Pareto-efficient tax breaks

Sebastian Koehne
IIES, Stockholm University

Dominik Sachs
European University Institute

November 2, 2016

Abstract

We analyze Pareto-efficient tax breaks for work-related expenses in a Mirrleesian environment. In addition to a labor-leisure choice, they decide how to spend their money between consumption and work-related goods. We derive an efficiency condition that relates the (implicit) subsidy rate for work-related goods to the marginal tax rate. The condition holds irrespective of the skill distribution and the taste for redistribution. If the efficiency condition is violated, we show how to Pareto improve such allocations. For the policy-relevant case where the work-related good contributes to the time available for market work (e.g., domestic services, child care or care for elderly relatives), Pareto-efficiency requires that the implicit time-investment subsidy lies strictly between zero and the labor wedge. Moreover, the efficient relation between labor wedges and implicit time-investment subsidies is solely determined by the Frisch elasticity of labor supply. This relationship also allows for a simple back-of the envelope calculation of the resource gains from moving towards a Pareto efficient allocation.

JEL codes: D82, H21

Keywords: Optimal Taxation, Tax Deduction, Pareto Improving Tax Reform

1 Introduction

Tax breaks are ubiquitous. In the United States, estimated revenue losses from numerous tax exemptions, deductions, and other special provisions in personal taxation exceed $1.3 trillion in 2016 (US Congress, Joint Committee on Taxation, 2015).[1] In the presence of tax breaks, the design of income taxation becomes multidimensional. It is no longer sufficient to define a schedule of marginal tax rates – the tax base becomes an object of tax design in itself.

Many existing tax breaks concern work-related expenses.[2] For example, the US tax code promotes investments in an individual's earnings ability or human capital (broadly defined, including health). In particular, there are tax credits for the tuition for post-secondary education, deductions for medical expenses and exclusions of employer contributions for health care. Expenses for professional services that provide child care or long-term care for elderly parents can also be deducted in various ways. Because the ability to engage in market work depends on the personal involvement in dependent care, such tax breaks are also work related in a general sense. Finally, there are miscellaneous itemized deductions in the US tax code that cover job-related clothing or equipment and unreimbursed work-related expenses.[3]

In this paper, we study the efficient design of such work-related tax breaks. Despite the substantial body of research emerging from the seminal optimal taxation model of Mirrlees (1971), so far the design of efficient tax breaks has received relatively little attention in the literature. We explore tax breaks in a Mirrleesian environment with a work-related consumption good and a general consumption good, similar to the multi-good taxation models of Atkinson and Stiglitz (1976) or Mirrlees (1976). An important innovation over earlier studies is that we explore the possibility of Pareto improvements rather than the maximization of a given social objective. Another key difference to earlier studies is that we focus on tax breaks rather than commodity taxes to achieve a desired outcome. This difference is more than just semantics. Although earlier studies have suggested cases where some consumption goods should be taxed less than others (e.g., Atkinson and Stiglitz, 1976; Christiansen, 1984; Jacobs and Boadway,

[1]This paper focuses on tax breaks in personal taxation. There are also substantial tax breaks in the taxation of corporations.

[2]Other major tax breaks include the net exclusion of pension contributions and earnings, reduced rates of tax on dividends and long-term capital gains, and deductions for mortgage interest on owner-occupied residences.

[3]Similar tax breaks also exist in other countries. In Germany and Sweden, for instance, there are tax breaks for acquiring domestic services such as child care, cleaning or reparations in the personal home. Moreover, expenses for education and health (insurance premiums and/or out-of-pocket spending) are also treated favorably by the tax systems of many OECD countries.

2014), the question as to how preferential tax rates should relate to the marginal income tax remains unanswered.

Our main theoretical result is a sharp condition for the Pareto-efficient design of tax breaks. We show that a tax system can only be Pareto efficient if an *incentive-adjusted no-arbitrage principle holds.* That is, the labor wedge and the implicit subsidy on work-related consumption should be set in proportion to the relative impact of labor income and work-related consumption on informational rents. Equivalently, the incentive-adjusted cost of a marginal unit of general consumption must be the same whether it is financed through more work or less work-related consumption. This condition has to hold independent of the redistributive preferences – it describes the Pareto frontier. Further, it is independent of the distribution of skills.

Besides providing an intuitive understanding of efficient allocations, this result generates a recipe on how to improve allocations that do not fulfill the efficiency condition. We show that the description of Pareto-superior reforms becomes particularly simple if the work-related good is a time investment (e.g., a professional service that frees up time from domestic work or dependent care). Starting from a situation with a zero deductibility (full deductibility) for such expenses, we show that there exists a Pareto-improving reform in which the consumption of this good and labor supply are increased (decreased) and other consumption is decreased (increased).

An important question is of course how large such an inefficiency is. In ongoing work, we quantitatively evaluate the Pareto improvement that is possible if the US tax rules for household services are reformed.

1.1 Related literature

The paper is closely related to Werning (2007) who studies the Pareto-efficient design of non-linear income taxes in a classical Mirrlees (1971) environment where individuals only make a labor supply decision. As he shows, Pareto efficiency alone does not provide a strong restriction on the design of tax schedules. For each tax schedule, there exists a distribution of skills that justifies this schedule as Pareto efficient. By contrast, in our environment with one dimension of heterogeneity and two choices (how much to work and how to spend the income between normal consumption and work-related consumption), Pareto efficiency provides a sharp restriction on tax schedules that is independent of the skill distribution. In line with Werning (2007), we then also consider Pareto improving reforms.[4]

The findings by Corlett and Hague (1953), Atkinson and Stiglitz (1976), Christiansen (1984) and Jacobs and Boadway (2014) show that complements to work should be subsidized (taxed at lower rates) as compared with other consumption. Similarly, Kleven (2004) shows in a Ramsey framework that any consumption good that requires little time (or even saves time) should be taxed at low rates. Although these contributions generally suggest that work-related goods (in particular, investments in the endowment of time for market work) should be favored by the tax system, they do not address how the optimal subsidy rate relates to the marginal income tax. Moreover, none of these papers consider the possibility of Pareto-improving reforms.[5] Mirrlees (1976) also explores a nonlinear taxation model with many goods. Among other things, he notes that the first-order conditions from utilitarian welfare maximization can be informative more generally for Pareto-efficient allocations after some manipulations, but he does not further explore this insight or pursue its policy implications.

We establish a set of necessary conditions for Pareto efficiency by minimizing the aggregate resource costs within a class of incentive-neutral allocation perturbations. This approach is similar in spirit to explorations of intertemporal perturbations (Rogerson, 1985; Golosov et al., 2003; Farhi and Werning, 2012).[6] Our paper is particularly related to Farhi and Werning (2012), who evaluate the potential gains from variations of consumption across time without making any individuals worse off. Crucially, all these contributions rely on a separability between work and consumption. In the present model, by contrast, we study work-related goods whose impact is by definition not separable from labor-leisure choices. Hence, for variations that involve work-related consumption, holding utilities fixed is no longer equivalent with preserving incentive compatibility and, thus, incentive-neutral perturbations need to manipulate labor supplies in addition to general and work-related consumption.

The paper is also related to the literature on human capital subsidies. In particular, we show how the "Siamese Twins" result by Bovenberg and Jacobs (2005) can be interpreted in

[4]Relatedly, Lorenz and Sachs (2015) extend the approach of Werning (2007) to an environment with a participation margin. Applied to Germany, they find a Pareto-inefficient structure of marginal tax rates.

[5]Laroque (2005) and Kaplow (2006) study Pareto improvements when the preferences for consumption goods are weakly separable from labor. In their framework, a uniform taxation of all consumption goods is optimal.

[6]Similarly, Koehne (2015) studies a class of incentive-neutral consumption perturbations in a case with durable and nondurable goods.

our environment. We also relate to a finding by Kapicka (2015), who studies the evolution of labor wedges across time in a learning-by-doing framework. Further, our paper is related to Ho and Pavoni (2016), who study the efficient design of child care subsidies. Their focus is on describing properties of particular welfare optima in the context of child care, whereas our focus is on Pareto-improving reforms. We therefore consider our work as complementary.

Finally, the general topic of efficient tax breaks is also emphasized by Saez (2004), who studies deductions for charitable giving in a model with a contribution good, linear taxes and subsidies, and social welfare maximization. Based on numerical simulations, his paper suggests that subsidy rates on charitable giving should typically lie below the earnings tax rate. Our results identify another policy-relevant class of expenses that should be imperfectly deductible from taxable income: services that enhance the time endowment for market work.

2 Model

We explore the concept of Pareto efficiency in a Mirrleesian taxation model with two consumption goods. One of these goods is nonseparable with labor/leisure and represents work-related consumption. Examples include job-related equipment, apparel, books, home offices, and professional services that free up time for market work. More broadly, the work-related good may also capture a health investment. The second consumption good represents general consumption and is separable from labor/leisure (but possibly nonseparable with work-related consumption). In Section 3.2.1, we provide an extension to many work-related goods.

2.1 Preferences

Individual agents are heterogeneous in their skill $n \in \mathcal{N} := [n_0, n_1] \subset \mathbb{R}_+$. The distribution of skill types in the economy is defined by a smooth probability density $f : \mathcal{N} \to \mathbb{R}$ with full support. Agents' preferences are described by a concave and continuously differentiable function $u : \mathbb{R}_+^4 \to \mathbb{R}$. Utility $u(c, d, y; n)$ is strictly increasing in general consumption c and work-related consumption d, and strictly decreasing in output y.

Throughout the paper, we assume that utility is additively separable between general consumption and output:

$$u(c, d, y; n) = w(c, d) + v(d, y; n), \tag{1}$$

where w and v are concave and continuously differentiable and $w(c,d)$ is strictly increasing in both arguments, whereas $v(d,y;n)$ is strictly increasing in d and strictly decreasing in y. This functional form draws a clear distinction between general consumption c and work-related consumption d based on the separability of the former from the disutility of work.

A leading example in our analysis are work-related expenses that affect the time endowment for (market) work and leisure. For instance, the work-related good may be a service that replaces the agent's engagement in non-market work (e.g., child care, domestic services, elderly care for close relatives). Similarly, the work-related good may represent a (curative or preventive) health investment that reduces the number of sick days in a given year or delays the worker's retirement.

Assumption 1 (Time-endowment model) *The utility function is given by*

$$u(c,d,y;n) = w(c) + \tilde{v}\left(E(d) - \frac{y}{n}\right) \tag{2}$$

where $\tilde{v}' > 0 > \tilde{v}''$ *and* $E' > 0 > E''$.

Under Assumption 1, the worker has a concave utility function $\tilde{v}$ defined over leisure, where leisure is the difference between the endowment of time $E(d)$ (net of sick days and non-market work) and hours of labor supply $l = y/n$. To avoid indeterminacies between labor supply and time-enhancing consumption, we assume that the time-endowment function is strictly concave. Alternatively, we could assume that individual output were strictly concave in hours worked.

2.2 Technology

The output good can be transformed one-for-one into general consumption and work-related consumption. Therefore, an allocation $(c(n), d(n), y(n))_{n\in\mathcal{N}}$ is *resource feasible* if

$$\int_{n_0}^{n_1} (y(n) - c(n) - d(n)) f(n)\, dn \geq 0. \tag{3}$$

2.3 Incentive compatibility

The skill realizations are private information. Therefore, allocations $(c(n), d(n), y(n))_{n\in\mathcal{N}}$ need to be *incentive compatible*, i.e., they must satisfy

$$u(c(n), d(n), y(n); n) \geq u(c(n'), d(n'), y(n'); n) \quad \forall n, n' \in \mathcal{N}. \tag{4}$$

An allocation is *incentive feasible* if it is resource feasible and incentive compatible.

2.4 Definition of wedges

The *labor wedge* τ_y is a well-known concept and defined by

$$1 - \tau_y(n) = -\frac{u_y(c(n), d(n), y(n); n)}{u_c(c(n), d(n), y(n); n)}. \tag{5}$$

We define the *work-related consumption wedge* as the gap between the marginal rate of transformation between work-related consumption and other consumption and the marginal rate of substitution between the two. Formally, we set

$$1 - \tau_d(n) = \frac{u_d(c(n), d(n), y(n); n)}{u_c(c(n), d(n), y(n); n)}. \tag{6}$$

According to our definition, a positive value of τ_d implies that expenses for work-related consumption are subsidized (relative to other consumption). If $\tau_d = 0$, individuals bear the full marginal cost of work-related consumption.

Implicit deduction rate. The ratio of wedges τ_d/τ_y corresponds to an implicit deduction rate for work-related expenses. In particular, if expenses for work-related goods are fully deductible from taxable income at the margin, then we have $\tau_d = \tau_y$.

Alternatively, the work-related consumption wedge may be interpreted as a type of commodity tax. However, since this wedge generally differs across agents, it cannot be created by standard forms of (linear) commodity taxation. Rather, the underlying tax needs to be income-dependent and/or nonlinear in the quantity of work-related consumption. For this reason, it appears more natural to interpret the work-related consumption wedge as resulting from deduction rules in the income tax system. We now present a formal decentralization of incentive-feasible allocations by means of income taxation below.

2.5 Decentralization

Next, we justify our mechanism design approach to taxation and discuss how to implement an allocation as a competitive equilibrium with taxes. We demonstrate that any incentive-feasible allocation $(c(n), d(n), y(n))_{n\in\mathcal{N}}$ can be decentralized through a general (non-linear and non-separable) income tax that depends on labor incomes and work-related expenses. Equivalently, there can be a labor income tax with a non-linear, non-separable deduction rate for work-related expenses.

A simple application of the taxation principle (Hammond, 1979; Rochet, 1985) implies that any incentive-feasible allocation can be implemented by a tax function $\bar{\mathcal{T}}(\cdot,\cdot)$ defined as

$$\bar{\mathcal{T}}(y(n), d(n)) := y(n) - d(n) - c(n),$$

and $\bar{\mathcal{T}}(y, s) := \infty$ for any pair (y, s) that is not part of the incentive-feasible allocation.

In order to construct a less extreme implementation, note that a tax function $\mathcal{T}(\cdot,\cdot)$ implements the given allocation $(c(n), d(n), y(n))_{n\in\mathcal{N}}$ if and only if, for all n,

$$(d(n), y(n)) \in \arg\max_{d,y} \; u\,(c, d, y; n) \quad \text{s.t.} \;\; c + d \leq y - \mathcal{T}(y, d). \tag{7}$$

In fact, many functions $\mathcal{T}(\cdot,\cdot)$ exist that satisfy this set of conditions. What they need to satisfy for sure is $\mathcal{T}(y(n), d(n)) = \bar{\mathcal{T}}(y(n), d(n))$ for all n. We now derive the lower envelope of the set of tax schedules that satisfy (7) using an approach similar to that of Werning (2011), who studies the lower envelope of tax schedules in a framework with income and savings taxes. The lower envelope is least extreme in punishing choices that are not part of the incentive-feasible allocation.

Let us construct for each type n a function $\mathcal{T}_n(\cdot,\cdot)$ such that:

$$u\,(y - d - \mathcal{T}_n\,(y, d)\,, d, y; n) = u\,(c(n), d(n), y(n); n) \quad \forall\,(y, d)\,. \tag{8}$$

Note that this construction is possible if u is continuous and unbounded above and below in general consumption.[7] We know by construction that $\mathcal{T}_n(y(n), d(n)) = \bar{\mathcal{T}}(y(n), d(n))$, because otherwise Equation (8) would not hold for $(y, d) = (y(n), d(n))$. For this tax schedule, the agent of type n is indifferent between $(y\,(n)\,, d\,(n))$ and any other pair (y, d).

We claim that the upper envelope of the tax functions $\mathcal{T}_n$ implements the incentive-feasible

[7]In particular, this approach does not rely on the separability assumption of Equation (1).

allocation. Define

$$\mathcal{T}^*(\cdot,\cdot) := \sup_n \mathcal{T}_n(\cdot,\cdot).$$

In this definition, the supremum is in fact a maximum because the type space is compact and $\mathcal{T}_n$ is continuous in n.

Proposition 1 (Implementation) *The tax function $\mathcal{T}^*$ implements the incentive-feasible allocation $\{c(n), d(n), y(n)\}_{n\in\mathcal{N}}$. Moreover, if $\mathcal{T}$ is another tax function that implements the allocation, then $\mathcal{T} \geq \mathcal{T}^*$.*

Because $\mathcal{T}^*$ is the point-wise supremum over a set of continuous functions, $\mathcal{T}^*$ is lower semi-continuous. It is not necessarily continuous or differentiable. Yet, we conjecture that the differentiability of $\mathcal{T}^*$ can be established based on envelope theorems similar to Milgrom and Segal (2002).

Tax systems with deductions. Note that for any nonseparable tax system $\mathcal{T}(y,d)$ there exists an equivalent tax system with deduction rules for work-related consumption, $\hat{\mathcal{T}}(y - \kappa(y,d)\,d)$, that yields the same outcome for all choices (y,d). For example, let $\hat{\mathcal{T}}$ be the identity function and set

$$\kappa(y,s) = \frac{y - \mathcal{T}(y,d)}{d}.$$

Then, by definition,

$$\mathcal{T}(y,d) = \hat{\mathcal{T}}(y - \kappa(y,d)\,d) \quad \forall (y,d).$$

3 Pareto-efficient allocations

Throughout this section, we analyze a social planning problem under private information on ability n. Rather than maximizing social welfare for a specific welfarist objective, we are interested in Pareto improvements of a given baseline allocation. Specifically, we ask: can we reduce aggregate resources without making any individual worse off?

3.1 The envelope condition

Following common practice in optimal tax theory, we replace the original incentive-compatibility constraint by a relaxed condition. Specifically, we define the agents' indirect utilities as

$$U(n) = u(c(n), d(n), y(n); n)$$

and replace the incentive-compatibility constraint (4) by the envelope condition,

$$\dot{U}(n) = v_n(d(n), y(n); n). \tag{9}$$

It is well-known that the envelope condition is necessary for incentive compatibility (e.g., Mirrlees, 1976). The envelope condition is sufficient provided that the second-order condition of utility maximization with respect to the reported type is satisfied.[8]

3.2 Utility-neutral perturbations

Consider a given incentive compatible and resource feasible allocation $(c(n), d(n), y(n))$. For each type n, we change work-related consumption by some (positive or negative) amount ε and we adjust output and general consumption (up or down) such that utility remains unchanged and the envelope condition continues to hold. The planner seeks to minimize the amount of aggregate resources.

Formally, we define the perturbed allocation $(\hat{c}(n), \hat{d}(n), \hat{y}(n))$ as follows:

$$\begin{aligned} \hat{d}(n) &= d(n) + \varepsilon(n) \\ \hat{c}(n) &= c(n) + \gamma(n) \\ \hat{y}(n) &= y(n) + \delta(n) \end{aligned} \tag{10}$$

[8] As shown by Mirrlees (1976), the envelope condition is sufficient if for all n, n' we have

$$y'(n)\, v_{yn}\left(d(n), y(n); n'\right) + d'(n)\, v_{dn}\left(d(n), y(n); n'\right) \geq 0.$$

It is difficult to evaluate this condition without precise knowledge about the gradients of output and work-related goods with respect to skills. In our quantitative application, we verify ex post that this condition is satisfied at the computed allocations.

subject to the constraints

$$u\left(\hat{c}(n), \hat{d}(n), \hat{y}(n); n\right) = u(c(n), d(n), y(n); n) \quad (11)$$

$$\frac{du\left(\hat{c}(n), \hat{d}(n), \hat{y}(n); n\right)}{dn} = v_n\left(\hat{d}(n), \hat{y}(n); n\right) \quad (12)$$

which ensure that no individual is made worse off (11) and that the reform is incentive compatible (12). Intuitively, for each value of ε, there are unique values of γ and δ such that both of these conditions are fulfilled.

We express the consumption perturbation in terms of indirect utilities,

$$\gamma(n) = w^{-1}\left(U - v\left(d(n) + \varepsilon(n), y(n) + \delta(n); n\right); d(n) + \varepsilon(n)\right) - c(n)$$

where w^{-1} denotes the inverse of $w(c, d)$ with respect to its first argument. Now the planner problem becomes an optimal control problem with state variable U and controls ε and δ:

$$\max_{U,\varepsilon,\delta} \int_{n_0}^{n_1} \left[\delta - \varepsilon - w^{-1}\left(U - v(d+\varepsilon, y+\delta; n); d+\varepsilon\right) + c\right] f\, dn \quad (13)$$

$$\text{s.t.} \quad U = w(c,d) + v(d, y; n), \quad \dot{U}(n) = v_n(d+\varepsilon, y+\delta; n), \quad U(n_0), U(n_1) \text{ given.}$$

The maximum principle for problem (13) generates the following necessary condition for Pareto efficiency.[9]

Proposition 2 (Incentive-adjusted no-arbitrage principle) *Suppose $v_{ny}(d, y; n) \neq 0$ for all $(d, y; n)$. A necessary condition for Pareto efficiency is that the two equivalent conditions*

$$\frac{u_c - u_d}{-v_{nd}} = \frac{u_c + u_y}{v_{ny}} \quad (14)$$

$$\frac{-\tau_d}{v_{nd}} = \frac{\tau_y}{v_{ny}} \quad (15)$$

hold for almost all types n with $v_{nd}(d(n), y(n); n) \neq 0$ and that $\tau_d = 0$ holds for almost all types n with $v_{nd}(d(n), y(n); n) = 0$.

[9]Note that Mirrlees (1976, p. 337) has derived a first-order condition very similar to our Equation (15). He did not pursue the intuition ("No intuitive explanation of the result has occurred to me."(Mirrlees, 1976, p. 338)) nor the implication in greater detail though.

Proposition 2 provides a sharp condition for Pareto efficiency: an allocation (or tax system) cannot be Pareto efficient unless an *incentive-adjusted no-arbitrage principle* holds. Importantly, Equations (14) and (15) do not only help to understand properties of the optimum, but can also be applied to test for the efficiency of real-world tax systems. These conditions are particularly powerful in the sense that they can be evaluated without information on the distribution of skills. By contrast, in the standard Mirrlees model where individuals only make a labor supply decision, for each tax schedule there exists a skill distribution such that the tax schedule is Pareto efficient (Werning, 2007).

To gain intuition for the Pareto efficiency condition, note that an individual always has two ways to finance a marginal unit of general consumption: the individual can reduce her consumption of the work-related good d by one unit or work more to increase y by one unit. Those two options change individual utilities by $u_c - u_d$ and $u_c + u_y$, and affect the incentive problem through the envelope condition according to $-v_{nd}$ and v_{ny}, respectively. Thus, Equation (14) states that the price of general consumption in utility terms relative to incentive costs must be the same for both ways of financing consumption.

According to our terminology, the negative wedge $-\tau_d$ corresponds to a tax on work-related consumption, whereas the labor wedge τ_y represents a tax on income. The version of the no-arbitrage principle in Equation (15) states that the taxes on work-related consumption and income should be proportional to the influence of these goods on informational rents (the cross derivative of v). In particular, a full deductibility of work-related expenses ($\tau_d = \tau_y$) can only be Pareto efficient if a one-for-one change of work-related consumption and output would leave informational rents unaffected, meaning that the utility change of such a perturbation does not (locally) depend on the skill type, i.e., $v_{nd} + v_{ny} = 0$.

Equation (15) also encompasses two important benchmark results in the optimal taxation literature. First, as shown by Atkinson and Stiglitz (1976), the wedge between general consumption and work-related consumption should be zero if consumption and work enter preferences in a separable way ($v_{nd} = v_d = 0$).[10] Second, Bovenberg and Jacobs (2005) show that education

[10] We assume from the start that general consumption is additively separable from the preferences for work. Therefore, strictly speaking, we obtain the Atkinson-Stiglitz result only for the case of additively separable preferences. The uniform taxation result is true more generally whenever the consumption preferences are *weakly* separable from work. Moreover, this result can be established based on the principle of Pareto efficiency alone (Laroque, 2005; Kaplow, 2006), despite the fact that the original Atkinson-Stiglitz result relies on optimal income taxation.

should be subsidized at the exactly same rate as income is taxed. Thus, in our terminology, it would be optimal to have $\tau_d = \tau_y$ when d represents an educational investment. Proposition 2 helps to understand this well-known finding in the theory of optimal taxation subsidies from a different angle. It also highlights the generality of their finding by showing that marginal education subsidies and marginal income taxes in fact coincide along the entire Pareto frontier in their framework. We obtain their setup if we set

$$u(c, d, y; n) = w(c) - V\left(\frac{y}{n\phi(d)}\right)$$

where $\phi(\cdot)$ is concave and $V(\cdot)$ convex. In that case, we have $-\frac{v_{nd}}{v_{ny}} = \frac{1-\tau_d}{1-\tau_y}$.[11] Hence, by Equation (15), Pareto efficiency dictates

$$\frac{\tau_d}{1-\tau_d} = \frac{\tau_y}{1-\tau_y},$$

implying that $\tau_d = \tau_y$ holds in any Pareto-efficient allocation. In particular, the marginal informational rents of education d and income y sum up to zero if $\tau_d = \tau_y$.[12]

Finally, the perturbation approach that underlies Proposition 2 is also informative on the direction of Pareto-improving reforms. Evaluating the Lagrangian of problem (13) at any given allocation indicates whether (marginal) Pareto improvements are associated with higher or lower levels of work-related consumption relative to the status quo. We discuss this implication in more detail in the time-endowment model further below.

3.2.1 Extension: multiple work-related goods and dynamic labor wedges

The no-arbitrage principle of Proposition 2 is very general and extends without difficulty to multiple work-related goods. Specifically, consider an environment with a vector $d = (d_1, \ldots, d_K)$ of work-related goods and a utility function of the form $u(c, d, y; n) = w(c, d) + v(d, y; n)$. Analogous to Equation (6), define the work-related consumption wedge for good k in this environment

[11] More precisely, we obtain

$$-\frac{v_{nd}}{v_{ny}} = \frac{V'\frac{y\phi'}{n^2\phi^2} + V''\frac{y^2\phi'}{n^3\phi^3}}{V'\frac{1}{n^2\phi} + V''\frac{y}{n^3\phi^2}} = \frac{y\phi'}{\phi} = \frac{\frac{y\phi'}{n\phi^2}V'}{\frac{1}{n\phi}V'} = \frac{1-\tau_d}{1-\tau_y}.$$

[12] In a follow-up paper, Jacobs and Bovenberg (2011) discuss different human capital production functions for which it can become optimal to subsidize human capital less or more than income is taxed. Also these results can be understood in terms of marginal informational rents. Relatedly, Baake et al. (2004) establish a full-deductibility result for a multiplicative specification of the form $v(d, y; n) = g(d, y)h(n)$. Equation (15) implies that this result extends to the entire Pareto frontier as well.

as

$$1-\tau_d^k(n)=\frac{u_{d_k}(c(n),d(n),y(n);n)}{u_c(c(n),d(n),y(n);n)}.$$

Then, the proof of Proposition 2 establishes the following necessary condition for Pareto efficiency (assuming $v_{n,d_{k'}}\neq 0$):

$$\frac{\tau_d^k}{\tau_d^{k'}}=\frac{v_{n,d_k}}{v_{n,d_{k'}}}\quad \text{for all } 1\leq k,k'\leq K. \tag{16}$$

Once more, this condition states that distortions should be determined in proportion to the marginal informational rents.

If the work-related goods do not have a direct consumption value, they become similar to labor supplies in dynamic environments. Therefore, we can relate Equation (16) to characterizations of labor wedges across time (in frameworks without uncertainty). In particular, we can capture processes of human capital formation through *learning by doing* or *learning or doing* as studied by Kapicka (2015). In those cases, labor supply decisions affect agents' future productivities and, thus, the preferences over outputs become nonseparable across time.

Specifically, we can interpret the work-related good d_k as the negative of output produced in period k and interpret y as the output in an initial period. Suppose that the preferences take the form $u=w(c)-V(z_0,z_1,\ldots,z_K)$, where V is increasing and convex, and labor supplies are given by $z_0=y_0/n$ and $z_k=-d_k/n$ for $k\geq 1$. Then, $\tilde{\tau}_k:=\tau_d^k$ represents the labor wedge at time k. For this specification, we show in Appendix A.3 that Equation (16) implies

$$\frac{\frac{\tilde{\tau}_k}{1-\tilde{\tau}_k}}{\frac{\tilde{\tau}_{k'}}{1-\tilde{\tau}_{k'}}}=\frac{1+\sum_{t=0}^{K}z_t\frac{V_{t,k}}{V_k}}{1+\sum_{t=0}^{K}z_t\frac{V_{t,k'}}{V_{k'}}}.$$

This condition replicates a finding by Kapicka (2015) and imposes a sharp restriction on the evolution of labor wedges across time. Kapicka also provides an insightful economic interpretation of this condition and decomposes it into an effect on the contemporaneous information rent, an anticipation effect due to the link between contemporaneous and future labor supplies, and an accumulation effect through human capital formation.

3.3 Time-enhancing investments

We now turn to the time-endowment model (Assumption 1) as a special case. We think that this specification captures several real-world situations. Many individuals hire professionals such as housekeepers, gardeners or cleaning staff to free up time from domestic chores. Further, individuals can invest in their time endowment by paying someone to care for their children, an ill spouse or elderly relatives. The costs of these services are often tax deductible. In this subsection, we analyze the efficiency of such tax breaks through the lens of our model.

For this purpose, we now derive our result of Proposition 2 in a heuristic and intuitive fashion for time-enhancing investments. Consider an incentive-feasible allocation where income of type-n individuals is taxed at the marginal rate $\tau_y(n)$ and time investment d is subsidized in the form of a tax break at the marginal rate $\tau_d(n)$. Can we improve this allocation in a Pareto sense and how would that allocation reform look like?

Assume we increase time investment d by an amount ε. Then, in order not to violate incentive compatibility, output y has to be adjusted by $\delta(\varepsilon)$ such that the informational rent stays constant. Note that the informational rents in the time-endowment model are given by

$$v_n = \frac{y}{n^2}\tilde{v}'\left(E(d) - \frac{y}{n}\right). \tag{17}$$

As Equation (17) highlights, time investment d reduces the informational rent by lowering the marginal value of leisure $\tilde{v}'$. Hence, if time investment increases by a positive amount ε, output y has to increase by a positive $\delta(\varepsilon)$ such that

$$\tilde{v}'\left(E(d) - \frac{y}{n}\right)\frac{y}{n^2} = \tilde{v}'\left(E(d+\varepsilon) - \frac{y+\delta(\varepsilon)}{n}\right)\frac{y+\delta(\varepsilon)}{n^2}. \tag{18}$$

Importantly, in this setup, output has a twofold impact on informational rents: more output means that the marginal value of leisure $\tilde{v}'$ increases and that the leisure gain y/n^2 of a marginal shirker becomes bigger.

Now, to insure that individual utility is unaffected, consumption has to be changed by $\gamma(\varepsilon)$ to ensure that

$$w(c) + \tilde{v}\left(E(d) - \frac{y}{n}\right) = w(c+\gamma(\varepsilon)) + \tilde{v}\left(E(d+\varepsilon) - \frac{y+\delta(\varepsilon)}{n}\right). \tag{19}$$

Given that incentive compatibility is ensured and individual utility is also held constant by construction, the question remains what the resource gains or costs of such a reform are. Resources change according to $R(\varepsilon) = -\varepsilon + \delta(\varepsilon) - \gamma(\varepsilon)$ and a Pareto improvement is possible whenever this expression is positive for some (possibly negative) values of ε.

It is helpful to look at the marginal resource effect: $R'(\varepsilon) = -1 + \delta'(\varepsilon) - \gamma'(\varepsilon)$. In particular, note that implicit differentiation of (19) yields $\gamma'(\varepsilon) = -(1 - \tau_d^\varepsilon(n)) + \delta'(\varepsilon)(1 - \tau_y^\varepsilon(n))$, which implies that the marginal resource gain can be written as:

$$R'(\varepsilon) = -\tau_d^\varepsilon(n) + \delta'(\varepsilon)\tau_y^\varepsilon(n). \tag{20}$$

This expression is very intuitive. An increase in time investment d accompanied by an adjustment in general consumption that holds utility constant implies resource costs that are proportional to the marginal distortion τ_d. The higher the rate of subsidization of time investment, the more costly it is to further increase it. By contrast, an increase in output y (accompanied by an adjustment in general consumption that holds utility constant) increases resources in proportion to the labor wedge τ_y. Finally, these two effects are related to each other by the factor $\delta'(\varepsilon)$ which measures by how much output has to be adjusted to maintain incentive compatibility after a change in time investment.

Implicit differentiation of (18) yields:

$$\delta'(\varepsilon) = \frac{1 - \tau_d^\varepsilon(n)}{1 - \tau_y^\varepsilon(n)} \frac{1}{1 + e^\varepsilon(n)}, \tag{21}$$

where $e^\varepsilon(n)$ is the Frisch elasticity of labor supply with respect to the net-of-tax rate $1 - \tau_y$ (holding time investment fixed).[13] Hence, two factors determine the adjustment of output that keeps informational rents constant. First, the ratio of wedges $1 - \tau_d$ and $1 - \tau_y$ captures the relative productivities of time investment and output in generating leisure. The larger the

[13]Formally, we have

$$e^\varepsilon(n) = -\frac{\tilde{v}'\left(E(d-\varepsilon) - l^\varepsilon(n)\right)}{\tilde{v}''\left(E(d-\varepsilon) - l^\varepsilon(n)\right) l^\varepsilon(n)}, \quad \text{where } l^\varepsilon(n) = \frac{y + \delta(\varepsilon)}{n}.$$

In frameworks without time investment, this variable represents the standard Frisch elasticity of labor supply with respect to the net-of-tax rate. In the present environment, the concept becomes slightly more specific. We can interpret $e^\varepsilon(n)$ as the Frisch elasticity of labor supply with respect to the net-of-tax rate *holding time investment fixed* or, equivalently, as the Frisch elasticity of labor supply with respect to the net-of-tax rate when time investments are fully deductible from taxable income (not holding time investment fixed). See Appendix A.2 for further details.

subsidy rate τ_d, the smaller is the marginal effect of time investment on leisure (relative to the effect of output), and hence the smaller is the required change of output to hold information rents constant when time investment is raised. Second, there is a scaling factor $\frac{1}{1+e} < 1$ due to the fact that output has a double impact on informational rents (through the marginal value of leisure and the leisure gain of a marginal shirker). The larger the Frisch elasticity e, the larger is the discrepancy between the roles of output and leisure for informational rents, and the smaller is the necessary change of output to keep informational rents fixed.[14]

Overall, we can now write the resource gradient $R'(\varepsilon)$ as:

$$R'(\varepsilon) = -\tau_d^{\varepsilon}(n) + \tau_y^{\varepsilon}(n)\frac{1-\tau_d^{\varepsilon}(n)}{1-\tau_y^{\varepsilon}(n)}\frac{1}{1+e^{\varepsilon}(n)}. \tag{22}$$

This expression of the marginal resource gain only depends on the wedges and the Frisch elasticity of labor supply. Evaluating it at any particular allocation gives direct insights on whether the allocation can be Pareto improved by an allocation reform associated with an increase or decrease in time investment. Moreover, it provides a sharp condition for Pareto optimality. An allocation can only be Pareto efficient if $R'(0) = 0$. This reasoning leads to the following result.

Proposition 3 (Imperfect deduction of time investments) *Define* $e\,(n) := -\frac{n\tilde{v}'}{y\tilde{v}''}$. *Under Assumption 1, a necessary condition for Pareto efficiency is*

$$\frac{\tau_d(n)}{1-\tau_d(n)} = \frac{\tau_y(n)}{1-\tau_y(n)}\frac{1}{1+e(n)} \tag{23}$$

for almost all types $n \in \mathcal{N}$. *In particular, if* $0 < \tau_y < 1$, *we have* $0 < \tau_d < \tau_y$. *That is, investments in the time endowment should be imperfectly deductible from taxable income at the margin.*

This result immediately implies that time-enhancing investments should be less than fully deductible. Otherwise, a Pareto improvement is possible. In particular, the rate of deductibility should be lower, the higher the Frisch elasticity of labor supply. Further, in a Pareto optimum,

[14] As Equation (17) shows, output has a linear effect on informational rents through the leisure gain of a marginal shirker and a nonlinear effect through the change in the marginal value of leisure. The magnitude of latter depends on the curvature of leisure utility, captured by the inverse of the elasticity e. By contrast, time investment affects informational rents only via the marginal value of leisure (captured once more by the reciprocal of e). Hence, the bigger the elasticity e, the larger is the discrepancy between output and leisure in determining the informational rents.

the time-investment subsidy τ_d and the labor wedge τ_y co-move positively: the higher the labor supply distortion, the higher is the subsidy rate on time investment (holding the Frisch elasticity fixed). That is, Pareto-efficient tax breaks tend to be regressive in the specific sense that individuals with higher marginal tax rates also face higher implicit subsidies for time-enhancing investments.

Finally, we state a result about the directions of Pareto-improving reforms that seem particularly policy relevant.

Proposition 4 (Pareto-improving reforms) *Suppose that Assumption 1 holds. Starting from a tax system with positive labor wedges and no tax breaks for time-enhancing investment, a Pareto-improving reform exists where individuals invest more in their time endowment, have less general consumption, and work more.*

By contrast, starting from a tax system where time-enhancing investment is fully deductible and labor wedges are positive, a Pareto-improving reform exists where individuals invest less in their time endowment, have more general consumption, and work less.

Heuristically, evaluating the marginal resource gain at an allocation without a tax break ($\tau_d = 0$) yields:

$$R'(0) = \frac{\tau_y(n)}{1-\tau_y(n)} \frac{1}{1+e(n)} > 0.$$

More specifically, a Pareto improvement is obtained by increasing time-enhancing investment slightly by ε and increasing income by $\frac{1}{1-\tau_y}\frac{1}{1+e}\varepsilon$ to maintain the envelope condition. To hold utility constant, general consumption is reduced by $\frac{e}{1+e}\varepsilon$. The resource gain of this reform is $\frac{\tau_y}{1-\tau_y}\frac{1}{1+e}\varepsilon$ and increases in the size of the labor wedge. Moreover, the resource gain decreases in the labor supply elasticity. This finding is related to the fact that the efficient subsidy rates decrease in the labor supply elasticity according to Proposition 3.

By contrast, the marginal resource gain at an allocation with a full deductibility ($\tau_d = \tau_y$) is negative: $R'(0) = -\tau_y(n)\frac{e(n)}{1+e(n)} < 0$. Here, decreasing time-enhancing investment slightly by $|\varepsilon|$ and decreasing income slightly by $\frac{1}{1+e}|\varepsilon|$ maintains the envelope condition. To hold utility constant, general consumption needs to increase by $(1-\tau_y)\frac{e}{1+e}|\varepsilon|$. Hence, the reform saves $\tau_y\frac{e}{1+e}|\varepsilon|$ units of resources. The resource gain is bigger, the higher the elasticity of labor supply and the higher the labor wedge.

Back-of-the-envelope calculation. Our formula for the marginal resource gains allows for a simple back-of-the-envelope calculation. Consider for example a situation with a zero subsidy on time-enhancing investment, $\tau_d = 0$, a marginal tax rate of $\tau_y = 0.3$ and a Frisch elasticity of $e = 0.5$. The marginal resource gain then is $R'(0) = \frac{0.3}{0.7}\frac{1}{1+0.5} \approx 0.29$. This number means that for each additional marginal dollar spend on the time-investment good (in an incentive-compatible and utility-neutral way), \$0.29 in resources are freed up, which seems a very significant economic effect. For example, if the reform induces a household to spend \$500 more on the time-investment good, a first-order approximation of the gain would be \$145. Of course, this first-order approximation is problematic because the wedges τ_d and τ_y change as we increase time-enhancing investment in our incentive-compatible and utility-neutral reform. This caveat motivates our thorough quantitative exercise, in which we introduce implicit subsidies for housekeeping expenditures.

3.4 A sufficient-statistics approach

Our theoretical results were derived with a mechanism-design approach in terms of the parameters of the underlying model. This contrasts with the sufficient-statistics approach (Chetty, 2009), where welfare statements are made solely in terms of empirically measurable concepts like elasticities and expenditure shares, for example. The sufficient-statistics approach has the advantage that it relies less on structural assumptions. On the other hand, a general drawback is that even if one has empirical estimates of these sufficient statistics, these are typically local estimates that are mainly applicable to small policy reforms. Our goal is to think about larger reforms that move the economy from an inefficient allocation to the Pareto frontier, which calls for a more structural approach. By contrast, for a pure test of whether a given allocation is Pareto efficient, the sufficient-statistics approach seems most appropriate.

In this subsection, we show that efficient taxes and subsidies can indeed be related in terms of sufficient statistics. It turns out that this condition will rely on four different elasticities. However, three of those elasticities are not standard concepts and lack empirical evidence.

The sufficient-statistics approach can be most easily understood in a Ramsey taxation environment. Consider the optimization problem of a representative agent with utility $u(c, d, y)$

facing a linear labor income tax rate t_y and a linear subsidy on the work-related good t_d,

$$V(t_y, t_d) = \max_{y,d} \; u\left((1 - t_y)y - (1 - t_d)d, d, y\right).$$

The Ramsey planner solves

$$\max_{t_y, t_d} V(t_y, t_d) \; \text{ subject to } \quad - t_d d + t_y y \geq R,$$

where R is some exogenous revenue requirement. As we show in Appendix A.4, the solution to this problem satisfies

$$\varepsilon_{d,1-t_d} \frac{t_d}{1 - t_d} - \frac{t_y}{1 - t_d} \varepsilon_{y,1-t_d} \frac{y}{d} - \frac{t_y}{1 - t_y} \varepsilon_{y,1-t_y} + \frac{t_d}{1 - t_y} \varepsilon_{d,1-t_y} \frac{d}{y} = 0,$$

where $\varepsilon_{a,b} = \frac{da}{db}\frac{b}{a}$ is the elasticity of variable a with respect to variable b. This condition only depends on sufficient statistics, i.e., four elasticities, income and expenditures for the work-related good. If we had knowledge about the values of all these elasticities, this condition would be very useful to test for the efficiency of a given allocation. However, to the best of our knowledge, there is not a lot of reliable evidence for the elasticities $\varepsilon_{d,1-t_d}$, $\varepsilon_{y,1-t_d}$ and $\varepsilon_{d,1-t_y}$.

To sum up, we think that the sufficient-statistics approach is not the right one for our purpose because the relevant elasticities are not standard objects currently estimated in the empirical literature. Further, we explore potentially larger reforms, where local values of elasticities would not suffice.

4 Conclusion

In this paper, we studied the efficient design of tax breaks for work-related goods from a mechanism-design perspective. We derive a sharp efficiency condition that holds along the Pareto frontier and is independent of the distribution of types. If the work-related good enhances the time endowment for market work, this condition implies that expenditures should be less then fully deductible from taxable income. Further, the efficient subsidy rate decreases in the Frisch elasticity of labor supply.

Our approach also provides a recipe on how to improve Pareto-inefficient allocations: we show how to perturb an allocation in an incentive-compatible way such that resources are maximized holding individual utilities constant. In ongoing work, we quantitatively apply our method to assess the potential efficiency gains of reforming the US tax code.

A Appendix

A.1 Proofs

Proof of Proposition 1. First, we claim

$$\mathcal{T}_{\hat{n}}(y(n), d(n)) \leq \mathcal{T}_n(y(n), d(n)) \text{ for all } \hat{n}, n.$$

Suppose, to the contrary, that there exist some $\hat{n}, n$ with

$$\mathcal{T}_{\hat{n}}(y(n), d(n)) > \mathcal{T}_n(y(n), d(n)).$$

Equivalently,

$$y(n) - d(n) - \mathcal{T}_{\hat{n}}(y(n), d(n)) < y(n) - d(n) - \mathcal{T}_n(y(n), d(n)) = c(n).$$

By the construction of $\mathcal{T}_{\hat{n}}$, we have

$$u\left(y(n) - d(n) - \mathcal{T}_{\hat{n}}\left(y(n), d(n)\right), d(n), y(n); \hat{n}\right) = u\left(c(\hat{n}), d(\hat{n}), y(\hat{n}); \hat{n}\right).$$

Hence, the previous inequality implies

$$u\left(c(n), d(n), y(n); \hat{n}\right) > u\left(c(\hat{n}), d(\hat{n}), y(\hat{n}); \hat{n}\right),$$

which violates the incentive compatibility constraint.

Hence, we have established $\mathcal{T}_{\hat{n}}(y(n), d(n)) \leq \mathcal{T}_n(y(n), d(n))$ for all $\hat{n}, n$. This implies

$$\mathcal{T}^*(y(n), d(n)) = \sup_{\hat{n}} \mathcal{T}_{\hat{n}}(y(n), d(n)) = \mathcal{T}_n(y(n), d(n)) \quad \forall n.$$

Moreover, by construction, the weak inequality $\mathcal{T}^*(y, d) \geq \mathcal{T}_n(y, d)$ holds for all pairs (y, d).

Because agent n was indifferent between all pairs (y, d) under the tax system $\mathcal{T}_n$, it follows that the agent weakly prefers $(y(n), d(n))$ under the tax system $\mathcal{T}^*$.

Finally, let $\mathcal{T}$ be another tax function that implements the allocation. Suppose, to the contrary, that there exists some pair (y, d) with $\mathcal{T}(y, d) < \mathcal{T}^*(y, d)$. Then, by the definition of $\mathcal{T}^*$, there exists some n with

$$\mathcal{T}(y, d) < \mathcal{T}_n(y, d).$$

However, because $\mathcal{T}_n(y, d)$ was constructed to make the type n agent indifferent between (y, d) and $(y(n), d(n))$, the agent will strictly prefer (y, d) over $(y(n), d(n))$ under the tax system $\mathcal{T}$. This contradicts the assumption that $\mathcal{T}$ implements the allocation. ■

Proof of Proposition 2. We extend the optimal control problem (13) to a framework where $d = (d_1, \ldots, d_K)$ and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_K)$ are real vectors. We set up the augmented Hamiltonian

$$\begin{aligned} H =& f(n) \left[\delta - \sum_{k=1}^{K} \varepsilon_k - w^{-1} \left(U - v \left(d + \varepsilon, y + \delta; n \right); d + \varepsilon \right) + c \right] \\ &+ \lambda(n) \left[U - w(c, d) + v \left(d, y; n \right) \right] \\ &+ \mu(n) v_n \left(d + \varepsilon, y + \delta; n \right) \end{aligned}$$

and derive the first-order conditions for ε_k and δ (evaluated at $\varepsilon_k = \delta = 0$):

$$\begin{aligned} f \left[1 - v_{d_k} \frac{1}{w_c} - \frac{w_{d_k}}{w_c} \right] &= \mu v_{n d_k} \\ f \left[1 + v_y \frac{1}{w_c} \right] &= -\mu v_{n y} \end{aligned}$$

where we have used the derivatives of the inverse function, $w_c^{-1} = 1/w_c$ and $w_{d_k}^{-1} = -w_{d_k}/w_c$.

If $v_{n d_k} = 0$, we obtain

$$\tau_d^k = 1 - \frac{w_{d_k} + v_{d_k}}{w_c} = 0.$$

Otherwise, if $v_{n d_i} \neq 0$, we can divide the first-order conditions and obtain

$$\frac{w_c - (v_{d_k} + w_{d_k})}{w_c + v_y} = \frac{1 - \frac{v_{d_k} + w_{d_k}}{w_c}}{1 + \frac{v_y}{w_c}} = -\frac{v_{n d_k}}{v_{n y}},$$

which establishes Equation (14). Now, Equation (15) follows directly from the definition of the wedges τ_d and τ_y. ■

Proof of Proposition 3. Under Assumption 1, the relationship between the cross-derivatives of leisure utility imply

$$-\frac{v_{nd}}{v_{ny}} = \frac{\frac{y}{n^2}E'\tilde{v}''}{\frac{y}{n^3}\tilde{v}'' - \frac{\tilde{v}'}{n^2}} = \frac{nE'}{1+e(n)}$$

where $e(n) := -\frac{n\tilde{v}'}{y\tilde{v}''}$. Moreover, using the definition of wedges we have

$$\frac{1-\tau_d}{1-\tau_y} = \frac{E'\tilde{v}'}{\frac{\tilde{v}'}{n}} = nE'.$$

Therefore, Equation (15) implies

$$\frac{\tau_d}{1-\tau_d} = \frac{\tau_y}{1-\tau_y}\frac{1}{1+e(n)}.$$

■

A.2 Frisch elasticity of labor supply in the time-endowment model

Consider an individual with skill n who maximizes utility subject to (locally) linear taxes and subsidies at rates t and s and a lump-sum transfer g. The decision problem is

$$\max_{c,d,l} \quad w(c) + \tilde{v}(E(d)-l) \quad \text{s.t.} \quad c + (1-s)d \leq (1-t)nl + g$$

Denoting the Lagrange multiplier for the budget constraint by λ, the first-order conditions of this problem are

$$\begin{aligned} w'(c) &= \lambda \\ E'(d)\tilde{v}'(E(d)-l) &= \lambda(1-s) \\ \tilde{v}'(E(d)-l) &= \lambda(1-t)n. \end{aligned}$$

Holding fixed the marginal utility of consumption λ and the time investment d, differentiation of the last equation with respect to $1-t$ yields

$$-\tilde{v}''(E(d)-l)\frac{\partial l}{\partial(1-t)} = \lambda n.$$

Therefore, the Frisch elasticity of labor supply (holding time investment fixed) is given by

$$e = \frac{\partial l}{\partial (1-t)} \frac{(1-t)}{l} = -\frac{\lambda n (1-t)}{\tilde{v}'' (E(d)-l) \, l} = -\frac{\tilde{v}' (E(d)-l)}{\tilde{v}'' (E(d)-l) \, l},$$

where we have used the first-order condition for labor supply.

Alternatively, consider a (locally) linear tax system where time investment is perfectly deductible from taxable income. Then the first-order conditions for time investment and labor supply are

$$\begin{aligned} E'(d)\tilde{v}' (E(d)-l) &= \lambda (1-t) \\ \tilde{v}' (E(d)-l) &= \lambda (1-t) \, n \end{aligned}$$

and differentiation (holding fixed λ) implies

$$\begin{aligned} E''\tilde{v}' \frac{\partial d}{\partial (1-t)} + E'\tilde{v}'' \left[E' \frac{\partial d}{\partial (1-t)} - \frac{\partial l}{\partial (1-t)} \right] &= \lambda \\ \tilde{v}'' \left[E' \frac{\partial d}{\partial (1-t)} - \frac{\partial l}{\partial (1-t)} \right] &= \lambda n. \end{aligned}$$

We substitute the second equation into the first and obtain

$$\frac{\partial d}{\partial (1-t)} = \frac{\lambda - E'\lambda n}{E''\tilde{v}'}.$$

Now the second equation yields

$$\frac{\partial l}{\partial (1-t)} = E'\lambda \frac{1-E'n}{E''\tilde{v}'} - \frac{\lambda n}{\tilde{v}''}.$$

Note that the first-order conditions for d and l imply $E' = \frac{1}{n}$. Hence, we obtain the Frisch elasticity of labor supply as

$$e = \frac{\partial l}{\partial (1-t)} \frac{(1-t)}{l} = -\frac{\lambda n (1-t)}{\tilde{v}'' l} = -\frac{\tilde{v}'}{\tilde{v}'' l}$$

where the last identity follows from the first-order condition for l.

A.3 Dynamic labor wedges

Consider a utility function of the form $w(c) - V(z_0, z_1, \ldots, z_K)$, where $z_0 = \frac{y_0}{n}$ and $z_k = \frac{-d_k}{n}$ for $k \geq 1$. Note that the informational rents in this model are given by

$$v_n = \sum_{t=0}^{K} \frac{z_t}{n} V_t(z_0, z_1, \ldots, z_K).$$

Hence, the marginal effect of d_k on the informational rent is

$$v_{n,d_k} = -\frac{1}{n^2}\left(V_k + \sum_{t=0}^{K} z_t V_{t,k}\right),$$

which implies

$$\frac{v_{n,d_k}}{v_{n,d_{k'}}} = \frac{V_k + \sum_{t=0}^{K} z_t V_{t,k}}{V_{k'} + \sum_{t=0}^{K} z_t V_{t,k'}}.$$

The definition of the wedges implies

$$\frac{1-\tau_d^k}{1-\tau_d^{k'}} = \frac{V_k}{V_{k'}}.$$

Hence, we can rewrite the previous condition as

$$\frac{v_{n,d_k}}{v_{n,d_{k'}}} = \frac{1-\tau_d^k}{1-\tau_d^{k'}} \frac{1+\sum_{t=0}^{K} z_t \frac{V_{t,k}}{V_k}}{1+\sum_{t=0}^{K} z_t \frac{V_{t,k'}}{V_{k'}}}.$$

Therefore, Equation (16) implies

$$\frac{\tau_d^k}{\tau_d^{k'}} = \frac{v_{n,d_k}}{v_{n,d_{k'}}} = \frac{1-\tau_d^k}{1-\tau_d^{k'}} \frac{1+\sum_{t=0}^{K} z_t \frac{V_{t,k}}{V_k}}{1+\sum_{t=0}^{K} z_t \frac{V_{t,k'}}{V_{k'}}}.$$

A.4 Sufficient statistics result

The individual problem of a representative household with linear taxes is

$$V(t_y, t_d) = \max_{y,d}\ u\left((1-t_y)y - (1-t_d)d, d, y\right).$$

The Ramsey problem reads as

$$\max_{t_y, t_d} V(t_y, t_d) \text{ s.t. } \; - t_d d + t_y y \geq R.$$

The first-order condition for t_y (using the envelope theorem) is

$$-y + \lambda \left(y + t_y \frac{\mathrm{d}y}{\mathrm{d}t_y} - t_d \frac{\mathrm{d}d}{\mathrm{d}t_y} \right) = 0.$$

The first-order condition for t_d (using the envelope theorem again) is

$$d + \lambda \left(-d + t_y \frac{\mathrm{d}y}{\mathrm{d}t_d} - t_d \frac{\mathrm{d}d}{\mathrm{d}t_d} \right) = 0.$$

The first condition can be rewritten as

$$\lambda - 1 = \frac{t_y}{1 - t_y} \varepsilon_{y,1-t_y} - \frac{t_d}{1 - t_y} \varepsilon_{d,1-t_y} \frac{d}{y}.$$

The second one we can write as

$$\lambda - 1 = \varepsilon_{d,1-t_d} \frac{t_d}{1 - t_d} - \frac{t_y}{1 - t_d} \frac{y}{d} \varepsilon_{y,1-t_d}.$$

Putting both conditions together yields

$$\varepsilon_{d,1-t_d} \frac{t_d}{1 - t_d} - \frac{t_y}{1 - t_d} \varepsilon_{y,1-t_d} \frac{y}{d} - \frac{t_y}{1 - t_y} \varepsilon_{y,1-t_y} + \frac{t_d}{1 - t_y} \varepsilon_{d,1-t_y} \frac{d}{y} = 0,$$

which is the formula in the main body of the text.

References

Atkinson, A. and J. Stiglitz (1976): "The design of tax structure: Direct versus indirect taxation," *Journal of Public Economics*, 6, 55 – 75.

Baake, P., R. Borck, and A. Löffler (2004): "Complexity and progressivity in income tax design: Deductions for work-related expenses," *International Tax and Public Finance*, 11, 299–312.

Bovenberg, A. L. and B. Jacobs (2005): "Redistribution and education subsidies are Siamese twins," *Journal of Public Economics*, 89.

Chetty, R. (2009): "Sufficient Statistics for Welfare Analysis: A Bridge Between Structural and Reduced-Form Methods," *Annu. Rev. Econ.*, 1, 451–488.

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

Corlett, W. J. and D. C. Hague (1953): "Complementarity and the excess burden of taxation," *Review of Economic Studies*, 21–30.

Farhi, E. and I. Werning (2012): "Capital taxation: Quantitative explorations of the inverse Euler equation," *Journal of Political Economy*, 120, 398–445.

Golosov, M., N. Kocherlakota, and A. Tsyvinski (2003): "Optimal Indirect and Capital Taxation," *Review of Economic Studies*, 70, 569–587.

Hammond, P. J. (1979): "Straightforward individual incentive compatibility in large economies," *Review of Economic Studies*, 46, 263–282.

Ho, C. and N. Pavoni (2016): "Efficient Child Care Subsidies," *Working Paper*.

Jacobs, B. and R. Boadway (2014): "Optimal linear commodity taxation under optimal non-linear income taxation," *Journal of Public Economics*, 117, 201 – 210.

Jacobs, B. and A. L. Bovenberg (2011): "Optimal Taxation of Human Capital and the Earnings Function," *Journal of Public Economic Theory*, 13, 957–971.

Kapicka, M. (2015): "Pareto Efficient Taxation with Learning by Doing," 2015 meeting papers, Society for Economic Dynamics.

Kaplow, L. (2006): "On the undesirability of commodity taxation even when income taxation is not optimal," *Journal of Public Economics*, 90, 1235 – 1250.

Kleven, H. J. (2004): "Optimum taxation and the allocation of time," *Journal of Public Economics*, 88, 545–557.

Koehne, S. (2015): "On the Taxation of Durable Goods," Working Paper 5194, CESifo Working Paper Series.

Laroque, G. R. (2005): "Indirect taxation is superfluous under separability and taste homogeneity: a simple proof," *Economics Letters*, 87, 141 – 144.

Lorenz, N. and D. Sachs (2015): "Identifying Laffer Bounds: A Sufficient-Statistics Approach with an Application to Germany," *Scandinavian Journal of Economics (forthcoming)*.

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

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

——— (1976): "The Optimal Structure of Incentives and Authority within an Organization," *Bell Journal of Economics*, 7, 105–131.

Rochet, J.-C. (1985): "The taxation principle and multi-time Hamilton-Jacobi equations," *Journal of Mathematical Economics*, 14, 113–128.

Rogerson, W. P. (1985): "Repeated Moral Hazard," *Econometrica*, 53, 69–76.

Saez, E. (2004): "The optimal treatment of tax expenditures," *Journal of Public Economics*, 88, 2657–2684.

US Congress, Joint Committee on Taxation (2015): "Estimates Of Federal Tax Expenditures For Fiscal Years 2015-2019," Publication JCX-141R-15, https://www.jct.gov.publications.html.

Werning, I. (2007): "Pareto Efficient Income Taxation," MIT. Mimeo.

——— (2011): "Nonlinear capital taxation," MIT. Mimeo.