Pareto-efficient tax deductions*

Sebastian Koehne
Kiel University

Dominik Sachs
LMU Munich

July 9, 2019

Abstract

We analyze Pareto-efficient tax deduction rules for work-related expenses (e.g. housekeeping services, child care or elderly care). Pareto efficiency dictates a tight rule for how the rate of deductibility should vary with income and expenditures. An immediate implication is a recipe for designing Pareto-improving tax reforms. We apply our theory to housekeeping services in the U.S.: Introducing deduction rules such that between 55% (low expenses) and 85% (high expenses) of housekeeping services can be marginally deducted from taxable income yields a Pareto improvement if combined with a slight increase in marginal tax rates. Nobody is made worse-off and tax revenue increases by 20 Dollars per capita.

JEL codes: D82, H21

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

*Contact: koehne@economics.uni-kiel.de, dominik.sachs@econ.lmu.de. An earlier version of this paper was circulated under the title "Pareto-efficient tax breaks". We thank Spencer Bastani, Felix Bierbrauer, Pierre Boyer, Christian Bredemeier, Emmanuel Farhi, Emanuel Hansen, Albert Jan Hummel, Bas Jacobs, Erzo F.P. Luttmer as well as seminar participants in Bristol, Cologne, Vaxjo and the CESifo Public Sector Economics Area Conference.

1 Introduction

Tax codes for personal taxation contain a large number of tax deductions, sometimes also called tax expenditures or tax breaks. Many of those deductions concern work-related expenses. For example, the U.S. tax code promotes investments in individual earnings ability, e.g. expenses for education and health. Expenses for professional services that provide child care or long-term care for elderly parents, and thereby free up time for market work, can also be deducted in various ways. Similar provisions exist in many other OECD countries.

Such deduction possibilities are often considered to be inequitable or regressive because high income earners make us of them to a larger extent.[1] In this paper, we argue that such concerns are likely to be misplaced because the design of such deduction rules can be separated from distributional concerns. We provide a simple recipe to redesign a given tax system (jointly changing marginal tax rates and deduction rules) in a Pareto-improving way: tax revenue is increased and nobody is made worse-off.

More formally, we study tax deduction rules for work-related goods in a Mirrleesian framework with multiple consumption goods. We define a work-related good as a good whose value is nonseparable with the (dis)utility of work. This notion encompasses a large range of special cases, including professional services that replace non-market work (housekeeping, child care, dependent care, etc.) and deterministic models of human capital formation. We show that the way how marginal deduction rates should vary with income and expenses can be studied without reference to a particular social welfare function.[2] Following this idea, we set up an explicit resource maximization problem holding individual utilities at given baseline levels (e.g. the utilities levels implied by current policies). Our objective is to fine-tune the marginal tax rates on income and the marginal deduction rates on work-related goods such that we obtain a Pareto-efficient outcome. This approach provides us with three results. First, we derive a condition for how the marginal deduction rate should vary with income and expenditures along the second-best Pareto frontier. Second, this condition is also a test for the Pareto efficiency of

[1] The Congressional Budget Office of the U.S. addresses this issue in their economic and budget outlook: "Tax expenditures are distributed unevenly across the income scale."(Congressional Budget Office, 2016, p.103). They also devoted an extra report to it (Congressional Budget Office, 2013). This report has been taken up also by the press, see e.g. Washington Post (2013).

[2] Some earlier theoretical studies of mixed taxation have pointed out that the efficiency condition for the ratio of different marginal tax rates is independent of the exact social welfare function. See Atkinson and Stiglitz (1976), Mirrlees (1976), Laroque (2005), and Kaplow (2006). We are more explicit on our relation to these papers below.

any given baseline tax system. Third, and most important, we establish a constructive procedure that Pareto improves suboptimal tax systems and quantitatively apply it to housekeeping services in the United States.

First Result: To put the first result into perspective, it is important to refer to the literature on commodity taxation. In the presence of deduction rules, individuals face effective prices for work-related goods that depend on their incomes and expenses for these goods. Therefore, the analysis of tax deductions can be viewed as a problem of nonlinear income and nonlinear commodity taxation (mixed taxation) in the tradition of Atkinson and Stiglitz (1976) and Mirrlees (1976). A well-known insight from this literature is that complements to work should be subsidized by the tax system (e.g., Christiansen, 1984), providing a rationale for work-related tax deductions. Close to this literature, we first derive a formal characterization of the second-best Pareto frontier in terms of allocation variables of our model (Proposition 1). We show how the characterization nests several special cases from the literature.

An important distinction is that we focus on tax deduction rules as an implementation device. We rewrite the Pareto-efficiency condition in terms of marginal deduction rates and show that Pareto efficiency dictates a tight relationship for how the rate of deductibility should vary with income and expenses (Corollary 1). We consider the implementation with tax deduction rules as particularly relevant. It can yield implicit prices for work-related goods that depend on income and expenses (and hence can yield Pareto-efficient outcomes) but, as opposed to nonlinear commodity taxation, is a policy instrument that is widely used in practice.

Second Result: An immediate implication is that we can use this Pareto-efficiency condition to test whether a given tax system is second-best Pareto efficient. The efficiency test is particularly powerful because it can be evaluated without information on the distribution of skills.[3] This result contrasts with the pure nonlinear income tax setting without work-related goods and without deduction rules: Werning (2007) shows that for each tax schedule in the classical Mirrlees setup, there exists a skill distribution such that the tax schedule is Pareto efficient.[4] We argue that standard deduction rules (e.g., a full or zero deductibility of an expense in a continuous income range), are unlikely to satisfy the efficiency test in our setup and

[3]This property is similar to Laroque (2005), and Kaplow (2006), who study Pareto efficiency for the special case where the goods are not work-related and therefore the famous uniform commodity taxation result by Atkinson and Stiglitz (1976) applies.

[4]Lorenz and Sachs (2016) 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.

hence a Pareto-improving reform of a given tax system will often exist. The reason is that the Pareto-efficiency test imposes a strict deduction rule at each point of the income distribution.

Third Result/Application: The formal problem that yields the conditions for Pareto efficiency also provides us with a recipe on how to Pareto improve a given tax system. The utility levels that are reached by different agents for the given tax system serve as constraints for the problem: no individual should be made worse off by the reform. To find a Pareto improvement, net resources (tax revenues) are maximized given these utility levels and incentive compatibility.[5]

We apply this approach to study an introduction of tax deductions for domestic (housekeeping) services in the United States. We focus on single households, where such expenses are currently non-deductible. As we document with data from the Consumer Expenditure Survey, annual expenses increase from 200 to more than 1500 Dollars along the income distribution. Despite the different expenditure levels, our analysis shows that tax deductions can be introduced in a distributionally neutral way. Our results indicate that the consumption of domestic services in the status quo is approximately 20 percent below the level of a Pareto-efficient outcome. We outline a Pareto-improving reform that increases the marginal deduction rates from zero to 55-85%. At the same time, marginal tax rates are shifted up slightly, in particular for lower incomes. The annual revenue gains of this reform are roughly $20 per household.

1.1 Related literature

The seminal findings by Corlett and Hague (1953), Atkinson and Stiglitz (1976) and Christiansen (1984) show that complements to work should be subsidized (taxed at lower rates) as compared with other consumption. Relatedly, 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. While these contributions generally suggest that work-related goods should be favored by the tax system, they abstract from tax deductions and they do not consider the possibility of Pareto-improving reforms. The present paper does explicitly consider tax deduction rules as an implementation device. Further, we describe a Pareto-improving reform and quantitatively evaluate it for a concrete example.

Similar to us, Laroque (2005) and Kaplow (2006) also study Pareto improvements but focus

[5] We could also consider alternative Pareto-improving reforms, where the utility levels are strictly increased for some agents. Selecting one reform would then require a welfare criterion. Similar to Werning (2007), we prefer to avoid this problem by focussing on the reform that maximizes net resources and keeps the utility levels fixed.

on the case where preferences for consumption goods are weakly separable from labor. In their framework, a uniform taxation of all consumption goods is optimal, i.e., the Atkinson-Stiglitz theorem holds. They show that whenever commodity taxation is not uniform, a joint reform of equalizing commodity taxes and changing the income tax exists that yields a Pareto improvement. We focus on the case where preferences for leisure and consumption are not separable. In this case, Pareto-efficient policy rules are more complicated and tax deductions are a plausible real-world implementation. We also provide a concrete quantitative application to housekeeping services.

Olovsson (2015) studies optimal taxation with home production in a quantitative Ramsey model. Deduction possibilities or nonlinear taxes are not present in his environment. He finds that it is optimal to tax services at a lower rate than goods in order to make home production less attractive. Our quantitative results show that household services should be strongly deductible from taxable income—a policy that drives a wedge between the price of household services and the price of other consumption goods.

An important class of household services (and work-related goods more generally) is child care. Our paper is therefore related to recent works by Ho and Pavoni (2016) and Bastani et al. (2017), who study the optimal 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, deduction rules as an implementation device, and household services other than care. We therefore consider our work as complementary. Domeij and Klein (2013) theoretically and quantitatively make a case for subsidizing child care in a Ramsey environment. While they do not study Pareto-improving reforms and focus on linear subsidies or constant rates of deductibility, their approach is generally related as they look at reforms that maximize a distribution-neutral welfare function.

The paper is also related to the recent literature on human capital subsidies. We especially relate to the work by Kapicka (2015), who studies the evolution of labor wedges across time in a learning-by-doing framework. In Section 3.3, we adapt our static model to include some of the dynamic considerations of his framework. Stantcheva (2017) explores human capital subsidies in a dynamic model with uncertainty. Among other things, she shows how the optimal education subsidy depends on the wage elasticity with respect to ability.

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 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.

2 Model

Individuals supply labor and choose how to allocate their income between two consumption goods. One of these goods is nonseparable with labor and represents work-related consumption. Examples include job-related equipment, apparel, books, home offices, and services that free up time for market work. More broadly, the work-related good may also capture a health investment. The work-related good is feasible for a tax deduction. The second consumption good represents general consumption and is separable from labor.

2.1 Preferences

Individuals 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. 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 (pre-tax labor income) y.

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

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

[6] Similarly, Koehne (2018) studies a class of incentive-neutral consumption perturbations in a case with durable and nondurable goods.

where w and v are continuously differentiable, concave in (c, d), and $w(c, d)$ is strictly increasing in c and weakly increasing in d, 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. The main purpose of the functional form is to facilitate the exposition and interpretation of the theoretical results. Yet, the general approach of the paper does not hinge on this assumption.

2.2 Tax system

Individuals face a nonlinear labor income tax schedule $\mathcal{T} : \mathbb{R} \to \mathbb{R}$ with a deduction rule $\mathcal{D} : \mathbb{R}_+^2 \to \mathbb{R}$ for work-related expenses. The deduction rule $\mathcal{D}(y, d)$ may be nonlinear and nonseparable between labor income y and work-related consumption d. The individual tax payment is given by $\mathcal{T}(y - \mathcal{D}(y, d))$, where $y - \mathcal{D}(y, d)$ represents the taxable income of the individual. We call the pair $(\mathcal{T}, \mathcal{D})$ a tax system.

Note that our specification of tax systems covers the entire universe of nonlinear, nonseparable tax functions. For any function $\hat{\mathcal{T}}(y, d)$, there trivially exists an equivalent tax system with a deduction rule for work-related expenses that yields the same tax payments. For example, let $\mathcal{T}$ be the identity function and set

$$\mathcal{D}(y, d) := y - \hat{\mathcal{T}}(y, d).$$

Then, by definition,

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

2.3 Individual problem and wedges

Individuals maximize their welfare given the tax system. They solve the following problem:

$$\max_{c,d,y} u(c, d, y; n) \text{ s.t. } c + d = y - \mathcal{T}(y - \mathcal{D}(y, d)) \tag{2}$$

The first-order conditions imply

$$-\frac{u_y}{u_c} = 1 - (1 - \mathcal{D}_y)\,\mathcal{T}'$$
$$\frac{u_d}{u_c} = 1 - \mathcal{D}_d\,\mathcal{T}'.$$

As these conditions show, the labor supply decision and the expenditure on the work-related good are both influenced by the marginal tax rate at the individual's taxable income as well as the marginal deduction rule. These two policy instruments jointly determine the respective wedges. First, there is the *labor wedge* τ_y given by

$$\tau_y := 1 + \frac{u_y}{u_c} = (1 - \mathcal{D}_y)\,\mathcal{T}'. \tag{3}$$

The labor wedge measures the gap between the marginal rate of transformation and the marginal rate of substitution between pre-tax income and consumption. As the right-hand side of Eq. (3) shows, the gap is induced by the tax system in the following way: an extra dollar of pre-tax income increases the individual's taxable income by $(1 - D_y)$ dollars; therefore, the tax bill grows by $(1 - D_y)\mathcal{T}'$ dollars. If the deduction rule depends on labor income, the labor wedge does not simply equal the marginal tax rate. For example, if higher income allows for more deductions ($D_y > 0$), the distortion on labor supply is lower than the mere accounting for the marginal tax rate would suggest.

The second relevant wedge in our environment is the *work-related expenditure wedge* τ_d. We define this wedge as the gap between the marginal rate of transformation between work-related goods and general consumption and the marginal rate of substitution between the two. Formally, we set

$$\tau_d := 1 - \frac{u_d}{u_c} = \mathcal{D}_d\,\mathcal{T}'. \tag{4}$$

This wedge captures the implicit subsidy to work-related goods relative to general consumption. Note that an extra dollar of work-related spending reduces the individual's taxable income by $\mathcal{D}_d$ dollars and diminishes the tax bill by $\mathcal{D}_d\mathcal{T}'$ dollars. By contrast, an extra dollar spent on general consumption is not deductible.

There are two natural benchmarks for the work-related expenditure wedge. If work-related spending is not deductible from taxable income ($\mathcal{D} = 0$), we have $\tau_d = 0$. By contrast, if work-related goods can be paid with pre-tax income ($\mathcal{D} = d$), i.e, if these expenses are fully deductible, we have $\tau_d = \tau_y$.

More generally, the difference between the two wedges captures the overall distortion to work-related spending induced by the tax system. We define the *net expenditure wedge* as

$$\bar{\tau}_d := \tau_y - \tau_d = (1 - \mathcal{D}_y - \mathcal{D}_d)\,\mathcal{T}'. \tag{5}$$

A marginal dollar of labor income that is spent on the work-related good increases the agent's tax bill by $\bar{\tau}_d$ dollars. Therefore, a zero net wedge means that the tax system is neutral with respect to work-related spending. For instance, if work-related expenditures are fully deductible irrespective of the level of income, we have $\mathcal{D}_d = 1$, $\mathcal{D}_y = 0$ and therefore $\bar{\tau}_d = 0$. A positive net wedge, by contrast, implies that work-related spending is at most imperfectly deductible.

Because the net wedge is mathematically redundant, we will formulate our theoretical findings mainly in terms of the labor wedge and the work-related expenditure wedge. Yet, the net wedge will be a useful measure of the optimal degree of deductibility in our quantitative evaluation further below.[7]

2.4 Applying the revelation principle

To characterize Pareto-improving reforms of the tax system, we make use of the revelation principle. By the revelation principle, any allocation that can be implemented through a tax system $\mathcal{T}\left(y - \mathcal{D}\left(y, d\right)\right)$, can also be implemented through an incentive-compatible direct mechanism. Formally, an allocation $(c(n), d(n), y(n))_{n \in \mathcal{N}}$ is *incentive compatible* if it satisfies

$$u\left(c\left(n\right), d\left(n\right), y\left(n\right); n\right) \geq u\left(c\left(n'\right), d\left(n'\right), y\left(n'\right); n\right) \quad \forall n, n' \in \mathcal{N}. \tag{6}$$

As usual, individual welfare maximization subject to the tax system establishes an incentive-compatible allocation. A simple application of the taxation principle (Hammond, 1979; Rochet, 1985) implies that the reverse is also true: for any incentive-compatible allocation, there ex-

[7]The concept of the net wedge is similar to the terminology of Bovenberg and Jacobs (2005) measuring whether education is taxed or subsidized on a net basis and to the concept of a net human capital subsidy in Stantcheva (2017).

ists a tax system that implements the allocation. This result also implies that if we find a Pareto-improving and incentive-compatible allocation perturbation, there exists a tax reform that implements this allocation perturbation.[8]

Following common practice in optimal tax theory, we replace the incentive-compatibility constraint by an envelope 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 (6) by the following condition:

$$\dot{U}(n) = v_n(d(n), y(n); n). \tag{7}$$

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.[9]

3 Pareto-improving reforms of tax deduction rules

Consider a given tax system $\mathcal{T}(y - \mathcal{D}(y, d))$ that generates an allocation $(c(n), d(n), y(n))_{n\in\mathcal{N}}$. In this section, we describe how a reform can be constructed in order to achieve a Pareto improvement. We show that such a reform exists unless the wedges implied by the tax system satisfy a continuum of efficiency conditions.

If an allocation is Pareto inefficient, then typically many Pareto-improving reforms exist. We will elaborate on those reforms that maximize tax revenue without altering the levels of individual utilities. Given an incentive-compatible allocation $(c(n), d(n), y(n))_{n\in\mathcal{N}}$ that is induced by the baseline tax system, we seek to find an incentive-compatible allocation $\left(\hat{c}(n), \hat{d}(n), \hat{y}(n)\right)_{n\in\mathcal{N}}$

[8] We present an implementation with an income-independent deduction rule in our quantitative analysis in Section 4.2.5. A general implementation result is derived in Appendix A.4.

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

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

Because we work with arbitrary baseline allocations, we cannot validate this condition theoretically. In our quantitative application, we verify ex post that the condition is satisfied at the computed allocations.

that maximizes resources

$$\int_{n_0}^{n_1} \left(\hat{y}(n) - \hat{c}(n) - \hat{d}(n)\right) f(n)\, dn$$

subject to:

$$u\left(\hat{c}(n), \hat{d}(n), \hat{y}(n); n\right) = u(c(n), d(n), y(n); n) \quad \forall n \in \mathcal{N}. \tag{8}$$

Eq. (8) ensures that no individual is worse off after the reform.

3.1 Constructing the reform

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. We seek to reduce the amount of resources needed for the given levels of utility. Formally, for every $n \in \mathcal{N}$, we define the elements $(\hat{c}(n), \hat{d}(n), \hat{y}(n))$ of the perturbed allocation as follows:

$$\begin{aligned} \hat{d}(n) &= d(n) + \varepsilon(n) \\ \hat{c}(n) &= c(n) + \gamma(n) \\ \hat{y}(n) &= y(n) + \delta(n) \end{aligned} \tag{9}$$

subject to the constraints

$$u\left(\hat{c}(n), \hat{d}(n), \hat{y}(n); n\right) = u(c(n), d(n), y(n); n), \tag{10}$$

$$\frac{du\left(\hat{c}(n), \hat{d}(n), \hat{y}(n); n\right)}{dn} = v_n\left(\hat{d}(n), \hat{y}(n); n\right), \tag{11}$$

where Eq. (10) ensures that no individual is made worse off and Eq. (11) ensures that the reform is incentive compatible.

We make a change of variables and express the consumption perturbation in terms of utility levels $U(n)$:

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

where w^{-1} denotes the inverse of $w(c, d)$ with respect to its first argument and $U(n)$ represents the utility of an agent with skill n at the original (and perturbed) allocation. Now we obtain an optimal control problem with state variable $U(n)$ and controls $\varepsilon(n)$ and $\delta(n)$:

$$\max_{U(n),\varepsilon(n),\delta(n)} \int_{n_0}^{n_1} \Big[\delta(n) - \varepsilon(n) - w^{-1}\Big(U(n) - v\Big(d(n) + \varepsilon(n), y(n) + \delta(n); n\Big), d(n) + \varepsilon(n)\Big) + c(n)\Big] f(n)dn \tag{12}$$

subject to

$$U(n) = w(c(n), d(n)) + v\,(d(n), y(n); n)$$
$$\dot{U}(n) = v_n\,(d(n) + \varepsilon(n), y(n) + \delta(n); n)\,.$$

3.2 Properties of Pareto-efficient allocations

Problem (12) implies a set of necessary conditions for Pareto efficiency. These conditions not only describe the allocation after a Pareto-improving reform has been implemented. More importantly for our purposes, the conditions are also a test for whether a given allocation is Pareto efficient and, hence, whether or not the reform described in Section 3.1 can yield a Pareto improvement.

By applying the maximum principle for Problem (12), we obtain the the following property of Pareto-efficient allocations.

Proposition 1 (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 following condition:*

$$\frac{u_c + u_y}{v_{ny}} = \frac{u_c - u_d}{-v_{nd}} \tag{13}$$

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

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 and increase income y by one unit. These 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, Eq. (13) shows that the cost of a marginal unit of general consumption (measured in utility terms relative to incentive costs) must be the same for both ways of financing the marginal unit of consumption. In this sense, Eq. (13) can be interpreted as an *incentive-adjusted no-arbitrage principle.*

Proposition 1 implies that Pareto-efficient deduction rules strongly depend on preference nonseparabilites between work-related goods, skills and labor supply. An alternative way of writing Eq. (13) is to express it in the form of wedges:

$$\tau_d = -\frac{v_{nd}}{v_{ny}} \tau_y. \tag{14}$$

which shows that Pareto efficiency dictates a tight relation between the labor and the work-related good wedge. For the purpose of testing the efficiency of a given tax system $(\mathcal{T}, \mathcal{D})$, we now rephrase Eq. (13) in terms of the tax system.

Corollary 1 *Consider a tax system* $(\mathcal{T}, \mathcal{D})$*. The allocation implemented by this tax system is Pareto efficient if and only if*

$$\mathcal{D}_d = -\frac{v_{nd}}{v_{ny}} (1 - \mathcal{D}_y) \tag{15}$$

for all types n *where* $\mathcal{T}'(y - D(y, d)) \neq 0$.

Although the right-hand side of Eq. (15) depends on the functional form of the leisure utility function v, this equation already suggests that Pareto-efficient deduction rules may be relatively complex. In particular, unless the ratio of cross derivatives v_{nd}/v_{ny} happens to take a very simple form, standard deduction rules (e.g., zero or full deductibility of work-related expenses within a continuous income range) are unlikely to be Pareto efficient. In such a case, the reform constructed in Section 3.1 will yield a Pareto improvement. In Section 4, we will further elaborate on how such reforms look like for the special case where d represents a time investment. We will also quantitatively apply the insight and construct a Pareto-improving reform for household-service expenditures in the United States. Before moving to the case of time investment, we show in Section 3.3 how our conditions for Pareto efficiency are related to various results in the literature.

3.3 Special cases and extensions

Proposition 1 encompasses several important benchmark results in the optimal taxation literature. Next, we present four well-known cases that have been discussed in the literature. In Section 4 we consider another practically relevant application: expenditures that enhance the time endowment for market work and leisure.

3.3.1 Uniform commodity taxation

As shown by Atkinson and Stiglitz (1976), the consumption choice should be undistorted if the preferences are separable between consumption and work. For separable preferences ($v_{nd} = v_d = 0$), Proposition 1 indeed implies $\tau_d = 0$.[10] Hence, Proposition 1 shows that uniform commodity taxes are necessary conditions for Pareto efficiency if the preferences are separable between consumption and work, reiterating the findings by Laroque (2005) and Kaplow (2006).

3.3.2 Subsidies to work-complementary goods

If the utility function of leisure takes the common form $v(d, y; n) = \tilde{v}(d, \frac{y}{n})$, we can interpret $l := \frac{y}{n}$ as hours worked and the skill level n as the individual's hourly productivity. In that case, Eq. (14) implies that τ_d has the same sign as the labor wedge if the utility of leisure has a positive cross derivative with respect to work-related consumption d and hours worked l.[11] In other words, Proposition 1 implies that work-complementary goods are subsidized in any Pareto-efficient allocation. This result is in line with the findings by Christiansen (1984).

[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.

[11] With $v(d, y; n) = \tilde{v}(d, l)$, we obtain

$$\tau_d = -\frac{v_{nd}}{v_{ny}}\tau_y = \frac{\frac{y\tilde{v}_{dl}}{n^2}}{-\frac{\tilde{v}_l}{n^2} - \frac{y\tilde{v}_{ll}}{n^3}}\tau_y.$$

The denominator of this expression is positive (assuming that the utility of leisure is decreasing and concave in hours worked). The sign of the right-hand side is hence determined by the cross derivative $\tilde{v}_{dl}$.

3.3.3 Human capital subsidies

Bovenberg and Jacobs (2005) show that education 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 1 helps to understand this well-known finding in the theory of optimal education 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 $-\frac{v_{nd}}{v_{ny}} = \frac{1-\tau_d}{1-\tau_y}$.[12] Hence, by Eq. (14), 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. Or, another way to put it, the net wedge is zero, i.e., we have $\bar{\tau}_d = 0$.

3.3.4 Multiple work-related goods and dynamic labor wedges

The no-arbitrage principle of Proposition 1 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 Eq. (4), define the work-related expenditure wedge for good k in this environment as

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

[12] More precisely, we obtain

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

Then, the approach of Proposition 1 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 the wedges should be determined in proportion to the marginal incentive effects of the respective goods.

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 Eq. (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.2 that Eq. (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.

4 An application to time-enhancing investments

We now turn to work-related expenses that affect the time endowment for (market) work and leisure. In particular, we consider services that replace the agent's engagement in non-market work.[13] This environment captures several real-world situations. Many individuals hire housekeepers, gardeners or cleaning staff to free up time from domestic chores. Further, individuals pay professionals to care for their children, an ill spouse or elderly relatives. The costs of these services are tax deductible in a number of countries (including Germany, Sweden and Denmark). In this section, we analyze the efficiency of such deductions through the lens of our model.

Throughout this section, we maintain the following assumption.

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

$$u(c, d, y; n) = w(c) + \tilde{v}\left(E(d) - \frac{y}{n}\right) \tag{17}$$

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 (net of non-market work) and hours of labor supply $l = y/n$.

4.1 Imperfect deductibility

Proposition 2 (Imperfect deduction of time-enhancing investments) *Under Assumption 1, a necessary condition for Pareto efficiency is* $0 < \tau_d(n) < \tau_y(n)$ *whenever the labor wedge of the considered type satisfies* $0 < \tau_y(n) < 1$.

By Proposition 2, Pareto efficiency requires a positive net wedge for time-enhancing investment: $\bar{\tau}_d(n) > 0$, implying that these investments should be positively, but imperfectly, deductible at the margin. To understand this result and to elaborate on possible Pareto-improving reforms, we now consider two particularly relevant benchmarks.

Proposition 3 (Introducing a tax deduction for time-enhancing investment) *Suppose that Assumption 1 holds. Starting from a tax system where time-enhancing investment is not deductible and labor wedges are positive, i.e.,* $0 = \tau_d(n) < \tau_y(n) < 1$*, a Pareto-improving reform exists where type-n individuals invest more in their time endowment, work more and consume less.*

[13] 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.

The virtue of deductions for time-enhancing investments can be most easily understood in a model version with discrete types. Consider an agent with skill n and a hypothetical shirker with skill $\hat{n} > n$ who mimics the type-n agent. A tax system without deductions of time-enhancing investment can be improved in the following steps. First, we increase time-enhancing investment d by a marginal unit and reduce general consumption c by the same amount. (This step raises the work-related expenditure wedge to a positive level.) Because the margin between work-related expenses and general consumption was undistorted at the baseline tax system, the perturbation has no impact on the utility of the truth-telling agent:

$$du = -w' + E'(d)\tilde{v}'\left(E(d) - \frac{y}{n}\right) = 0.$$

The shirking agent, however, consumes more leisure and therefore values a unit of time investment relatively less. Her utility thus falls:

$$d\hat{u} = -w' + E'(d)\tilde{v}'\left(E(d) - \frac{y}{\hat{n}}\right) < 0.$$

Hence, the joint change of consumption and time-enhancing investment has relaxed the incentive problem without affecting the utility of the truth-telling agent. In the next step, due to the relaxed incentive-compatibility constraint, it becomes possible to increase consumption and income one-for-one. Given a positive labor wedge, this step will increase the agent's utility. In the final step, we can extract resources to reset the agent's utility to its baseline level.

While Proposition 3 shows that a zero deductibility of time-enhancing investment is inefficient, the following result highlights that the opposite case of a full deductibility is inefficient too. Overall, these results imply that Pareto-efficient tax systems necessarily include imperfect deduction possibilities for time-enhancing investment.

Proposition 4 (Reducing a full tax deduction of time-enhancing investments) *Suppose that Assumption 1 holds. Starting from a tax system where time-enhancing investment is fully deductible and labor wedges are positive, i.e.,* $0 < \tau_d(n) = \tau_y(n) < 1$*, a Pareto-improving reform exists where type-n individuals invest less in their time endowment, work less and consume more.*

To understand Proposition 4, we construct a counterpart to the perturbation described above. Once more, we consider a truth-telling agent with skill n and a hypothetical shirker with skill $\hat{n} > n$. A tax system with a full deduction of time-enhancing investment can be improved in the following steps. First, we perform a joint reduction of income y and time-enhancing investment d by one marginal unit.[14] The utility of the truth-telling agent changes by

$$du = -\left[E' - \frac{1}{n}\right] \cdot \tilde{v}'\left(E(d) - \frac{y}{n}\right),$$

whereas the utility of the shirker changes by

$$d\hat{u} = -\left[E' - \frac{1}{\hat{n}}\right] \cdot \tilde{v}'\left(E(d) - \frac{y}{\hat{n}}\right).$$

A full deductibility means that the margin between income and time-enhancing investment is undistorted at the baseline tax system, i.e., we obtain $du = 0$.[15] For the shirking agent, however, the perturbation causes a first-order utility loss, $d\hat{u} < 0$, because reducing the earnings by a dollar does not bring as much extra leisure as for the truth-telling agent. Hence, in a second step, we can increase consumption and income by a small amount without violating the incentive-compatibility constraint. Given a positive labor wedge, this step will increase the agent's utility. Finally, we can extract some resources to reset the agent's utility to its baseline level.

4.2 A quantitative study of tax deductions for household services

Next, we apply the time-endowment model to assess the potential welfare consequences of introducing tax deductions for household services in the United States. Unlike some European countries (e.g., Germany, Sweden and Denmark), the current US tax code does not provide general tax breaks to households that hire professional services for housekeeping, cleaning, gardening, et cetera. In this section, we quantitatively evaluate Pareto-improving reforms that stimulate the consumption of such services through tax deductions.

[14] Intuitively, time-investment and reductions of labor supply are substitutable inputs for the production of leisure. The proposed perturbation replaces a unit of time-investment by an equivalent amount of reduced labor supply, holding constant the level of leisure. As a consequence, the work-related expenditure wedge falls, whereas the labor wedge remains constant.

[15] Formally, we have $\tau_d = \tau_y$ if and only if $u_d = -u_y$ if and only if $E'\tilde{v} = -\tilde{v}/n$.

4.2.1 Institutional background

Tax breaks for household service expenditures exist in a number of countries including Denmark, Germany and Sweden. The classification of household services that are eligible for a tax break is similar between those three countries and includes services such as cleaning, gardening and child care.

In Denmark, individuals can deduct up to 6,000 DKK (2018) per person per year for expenses on household services ("servicefradrag"). The deduction reduces the individual's taxable income, implying that the monetary value depends on the personal tax rate. By contrast, Germany and Sweden provide tax credits that directly reduce the income tax liability of the claimant. In Germany, expenses on household services are credited at a rate of 20 percent against the tax liability of the household ("Abzug für haushaltsnahe Dienstleistungen"). The maximum credit per year and household is 4,000 EUR (2018). In Sweden, expenses on household services are credited at a rate of 50 percent and the tax credit is limited to 25,000 SEK (2018) per person per year ("RUT-avdrag"). The marginal rate of the income tax (municipal, county and national income tax combined) in Sweden has three steps at approximately 32, 52 and 57 percent. Hence, the tax credit rate of 50 percent corresponds to a strong degree of deductibility of household service expenses and for individuals in the lowest tax bracket, in fact, exceeds a full deductibility.

In the United States, the "Child and Dependent Care Credit" provides a tax credit for the care of children (under age 13) and the care of dependents incapable of self-care. Eligible annual expenses are limited to $3,000 (for one qualifying person) or $6,000 (for two or more qualifying persons). For taxpayers with incomes greater than $43,000, the tax credit amounts to 20 percent of the care costs. For taxpayers with lower incomes, the rate can rise to 35 percent. Expenses for other household services qualify only if a part of the service is for the care of qualifying persons. For example, the tax credit is not given for standard cleaning or gardening services. However, the tax credit applies for the employment of a nanny that provides child care and completes ancillary household tasks. As an alternative to the "Child and Dependent Care Credit", some employees have access to a pre-tax dependent care account offered by their employer. The account can be used to deduct up to $5,000 per year from taxable income to pay for the care of children (under age 13) or dependents incapable of self-care. Once more, household services qualify for the tax break only if the service includes care for a qualifying person.

Summing up, for households without young children or other dependents in need of care, the US tax code does not grant a tax break for expenses on household services. In our quantitative evaluation, we will therefore focus on single, prime-age households.

4.2.2 Model specification

We consider a quasi-linear version of the time-endowment model with a utility function of the following form:

$$u\left(c,d,y;n\right)=c+\gamma\frac{\left(1-\left(\frac{\bar{s}-d}{\alpha_1}\right)^{\frac{1}{\alpha_2}}-\frac{y}{n}\right)^{1-\frac{1}{\phi}}}{1-\frac{1}{\phi}}$$

where the parameters $(\alpha_1, \alpha_2, \bar{s}, \gamma, \phi)$ are positive. As usual, market labor is given by $l = y/n$ in the above specification.

The utility function is derived from a household production model based on Sandmo (1990) and Kleven et al. (2000). Households produce domestic services with a concave technology: $d_n = \alpha_1 l_n^{\alpha_2}$, where l_n represents non-market work and (α_1, α_2) are parameters with $0 < \alpha_1$ and $0 < \alpha_2 < 1$. Household-produced domestic services d_n and domestic services d obtained from the market are perfect substitutes: $s = d + d_n$. Households have a fixed demand for domestic services $s = \bar{s}$ and a fixed time endowment $\bar{E} = 1$ that can be used for market work l, non-market domestic work l_n and leisure.

The household decision problem has a convenient closed-form solution. Given a marginal tax rate $\mathcal{T}'$ and a zero subsidy on household services, the first-order conditions of the household problem can be solved for household service expenditure d and labor supply l as follows:

$$d = \bar{s}-\left(\frac{\alpha_1^{\frac{1}{\alpha_2}}\alpha_2}{\left(1-\mathcal{T}'\right)n}\right)^{\frac{1}{\frac{1}{\alpha_2}-1}}, \tag{18}$$

$$l = 1-\left(\frac{\bar{s}-d}{\alpha_1}\right)^{\frac{1}{\alpha_2}}-\left(\frac{\gamma}{\left(1-\mathcal{T}'\right)n}\right)^{\phi}. \tag{19}$$

4.2.3 Calibration strategy

We calibrate the parameters $(\alpha_1, \alpha_2, \bar{s}, \gamma, \phi)$ by matching a set of moments. Intuitively, the parameters $(\alpha_1, \alpha_2, \bar{s})$ govern the cross-sectional expenditure pattern for household services. The parameter γ determines the time share of labor. Finally, the value of ϕ corresponds to the Frisch elasticity of *leisure* (holding non-market work fixed), which is one-to-one related to the Frisch elasticity of labor supply.[16] Consequently, we target the time share of labor (i.e., market work), the Frisch elasticity of labor supply, and the pattern of household service expenditures of US households by income. Table 1 summarizes the parameters and data moments of our calibration procedure.

Parameter	Target	Source	Data	Model
ϕ	Mean Frisch elasticity of labor	Chetty et al. (2013)	0.5	0.501
γ	Mean time share of labor	Standard (40h/week)	0.238	0.234
$\alpha_1, \alpha_2, \bar{s}$	Household service expenditure (by income group)	CEX 2015	See Figure 1	

Table 1: Calibration strategy

Our measure of household service expenditure is based on the Consumer Expenditure Survey (CEX) 2015 and includes housekeeping services, gardening services and a number of other household services.[17] Child care services are excluded in our definition. For the sake of comparability across households, we restrict the data set to single, prime-age households (age 25–54 years).[18] We group those households into eight income bins (less than 20k, 20-40k, 40-60k, 60-80k, 80-100k, 100-150k, 150-200k, more than 200k) based on their annual employment income in US$. For each income bin, we compute the means of income and household service expenditure in the CEX data. Mean household service expenditures increase with the income group (except between the two highest groups) and range from $360 to $1,590 in our sample (Figure 1).

[16] More precisely, the Frisch elasticity of labor equals $(1 - l_n - l)\phi/l$, where ϕ is the Frisch elasticity of leisure and $(1 - l_n - l)/l$ is the leisure/labor ratio.

[17] Our measure follows the CEX category 'other household expenses' and includes housekeeping services, gardening and lawn care services, coin-operated laundry and dry-cleaning, termite and pest control products and services, home security systems service fees, moving, storage, and freight expenses, repair of household appliances and other household equipment, repair of computer systems for home use, computer information services, reupholstering and furniture repair, rental and repair of lawn and gardening tools, and rental of other household equipment.

[18] We drop households with incomes below the federal poverty level.

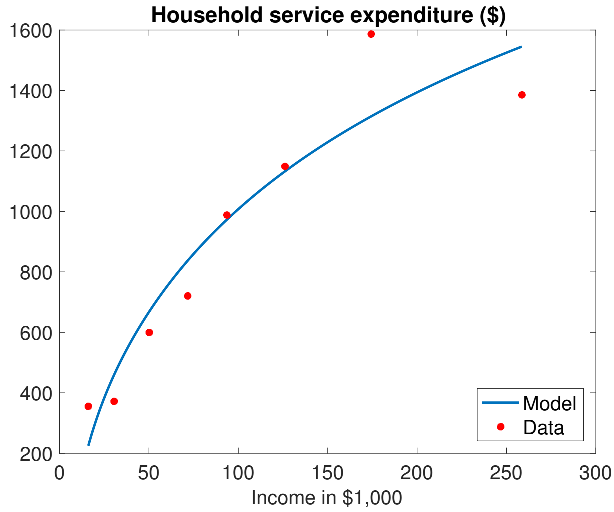


Figure 1: Household service expenditure in model and CEX data

For the baseline allocation, we assume that the households face marginal tax rates on labor income according to the parametric tax function of Gouveia and Strauss (1994),

$$\mathcal{T}' = b\left[1 - (s\hat{y}^p + 1)^{-\frac{1}{p}-1}\right],$$

where income $\hat{y}$ is measured in thousands of US dollars. Guner et al. (2014) provide recent estimates of these parameters for different specifications of the US tax system. We apply their parameters of the specification for unmarried households that includes state and local taxes. Hence, we set $b = 0.287$, $s = 0.006$, $p = 1.514$. In line with current tax practice in the US, we set the subsidy on household services to zero at the baseline allocation.

Our calibration proceeds in two steps. First, we construct a vector of skill types n that is consistent with the income groups of the CEX data. By substituting the solution for service expenditure (Eq. 18) into the solution for labor supply (Eq. 19), we obtain

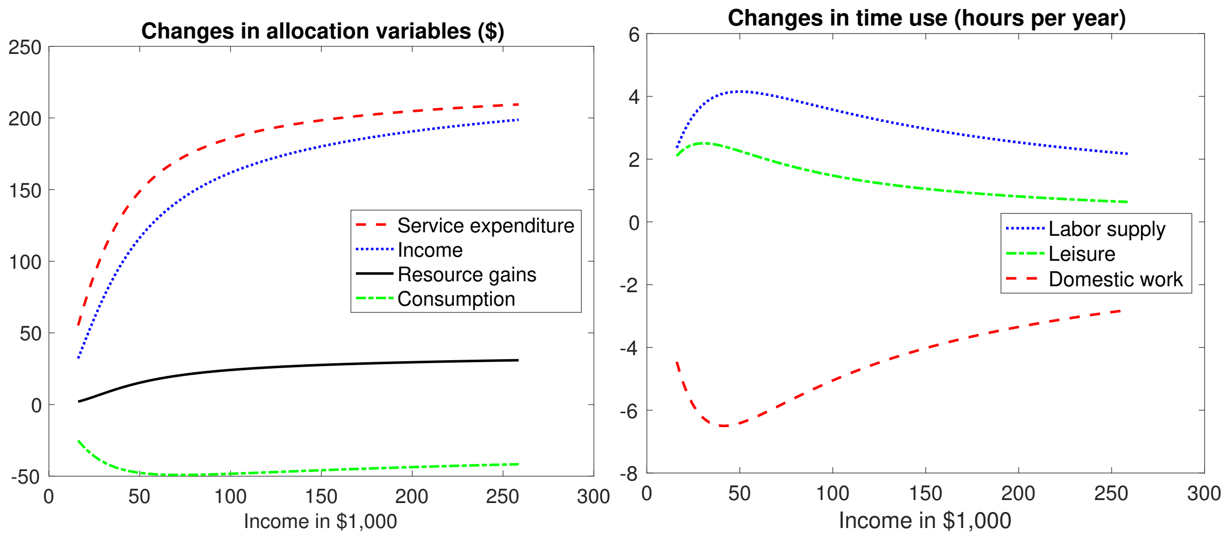
$$\frac{y}{n} = 1 - \left(\frac{\alpha_1 \alpha_2}{(1-\mathcal{T}')\, n}\right)^{\frac{1}{1-\alpha_2}} - \left(\frac{\gamma}{(1-\mathcal{T}')\, n}\right)^{\phi}.$$

For given parameters $(\alpha_1, \alpha_2, \bar{s}, \gamma, \phi)$ and the schedule of marginal tax rates $\mathcal{T}'$, this condition defines a nonlinear equation for skill n for every income level y, similar to the approach by Saez (2001). Based on this equation, we construct a skill vector such that the associated income levels match the incomes of the different groups in the CEX data. In the second step, we use the obtained skill vector, solve the individual decision problem and compare the model prediction with the calibration targets of Table 1. We calibrate the parameters $(\alpha_1, \alpha_2, \bar{s}, \gamma, \phi)$ by minimizing the distance to the calibration targets using standard numerical routines.

As shown by Table 1, our calibration matches the targets for the Frisch elasticity of labor supply and the time share of labor almost perfectly. We also obtain a reasonable fit for the expenditure pattern on household services (Figure 1).

4.2.4 Characterizing the Pareto-improving allocation reform

After solving for the baseline allocation, we explore a counterfactual Pareto-improving reform based on our theoretical results from Section 3 and maximize the resource gains of the reform. To approximate a continuous support of the skill distribution, we solve the model on a finer skill grid than the one used for the calibration procedure. Given that we start from a baseline allocation without deductions for household services, Proposition 3 implies that the reform induces households to work more and spend more on household services.



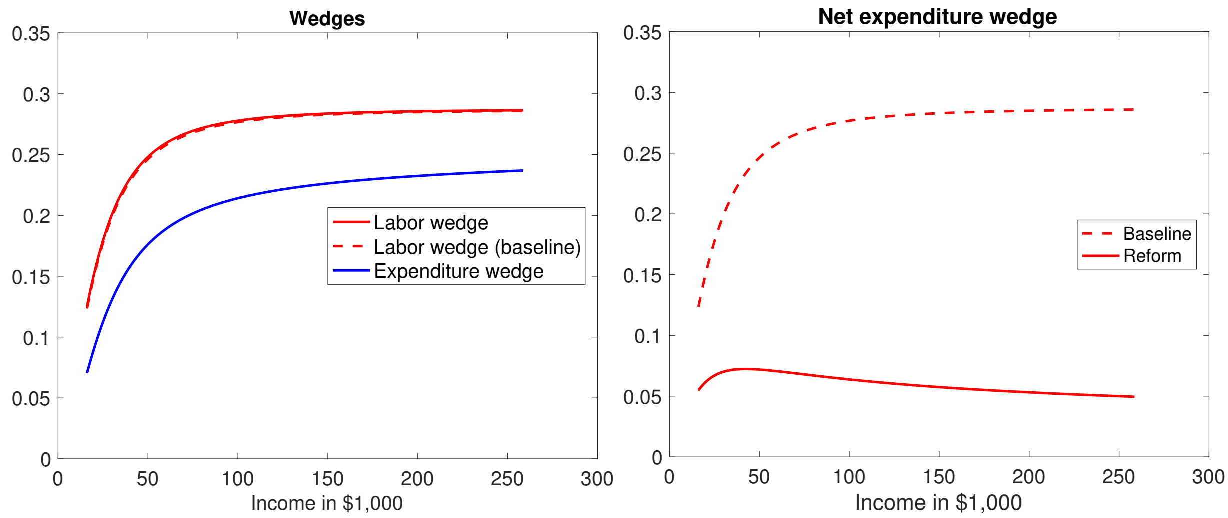
(a) Changes in allocation variables (b) Implied changes in time use

Figure 2: Pareto-improving reform

The changes in the allocation variables due to the Pareto-improving reform are illustrated in Figure 2. Evaluated at the median income level of US households in 2015 ($56,516), household service expenditures increase by $158 under the reform (roughly 1/4 of the baseline level). Income increases by $126 and other consumption falls by roughly $49, resulting in a resource gain of \$17 for a median household. Although the household-specific gains are moderate, aggregated across US households they imply annual gains in the range of several hundred million dollars.

Note that the pattern of the reform is exactly in line with Proposition 3; household services and income increase, whereas consumption decreases under the reform. Hence, the general idea of stimulating household services in order to facilitate labor supply in a situation with binding incentive constraints also applies to non-marginal reforms.

Figure 2b illustrates the time-use implications of the reform. Relative to a Pareto-efficient outcome, the baseline tax system induces households to work too much at home. The reform substitutes roughly 6 hours of annual household work by market services. The gained time is divided roughly 2 : 1 between market work and leisure.



(a) Labor and expenditure wedge

(b) Net wedge for household services

Figure 3: Wedges after Pareto-improving reform

The wedges for labor and household service expenditures are shown in Figure 3. Labor wedges remain largely unaffected by the reform, whereas the expenditure wedge increases notably from its baseline level of zero. The net expenditure wedges are relatively uniform across households and lie at approximately 6 percent (Figure 3b). That is, Pareto efficiency requires that expenses on household services are strongly deductible for all households.

4.2.5 A tax reform implementing the Pareto-efficient allocation

Finally, we also characterize a tax reform that implements the Pareto-efficient allocation. By the taxation principle, as described in Section 2.4, there exists a tax system implementing any incentive-compatible allocation. Next, we outline a particularly simple implementation of the Pareto-efficient allocation and numerically verify its validity.

We consider a deduction rule that depends only on the expenditure level, but not on income:

$$\mathcal{D}_y(d, y) = 0 \quad \forall\, d, y.$$

Having made this choice, Eqs. (3) and (4) imply that we have to set the marginal deduction rates and marginal income tax rates in the following way:

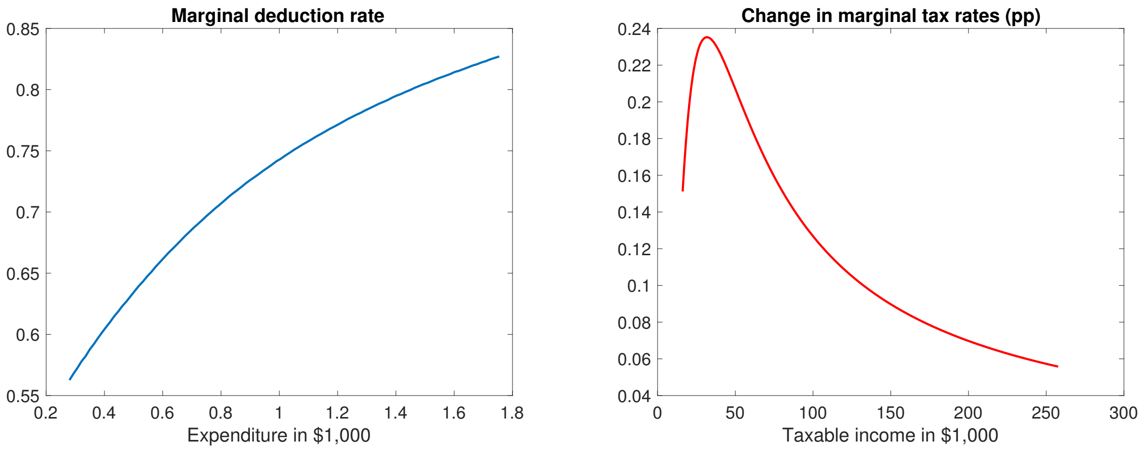
$$\mathcal{D}_d(d(n)) = \frac{\tau_d(n)}{\tau_y(n)} \quad \forall\, n,$$
$$\mathcal{T}'(y(n) - \mathcal{D}(d(n))) = \tau_y(n) \quad \forall\, n.$$

By construction, the first-order conditions of the individual choice problem are satisfied given these policy instruments. However, the open question is whether each individual bundle constitutes a global individual optimum. A theoretical analysis of this question is generally rather difficult (e.g., Renes and Zoutman, 2016). Therefore, our approach is to analyze the individual choice problem numerically.

It turns out that our proposed implementation is valid: individuals do choose the same (c, d, y) bundles as in the direct mechanism. Hence, this simple tax reform indeed implements the desired allocation. Figure 4a describes the marginal deduction rates $\mathcal{D}_d$ for the new tax system. They monotonically increase from 55% to 85% across the expenditure range. Figure 4b illustrates the change in marginal tax rates as a function of taxable income $y - \mathcal{D}(d)$. The increase in marginal tax rates is generally modest and never above 0.17 percentage points. This reflects the fact that housekeeping expenditures are generally at a rather low level. To sum up, Figure 4 shows that the Pareto-improving tax reform takes a very simple form here.

5 Conclusion

In this paper, we study the Pareto-efficient design of tax deductions for work-related goods. Our approach also provides guidelines on how to improve Pareto-inefficient tax systems. We quantitatively apply our method to assess the potential benefits of tax deductions for household



(a) Marginal deduction rates for household service expenditures

(b) Change in marginal tax rates

Figure 4: Tax reform implementing the Pareto-efficient allocation

services in the United States and find welfare gains that are non-negligible.

An added benefit may arise if deduction rules induce households to eliminate any informal employment of household helpers in order to qualify for tax benefits. In that case, the actual welfare gains from the introduction of deduction rules may be larger than the ones predicted by our model. Our analysis has also abstracted from general equilibrium effects. A possible direction for future research would be to explore effects of deduction rules on the tax-exclusive prices of deductible goods and services. Another interesting, but very challenging extension would be to explore the role of preference heterogeneity. For instance, households may differ in their productivity for domestic work. In such a framework, the utility function would explicitly depend on non-market work (in addition to the dependency on service expenditures), which creates a moral hazard problem on top of the private information problem for skills. We leave an extension along those lines for future research.

A Appendix

A.1 Proofs

Proof of Proposition 1. We extend the optimal control problem (12) to a framework where $d = (d_1, \ldots, d_K)$ and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_K)$ are K-dimensional real vectors. We set up the augmented

Hamiltonian

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

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

$$
f\left[1-v_{d_k}\frac{1}{w_c}-\frac{w_{d_k}}{w_c}\right]=\mu v_{nd_k} \tag{20}
$$

$$
f\left[1+v_y\frac{1}{w_c}\right]=-\mu v_{ny} \tag{21}
$$

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 $v_{nd_k}=0$, we obtain

$$
\tau_d^k=1-\frac{w_{d_k}+v_{d_k}}{w_c}=0. \tag{22}
$$

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

$$
\frac{w_c-(v_{d_k}+w_{d_k})}{w_c+v_y}=\frac{1-\frac{v_{d_k}+w_{d_k}}{w_c}}{1+\frac{v_y}{w_c}}=-\frac{v_{nd_k}}{v_{ny}}, \tag{23}
$$

which establishes Eq. (13). ■

Proof of Proposition 2. Under Assumption 1, the relationship between the cross-derivatives of leisure utility imply

$$
-\frac{v_{nd}}{v_{ny}}=\frac{\frac{y}{n^2}E'\tilde{v}''}{\frac{y}{n^3}\tilde{v}''-\frac{\tilde{v}'}{n^2}}=\frac{nE'}{1+e} \tag{24}
$$

where $e:=-\frac{n\tilde{v}'}{y\tilde{v}''}>0$. Moreover, using the definition of wedges we have

$$
\frac{1-\tau_d}{1-\tau_y}=\frac{E'\tilde{v}'}{\frac{\tilde{v}'}{n}}=nE'. \tag{25}
$$

Therefore, Eq. (14) implies

$$
\frac{\tau_d}{1-\tau_d}=\frac{\tau_y}{1-\tau_y}\frac{1}{1+e}. \tag{26}
$$

■

Proof of Proposition 3. We specialize the perturbation outlined in Section 3.1 for the time-endowment model. Suppose that time-enhancing investment d changes by an amount ε. Then, in order not to violate incentive compatibility, output y has to change by $\delta(\varepsilon)$ such that

$$\tilde{v}'\left(E(d) - \frac{y}{n}\right)\frac{y}{n^2} = \tilde{v}'\left(E(d+\varepsilon) - \frac{y+\delta(\varepsilon)}{n}\right)\frac{y+\delta(\varepsilon)}{n^2}. \tag{27}$$

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

$$w(c) + \tilde{v}\left(E(d) - \frac{y}{n}\right) = w(c+\gamma(\varepsilon)) + \tilde{v}\left(E(d+\varepsilon) - \frac{y+\delta(\varepsilon)}{n}\right). \tag{28}$$

The resource gain of the perturbation is $R(\varepsilon) = -\varepsilon + \delta(\varepsilon) - \gamma(\varepsilon)$. We are particularly interested in the marginal resource gain: $R'(\varepsilon) = -1 + \delta'(\varepsilon) - \gamma'(\varepsilon)$.

Implicit differentiation of (28) yields $\gamma'(\varepsilon) = -(1-\tau_d^\varepsilon) + \delta'(\varepsilon)(1-\tau_y^\varepsilon)$, which implies that the marginal resource gain can be written as:

$$R'(\varepsilon) = -\tau_d^\varepsilon + \delta'(\varepsilon)\tau_y^\varepsilon. \tag{29}$$

Implicit differentiation of (27) yields:

$$\delta'(\varepsilon) = \frac{1-\tau_d^\varepsilon}{1-\tau_y^\varepsilon}\frac{1}{1+e^\varepsilon}, \tag{30}$$

where e^ε is the Frisch elasticity of labor supply with respect to the net-of-tax rate $1-\tau_y$ (holding time investment fixed).[19] Formally, we have

$$e^\varepsilon = -\frac{\tilde{v}'\left(E(d-\varepsilon) - l^\varepsilon\right)}{\tilde{v}''\left(E(d-\varepsilon) - l^\varepsilon\right)l^\varepsilon}, \quad \text{where } l^\varepsilon = \frac{y+\delta(\varepsilon)}{n}. \tag{31}$$

[19] 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^ε 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.3 for further details.

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

$$R'(\varepsilon) = -\tau_d^\varepsilon + \tau_y^\varepsilon \frac{1-\tau_d^\varepsilon}{1-\tau_y^\varepsilon} \frac{1}{1+e^\varepsilon}. \tag{32}$$

Starting from an allocation with $0 = \tau_d < \tau_y < 1$, Equation (32) for the marginal resource gain takes the following form:

$$R'(0) = -\tau_d + \tau_y \frac{1-\tau_d}{1-\tau_y} \frac{1}{1+e} = \tau_y \frac{1}{1-\tau_y} \frac{1}{1+e} > 0. \tag{33}$$

Hence, we obtain a resource gain by marginally increasing d, i.e., choosing $\epsilon > 0$. Using $\delta'(0) > 0$ and $\gamma'(0) < 0$, we note that the associated change of income is positive, whereas the associated change of consumption is negative. ■

Proof of Proposition 4. Starting from an allocation with $0 < \tau_d = \tau_y < 1$, Equation (32) for the marginal resource gain takes the following form:

$$R'(0) = -\tau_d + \tau_y \frac{1-\tau_d}{1-\tau_y} \frac{1}{1+e} = -\tau_y \frac{e}{1+e} < 0. \tag{34}$$

Hence, we obtain a resource gain by marginally reducing d, i.e., choosing $\epsilon < 0$. Using $\delta'(0) > 0$ and $\gamma'(0) < 0$, we note that the associated change of income is negative, whereas the associated change of consumption is positive. ■

A.2 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, Eq. (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.3 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

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

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 fully 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.4 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 (nonlinear and nonseparable) income tax that depends on labor incomes and work-related expenses. Equivalently, there can be a labor income tax with a nonlinear, nonseparable deduction rule 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 $\bar{\mathcal{T}}(y, d) := \infty$ for any pair (y, d) that is not part of the incentive-feasible allocation.[20]

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,

$$(c(n), d(n), y(n)) \in \arg\max_{c,d,y} \; u(c, d, y; n) \;\text{ s.t. }\; c = y - d - \mathcal{T}(y, d). \tag{35}$$

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 (35) 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:

$$u\left(y - d - \mathcal{T}_n\left(y, d\right), d, y; n\right) = u\left(c(n), d(n), y(n); n\right) \quad \forall \left(y, d\right). \tag{36}$$

Note that this construction is possible if u is continuous and unbounded above and below in

[20]By incentive compatibility, if the pair $(y(n), d(n))$ is part of the allocation, the associated level of general consumption $c(n)$ is unique.

general consumption.[21] We know by construction that $\mathcal{T}_n(y(n), d(n)) = \bar{\mathcal{T}}(y(n), d(n))$, because otherwise Equation (36) 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 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 5 (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 $\mathcal{T} \geq \mathcal{T}^*$.*

Proof of Proposition 5. 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

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

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

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. ■

References

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

Bastani, S., S. Blomquist, and L. Micheletto (2017): "Child Care Subsidies, Quality, and Optimal Income Taxation," Mimeo.

Bovenberg, A. L. and B. Jacobs (2005): "Redistribution and education subsidies are Siamese twins," *Journal of Public Economics*, 89, 2005–2035.

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

Chetty, R., A. Guren, D. Manoli, and A. Weber (2013): "Does indivisible labor explain the difference between micro and macro elasticities? A meta-analysis of extensive margin elasticities," *NBER Macroeconomics Annual*, 27, 1–56.

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

CONGRESSIONAL BUDGET OFFICE (2013): "The Distribution of Major Tax Expenditures in the Individual Income Tax System," http://www.cbo.gov/sites/default/files/cbofiles/attachments/43768_DistributionTaxExpenditures.pdf.

——— (2016): "The Budget and Economic Outlook: 2016 to 2026," https://www.cbo.gov/publication/51129.

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

DOMEIJ, D. AND P. KLEIN (2013): "Should day care be subsidized?" *Review of Economic Studies*, 80, 568–595.

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.

GOUVEIA, M. AND R. P. STRAUSS (1994): "Effective federal individual income tax functions: An exploratory empirical analysis," *National Tax Journal*, 317–339.

GUNER, N., R. KAYGUSUZ, AND G. VENTURA (2014): "Income taxation of US households: Facts and parametric estimates," *Review of Economic Dynamics*, 17, 559–581.

——— (2017): "Child-Related Transfers, Household Labor Supply and Welfare," Mimeo.

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," Mimeo.

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.

Kleven, H. J., W. F. Richter, and P. B. Sørensen (2000): "Optimal taxation with household production," *Oxford Economic Papers*, 52, 584–594.

Koehne, S. (2018): "On the Taxation of Durable Goods," *International Economic Review*, 59, 825–857.

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 (2016): "Identifying Laffer Bounds: A Sufficient-Statistics Approach with an Application to Germany," *Scandinavian Journal of Economics*, 118, 646–665.

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

Olovsson, C. (2015): "Optimal taxation with home production," *Journal of Monetary Economics*, 70, 39–50.

Renes, S. and F. Zoutman (2016): "When a Price is Enough: Implementation in Optimal Tax Design," Working Paper.

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. (2001): "Using Elasticities to Derive Optimal Income Tax Rates," *Review of Economic Studies*, 68, 205–29.

Sandmo, A. (1990): "Tax distortions and household production," *Oxford Economic Papers*, 42, 78–90.

Stantcheva, S. (2017): "Optimal taxation and human capital policies over the life cycle," *Journal of Political Economy*, 125, 1931–1990.

Washington Post (2013): "Richest 20 percent get half the overall savings from U.S. tax breaks, CBO says," https://www.washingtonpost.com/business/economy/richest-20-percent-get-half-the-overall-savings-from-tax-breaks-cbo-says/2013/05/29/645f75c6-c894-11e2-9245-773c0123c027_story.html?noredirect=on&utm_term=.ae3b510adc90.

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

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