$), the degree of misoptimization ($\omega_i^j$), and the compensated elasticity of taxable income ($e_i^j$) are uncorrelated for individuals with the same income, this
reduces to:

\[ T^*_z - g^* \tilde{\omega}^* \frac{1}{1 - T^*_z} = 1 \frac{1 - H(z^*)}{z^* h(z^*)} \cdot \left( 1 - \bar{\alpha}_{z^*>z^*} \right). \]

**Proof.** Substituting eq. (19) and the definitions of the income distribution, income-conditional correlations, and the average social value of private resources into eq. (18) and rearranging yields eq. (21). If \( g^i, \omega^i \) and \( e^i_c \) are uncorrelated for individuals with taxable income \( z^* \), then \( \chi(g^*, \omega^*) = \chi(g^* \omega^*, e^*_c) = 0 \), and eq. (21) reduces to eq. (22).

Consider the case in which income-conditional covariances between \( g^i, \omega^i \), and \( e^i_c \) are nil. The first thing to notice is that the right-hand side of eq. (22) is virtually identical to the right-hand side of the optimality condition for utility-maximizing individuals in eq. (20). The elasticity term and the distribution term are exactly identical. There is a small adjustment implicit in the social marginal value of private resources \( (\alpha^i) \), as it now incorporates the fact that any behavioral income effects could affect the utility of individuals who earn more than \( z^* \). However, it is important to keep in mind that all the right-hand side terms are endogenous variables and therefore not likely to be independent from how individuals choose their labor income. Nevertheless, as long as we can measure these variables empirically, it is possible to evaluate existing tax systems without reference to the deeper model parameters that determine individuals’ decision making.\(^{11}\)

The most striking adjustment, however, is in the left-hand side of eq. (22). This term measures the social marginal costs associated with a compensated reduction in the tax base of an individual with income \( z^* \). With utility-maximizing individuals, this term only depends on the marginal tax rate: the higher the marginal tax rate, the larger the revenue losses from a reduction in the tax base. When individuals fail to maximize their utility, there is an offsetting welfare gain if individuals with income \( z^* \) on average chose their income too high \( (\tilde{\omega}^* > 0) \), or an even larger welfare loss if they

\(^{11}\)Naturally, since \( \alpha^i \) is to an important extent driven by social preferences for redistribution (e.g., by the Pareto weights \( \gamma^i \)), one cannot measure it empirically. However, by imposing the weak moral restriction that Pareto weights must be weakly positive \( (\gamma^i \geq 0) \), one could use the optimality condition to evaluate whether existing tax systems indeed satisfy this Pareto criterion.
on average chose to earn too little income ($\bar{\omega}^* < 0$). Recall from eq. (7) that $\bar{\omega}^*$ measures the average monetized marginal utility of reducing the gross income of individuals who earn $z^*$. Thus, $\bar{g}^*\bar{\omega}^*$ measures its social value. Generally, the larger the extent to which individuals with income $z^*$ choose their income too high (too low), the higher (lower) the optimal marginal tax rate at that income level. Intuitively, marginal tax rates are not only used to redistribute income away from higher-income individuals, but also to ‘correct’ individuals’ behavior.

It is important to note that $\bar{\omega}^*$ in eq. (22) refers to its value at the optimal tax system. This is problematic because even if one would know its value at the actual tax system, it would not necessarily be informative about its value at the optimum. The reason for this is that without imposing more structure on individual decision making, it is unclear whether a higher marginal tax rate increases or decreases the degree to which individuals mistakenly choose their income. On the one hand, individuals tend to reduce their taxable income in response to higher marginal taxes, thereby reducing the degree to which they choose to earn too much income. On the other hand, an increase in the marginal tax rate also reduces the utility-maximizing level of taxable income, thereby raising the degree to which individuals earn too much income.

There are two ways out of this conundrum. The first is to microfound the value of $\bar{\omega}^*$ by adopting one of many existing models of suboptimal individual decision making. For example, one could assume that individuals actually do try to maximize their utility but mistake marginal and average tax rates (as in Liebman and Zeckhauser, 2004), or that individuals have certain mistaken beliefs about their utility function (as in Loewenstein, O’Donoghue, and Rabin, 2003). Writing $\bar{\omega}^*$ in terms of the model’s parameters and structurally estimating these could then enable one to quantify the optimal tax schedule. However, to the best of my knowledge there is currently no consensus on what is the best alternative to the theory of util-

\[\frac{u_i}{u_i z} = 1 - T_i/z \]

Substituting this into eq. (7) yields $\omega = T_i - T_i/z$. However, note that it also implies that $c_i^e = 0$, which is refuted by empirical evidence (Saez, Slemrod, and Giertz, 2012). Thus, a model in which individuals are only incentivized by average tax rates rather than marginal tax rates is probably not realistic.
ity maximization within the context of individuals’ income decisions.\textsuperscript{13} An alternative approach is to empirically determine values of $\omega^i$ and its pattern across the income distribution at the existing tax system, and use this to indicate the direction in which tax rates should be adjusted to correct individual behavior. While such approach is not necessarily informative about the tax optimum, it could potentially provide information on the welfare implications of small reforms of the existing tax system.\textsuperscript{14}

To derive eq. (22), I assumed that $\omega^i$, $g^i$ and $e^i_c$ are uncorrelated across individuals with the same income. However, it might well be that government attaches a higher welfare weight to ‘hard-working’ individuals so that $g^i$ is increasing with $\omega^i$ and thus $\chi(g^i, \omega^i) > 0$. In that case, government is more interested in ‘correcting’ the behavior of high-$\omega^i$ individuals than that of low-$\omega^i$ individuals. As can be seen from eq. (21), this leads to higher marginal tax rates at the optimum. Similarly, it could be that individuals with larger deviations from utility maximization are less responsive to changes in tax rates.\textsuperscript{15} This would imply that the degree of misoptimization is negatively correlated with behavioral elasticities ($\chi(g^i\omega^i, e^i_c) < 0$) when individuals with income $z^i$ mistakenly earn too much on average. Conversely, this correlation would be positive ($\chi(g^i\omega^i, e^i_c) > 0$) when they earn too little on average. As can be seen from eq. (21), the corrective argument for taxation becomes weaker as a result, bringing optimal tax rates closer to the ones obtained with utility-maximizing individuals. In the extreme case in which only utility-maximizing individuals are responsive to taxation, tax rates cease to have a corrective function at all and optimal tax rates are once more given by eq. (20).

\textsuperscript{13}Though see Rees-Jones and Taubinsky (2016), who designed a survey experiment to structurally quantify the extent to which individuals confuse average and marginal tax rates among other potential perception biases regarding the income tax.

\textsuperscript{14}In Gerritsen (2015), I attempt to measure $\omega^i$ on the basis of British life-satisfaction data, which leads me to conclude that people at the bottom of the income distribution tend to work too little and people at the top of the income distribution tend to work too much. In order to correct individuals’ behavior, this would call for lower marginal tax rates at the bottom and higher marginal tax rates at the top of the income distribution. These findings are in line with those of Rees-Jones and Taubinsky (2016) who conclude from their survey experiment that individuals overestimate marginal taxes at low incomes and underestimate marginal taxes at high incomes.

\textsuperscript{15}Chetty et al. (2014) make an argument to this effect within the context of subsidies for retirement savings, see also Chetty (2015).
3.2.3 Asymptotic results

Proposition 4 also allows one to obtain results for the optimal tax rate at the top of the income distribution. For this, I assume that the top of the income distribution is well-described by a Pareto distribution with parameter $p$, such that $\frac{1-H(\hat{z})}{\hat{z}h(\hat{z})} = \frac{1}{p}$, with $\hat{z}$ indicating any income level at the top of the distribution (Saez, 2001; Piketty and Saez, 2013). I furthermore assume that the compensated elasticity $e^t$, the term for the income effect $\eta^t$, the social marginal value of private resources $g^t$, and the degree of misoptimization $\omega^t$ converge to $\hat{e}$, $\hat{\eta}$, $\hat{g}$, and $\hat{\omega}$ for top income earners. Substituting these definitions, as well as the definition of $\alpha^t$, into eq. (21) and rearranging yields the following optimality condition:

\begin{equation}
\frac{\hat{T}_z - \hat{g}\hat{\omega}}{1 - \hat{T}_z} = \frac{1 - \hat{g}}{p\hat{e} + \hat{\eta}},
\end{equation}

where $\hat{T}_z$ corresponds to the top tax rate. The right-hand side of eq. (23) perfectly corresponds with the optimal top tax wedge found by Saez (2001). Thus, top tax rates are decreasing in the compensated elasticities of top earners. They are also decreasing in the Pareto parameter – which measures the thinness of the upper tail of the income distribution and is therefore inversely related to the revenue that can be generated with a tax on top income earners. Moreover, as long as income effects on taxable income are negative ($\hat{\eta} < 0$), the optimal tax rate is increasing in the income responsiveness of top earners’ income.

Contrary to Saez (2001), the optimal tax wedge must take into account the degree of misoptimization by top income earners. This can be seen from the left-hand side of eq. (23). As with marginal tax rates generally, the optimal top tax rate serves both to redistribute income away from top income earners and to correct their behavior. Notice, however, that the corrective argument only plays a role if the welfare weight at the top is strictly positive ($\hat{g} > 0$). If government does not care about the very rich, it has no reason to correct their behavior either. In that case, the optimal tax rate simply equals the revenue-maximizing rate. Interestingly, if top-
income earners are mistakenly earning too much income ($\hat{\omega} > 0$), it might well be that a larger welfare weight for top-income earners leads to higher top tax rates in the optimum – i.e., to tax rates over and beyond the revenue maximizing rate in order to correct top earners’ mistaken behavior.

4 The desirability of limited reforms

4.1 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.

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. 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 ($\omega^i = 0$).\footnote{As noted by Saez, Slemrod, and Giertz (2012), there is little compelling evidence on significant income effects when it concerns taxable income.} 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 modelled 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. Proposition 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:

\[
\int_{I_z: z_i \in [z^a, z^b]} (z^i - z^a)(1 - g^i)\,di + \int_{I_z: z_i > z^b} (z^b - z^a)(1 - g^i)\,di
\]

\[
> \int_{z^a}^{z^b} \frac{T_z}{1 - T_z} \cdot \bar{e}_c \cdot z^i h(z^i) \cdot \,dz^i,
\]

where I substituted for the income density on the right-hand side. The left-hand-side of eq. (24) represents the redistributive 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 \((z^i - z^a)\,dk\), whereas the tax burden of an individual \(i\) above the tax bracket increases by \((z^b - z^a)\,dk\). The total redistributive benefits of the reform generally depend on welfare weights \(g^i\), which ultimately makes desirability a matter of political judgment.\(^\text{18}\)

Whereas the redistributive benefits of the reform importantly depend on political values, we can say more about the distortive costs of the reform, given by the right-hand side of eq. (24). 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 distortive costs do not simply equal the product of these three factors’ averages. As can be seen from eq. (24), it also matters how these factors are correlated. This issue is sidestepped by almost every study that measures the distortive 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 distortive costs indeed reduce to the product of the

\(^{18}\text{The only exception is if eq. (24) is strictly violated even if } g^i = 0 \text{ for all affected individuals. In that case, it is beneficial to lower the tax rate even if government does not care about the individuals who receive the tax cut. In other words, it would indicate that the status quo is a Pareto inefficient tax system with tax rates beyond the top of the Laffer curve.}\)
elasticity, the tax wedge, and the amount of income within the interval.\textsuperscript{19} However, in reality tax schedules are typically highly nonlinear, causing this approach to yield biased estimates of the marginal distortive costs of taxation. Nonlinearities in actual tax schedules stem from means-tested welfare arrangements such as an earned income tax credit, rental support, or child benefits, as well as different tax brackets. The phase-out intervals of means-tested programs typically combine falling marginal tax rates with increasing income concentrations. Eq. (24) then tells us that the distortive costs of a bracket’s tax rate are lower if this bracket overlaps with the phase-out of such welfare arrangements.

5 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...\textsuperscript{19}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 (2015), 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.
or other welfare programs, or changes to a quadratic tax schedule. 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 $\omega^i$ 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 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 furthermore 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

---

20 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.
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:

$$B = \int I(\pi^i(\kappa) T(z^i, \kappa) + (1 - \pi^i(\kappa)) T(0, \kappa)) \, di,$$

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:

$$\frac{dB}{d\kappa} = \int I\left(\pi^i(\kappa) \left(\tau(z^i) + T_s \frac{dz^i}{d\kappa}\right) + (1 - \pi^i(\kappa)) \tau(0) + (T^i - T^0) \frac{d\pi^i}{d\kappa}\right) \, di,$$

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 ($dz^i/d\kappa$), and an extensive behavioral effect on the tax base ($d\pi^i/d\kappa$). 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 fruitfully be 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,
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., 2015). 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.

6 Conclusion

This paper develops a method to solve for the optimal nonlinear income tax based on the dual approach. The procedure is not only intuitive, as it is close in spirit to actual tax policy, but also mathematically rigorous. 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 fellow scholars. I show 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 indicate 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.

References


Jacobs, Bas. 2013. “The marginal cost of public funds is one at the optimal tax system.” Mimeo.


Jacquet, Laurence and Etienne Lehmamn. 2015. “Optimal income taxation when skills and behavioral elasticities are heterogeneous.” Mimeo.


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 \( c^i = (1 - T_z^i)z^i + R^i \), with \( R^i \) 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 \( z^i = \tilde{z}(R^i, T_z^i) \). 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:

\[
\tilde{\varepsilon}_u^i \equiv \frac{1 - T_z^i}{z^i} \frac{\partial \tilde{z}(R^i, T_z^i)}{\partial T_z^i}.
\]

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:

\[
\tilde{\eta}_u^i \equiv (1 - T_z^i) \frac{\partial \tilde{z}(R^i, T_z^i)}{\partial R^i},
\]

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:

\[
\tilde{\eta}_c^i \equiv (1 - T_z^i) \frac{\partial \tilde{z}(R^i, T_z^i)}{\partial T_z^i}.
\]
by:

\[ \hat{e}_c^i \equiv \hat{e}_u^i - \hat{\eta}_c^i = \frac{1 - T_i^i}{z^i} \frac{\partial \hat{z}(R_i^i, T_i^i)}{\partial T_i^i} - (1 - T_i^i) \frac{\partial \hat{z}(R_i^i, T_i^i)}{\partial R}. \]

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 (2015) 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 \(\kappa\) as \(R_i^i = R(z^i, \kappa) = z^iTz_i(z^i, \kappa) - T(z^i, \kappa)\). Substituting for this and \(T_i^i = Tz_i(z^i, \kappa)\) into the labor income function yields \(z^i = \hat{z}(R_i^i(z^i, \kappa), T_i^i(z^i, \kappa)) = \hat{z}(z^iTz_i(z^i, \kappa) - T(z^i, \kappa), Tz_i(z^i, \kappa))\). Taking total derivatives with respect to \(z^i\) and \(\kappa\), and rearranging, yields:

\[ (1 + Tz_i^i \left( \frac{\partial \hat{z}^i}{\partial T_i^i} - z^i \frac{\partial \hat{z}^i}{\partial R_i^i} \right) ) \frac{dz^i}{d\kappa} = -\tau(z^i) \frac{\partial \hat{z}^i}{\partial R_i^i} - \tau(z^i) \left( \frac{\partial \hat{z}^i}{\partial T_i^i} - z^i \frac{\partial \hat{z}^i}{\partial R_i^i} \right). \]

Substituting for the virtual behavioral elasticities from eqs. (27)-(29) yields:

\[ \left( 1 + \frac{Tz_i^i z_c^i}{1 - Tz_i} \right) \frac{dz^i}{d\kappa} = -\frac{z^i}{1 - Tz_i^i} \left( \frac{\tau(z^i)}{z^i} \hat{\eta}_c^i + \tau(z^i) \hat{e}_c^i \right). \]
Now set \( \tau(z^i) = 0 \) and substitute for the definition of the actual compensated elasticity \( e^i_c \) from eq. (10) to obtain:

\[
(1 + \frac{T_{zz}^i z^i_c}{1 - T^i_z}) e^i_c = \tilde{e}^i_c.
\]

Similarly, set \( \tau_z(z^i) = 0 \) and substitute for the definition of the actual income effect \( \eta^i \) from eq. (12) to obtain:

\[
(1 + \frac{T_{zz}^i z^i_c}{1 - T^i_z}) \eta^i = \tilde{\eta}^i.
\]

This proves the statement in footnote 7: with knowledge on the tax schedule, \( e^i_c \) and \( \eta^i \) can easily be derived from \( \tilde{e}^i_c \) 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 \( \tilde{z}(R^i, T^i_z) \) and substituting for the virtual elasticities, I can write:

\[
\frac{dz^i}{z^i} = -\tilde{e}^i_c \frac{dT^i_z}{1 - T^i_z} + \tilde{\eta}^i \frac{dR^i - z^i dT^i_z}{z^i (1 - T^i_z)}.
\]

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 \( dT^i_z \). Moreover, notice that \( dR^i - z^i dT^i_z = -\tau_z(z^i) d\kappa \), for which one could substitute policy-induced changes in the tax burden for a given labor income.

However, one cannot simply estimate eq. (34) 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. (34) would therefore lead to an endogeneity bias. To see this clearly, note that the change in marginal tax rates is given by:

\[
dT^i_z = T^i_{zz} dz + \tau_z(z^i) d\kappa.
\]
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. (34) 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. (32)-(33). Alternatively, one could substitute for eq. (35) and \( dR_i - z^i dT_z^i = -\tau_z(z^i) d\kappa \) into eq. (34) and rearrange to obtain the reduced form regression equation:

\[
(36) \quad \frac{dz^i}{z^i} = -\tilde{e}_c^i \frac{\tau_z(z^i) d\kappa}{1 - T_z^i} + \tilde{\eta}^i \frac{-\tau_z(z^i) d\kappa}{z^i (1 - T_z^i)}.
\]

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. (36) problematic.