Optimal Income Tax Deductions for Mixed Business and Personal Expenditures*

Jacob Goldin Sebastian Koehne Nicholas Lawson

November 4, 2024

Abstract

We study the optimal taxation of mixed-purpose expenditures – i.e., expenditures that generate income while also serving a consumption function. We characterize the Pareto optimal income tax deduction for such expenditures within a generalized Atkinson-Stiglitz model. Pareto optimality requires a partial deduction for mixed-purpose expenditures, where the deduction rate depends on the fraction of an expenditure's marginal benefits that are attributable to income-generation rather than consumption. We extend our results to account for several practical considerations, including potential constraints relating to a uniform deduction rate or a fixed income tax schedule. Our results provide a rationale for non-uniform commodity taxation, distinct from existing models of preference heterogeneity or non-separability.

JEL codes: D82, H21

Keywords: Optimal Taxation, Tax Deduction

*Goldin: University of Chicago and NBER (jsgoldin@uchicago.edu); Koehne: Kiel University and CESifo (koehne@economics.uni-kiel.de); Lawson: Dalhousie University (nicholas.lawson@dal.ca). For helpful comments, we are grateful to Joe Bankman, Louis Kaplow, Dominik Sachs, David Weisbach, and seminar participants at Cornell, the University of Chicago, and the AMT Tax Conference.

1 Introduction

How should the tax system treat expenditures that are partly—but not entirely—motivated by the production of income? Mixed-purpose expenditures like these are pervasive in practice. Consider a small business owner who purchases a car for both professional and personal use; or a fancy home coffee machine that makes delicious coffee and also increases productivity at work; or an airplane ticket for a trip to meet with customers and visit family; or a fancy suit to attract clients and also to wear out-and-about around town. Reflecting on the taxation of such expenditures in 1938, Henry Simons wrote "There is here an essential and insuperable difficulty, even in principle."

In practice, the income tax rules governing the deductibility of mixed-purpose expenditures are varied and inconsistent. In the United States, for example, some are fully deductible (e.g., flying business class to meet a client), some are never deductible (e.g., purchasing a nice suit to wear at work), and some are partially deductible (e.g., business meals, limited to 50% of cost). In other cases, an expenditure's deductibility hinges on factors such as the taxpayer's primary purpose in undertaking the expense (e.g., travel expenses) or how the expense is used by the taxpayer in practice (e.g., home offices). In the words of one expert, "[I]f there is a rationale underlying these rules, it is most often unarticulated" (Halperin, 1973).

In this paper, we study the optimal taxation of mixed-purpose expenditures. The model we consider generalizes Atkinson and Stiglitz (1976) by allowing a commodity to enter into the taxpayer's budget constraint through its effect on pre-tax income (the income-production purpose), in addition to entering directly into the taxpayer's utility function (the consumption purpose). We assume that taxpayers' preferences are weakly separable between consumption and labor, and that taxpayers share the same subutility function for consumption. These assumptions rule out many of the rationales for non-uniform commodity taxation that have been studied in the public finance literature such as non-separability of preferences over consumption and labor effort, preference heterogeneity across taxpayers, or externalities (e.g., Christiansen, 1984; Saez, 2002, 2004; Kaplow, 2012).

We find that mixed-purpose goods provide an additional rationale for non-uniform commodity taxation, and we characterize the form that such tax treatment takes. We show that a necessary condition for Pareto optimality is the availability of a partial income tax deduction for expenditures on the mixed-purpose good. Under the optimal policy, the allowable deduction amount is equal to a fraction of total expenditures on the mixed-purpose good, with the fraction equal to the share of the total marginal benefits from additional expenditures on the good that stems from the production of income. Intuitively, if a taxpayer incurs an expense solely to generate income, Pareto optimality requires that the expenditure be undistorted by income taxation (Diamond and Mirrlees, 1971), and thus fully deductible from pre-tax income. Alternatively, if an expenditure represents personal consumption only, and is not undertaken to generate income, Pareto optimality requires that no income tax deduction be provided (Atkinson and Stiglitz, 1976). Between these two extremes, we find that the optimal policy takes a middle ground: the greater the extent to which the expenditure resembles a pure business input—for which the sole motivation is to produce income—the greater the share of the expenditure that is deductible, and vice-versa.

Although the optimal deduction rule we derive is intuitive, a number of its implications may not be immediately apparent. First, the optimal deduction rate for a mixed-purpose good will vary by taxpayer income if the ratio of income-producing to consumption benefits is different for high- versus low-income taxpayers. This can be implemented via phase-ins and phase-outs of deductions by income, which are a common feature of many tax systems. A related point is that the optimal deduction depends on the ratio of marginal rather than average benefits; for example, a taxpayer might require a basic phone plan for work, but marginal expenditures on the phone plan may be primarily consumption-motivated (such as an upgrade to obtain a higher quality camera). Finally, although we focus on goods that positively affect both income and utility, our results also apply to goods that positively enter utility but negatively affect income. The optimal deduction rate for such a good would be negative, which would entail calculating income tax based on pre-tax income plus some fraction of the taxpayer's expenditures on the good. We are not aware of real-world tax systems that incorporate such "additions to income" but the case for them follows directly from the basic Mirrleesian setup we assume.

We also consider extensions motivated by administrative practicalities or political constraints. For example, governments may need to provide the same deduction rate to all taxpayers, regardless of income or other sources of heterogeneity. Under such a constraint, we show that the optimal policy takes the form of a partial deduction, where the (uniform) deduction rate is equal to a weighted average of the taxpayer-specific deduction rates that would be optimal absent the uniformity constraint, and where the weights depend on taxpayers' marginal tax rates and elasticities of demand for the mixed-purpose good.

A second extension is motivated by the possibility that the range of available reforms is limited to adjustments to the deductibility of the mixed-purpose good, with the income tax schedule held fixed. For some institutional actors, such as a court or executive branch agency, this range of options may be more realistic than assuming that income tax rates can be freely adjusted. In this setting, we find that the optimal deduction needs to be modified to account for distributional concerns and labor supply effects associated with the mixed-purpose good. In our basic setup, both channels are (implicitly) addressed through adjustments to the schedule of income tax rates. But when that instrument is removed, the deduction rate becomes a useful tool for furthering distributional goals and provides additional fiscal benefits through its effect on labor effort.

Although fundamental to tax law and policy, prior research has not established the optimal tax treatment of mixed-purpose expenditures. The most closely related literature studies the taxation of expenditures that complement or substitute with labor effort, like child care, transportation or other services (Christiansen, 1984; Kleven, 2004; Kaplow, 2010; Bastani et al., 2019, 2020; Ho and Pavoni, 2020; Koehne and Sachs, 2022). By relaxing the Atkinson-Stiglitz separability assumption, these papers extend the Corlett and Hague (1953) intuition—that it can be optimal to tax complements to leisure and subsidize complements to labor—from a setting with linear commodity taxes to a Mirrleesian setup in which the planner can also raise revenue via a non-linear income tax. In contrast to the papers in this literature, we maintain weak separability of preferences for labor and other goods, while allowing pre-tax income to be determined by both labor effort and the mixed-purpose good. As we discuss below, models focused on preference non-separability generally cannot be reinterpreted to cover the income-generating process we consider. Consequently, our results shed new light on the optimal taxation of mixed-purpose expenditures.[1]

Distinct from this literature, Baake et al. (2004) also study an issue that is related to our focus: the inability of the tax authority to distinguish between two types of consumer expenditures – those that entirely constitute personal consumption versus those that contribute exclusively to income generation. Ideally, the government would provide no deduction for the former and a full deduction for the latter, but Baake et al. show that a partial deduction is optimal when the government is constrained to treat the two types of expenditure in the same manner. Although the motivation is related, the question we consider is quite different: the defining feature of our setting is the presence of a *single* good that contributes to both consumption and income. Consequently, in our model, in contrast to the model of Baake et al., consumers cannot separately adjust their expenditures on the mixed-purpose good to modify the good's income-producing or consumption-related characteristics. Another difference is that whereas Baake et al. characterize the signs of tax wedges under the optimal policy, they do not describe the optimal deduction rule, which we view as our main contribution.[2]

[1] As we discuss below, our framework can be cast into a model with a standard budget constraint and in which taxpayers have a specific form of non-separable preferences over consumption and *pre-tax income* (as opposed to labor effort or leisure). Viewed from that lens, our contribution is to characterize the optimal tax system associated with this specific form of preference non-separability, which prior work has not considered.

[2] Also related to our focus, Richter (2006) characterizes the efficient taxation of factors of production that generate non-taxable profit, which may be interpreted as utility from consumption, in a model with a representative agent and linear income taxation. We build on this result by considering a Mirrleesian setting with heterogeneous agents and non-linear income taxation, in which distributional considerations may enter into the social welfare function.

Outside of the economics literature, questions concerning the proper treatment of mixed-purpose expenditures have been debated for decades among tax lawyers and practitioners. Seminal treatments of the issue are provided in Halperin (1973) and Griffith (1993), which evaluate alternative deduction rules with respect to various policy goals, and a more recent discussion can be found in Givati (2020).[3] This literature highlights how an income tax system can avoid distorting pre-tax behavior by including only the personal portion of mixed-purpose expenditures in the tax base. Our findings build on this insight by establishing the conditions under which this efficient solution is in fact desirable within a conventional optimal tax setup where the planner also has distributional goals.

2 Setup

We extend the classic Mirrlees (1971) model by adding a mixed-purpose ("hybrid") good. The economy is populated by a continuum of individuals that differ with respect to their skill $w \in \mathcal{W} := [\underline{w}, \overline{w}] \subset \mathbb{R}_{++}$. The distribution of skill types in the economy is defined by a smooth probability density $f : \mathcal{W} \to \mathbb{R}_{++}$ with full support. The cumulative distribution function of this distribution is denoted by $F : \mathcal{W} \to [0, 1]$.

2.1 Preferences

Individuals have weakly separable preferences of the form

$$U(c, d, l, w) = u(v(c, d), l, w), \tag{1}$$

where c denotes regular consumption, d denotes the hybrid good, and l denotes labor effort. The sub-utility function $v : \mathbb{R}^2_+ \to \mathbb{R}$ is continuously differentiable, concave, and strictly increasing in both arguments. The outer utility function $u : \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}_{++} \to \mathbb{R}$ is continuously differentiable, concave, strictly increasing in its first argument, and strictly decreasing in its second argument. In this specification, the consumption sub-utility function is identical across individuals, while the overall utility function can vary by type, allowing for skill-specific labor preferences.

[3]The optimal deduction rate that we derive is close to what Griffith (1993) refers to as the allocation method for taxing mixed business and personal expenditures, and is also related to the efficient deduction rule derived by Givati (2020); one difference is that our result is tied to the share of total *marginal* benefits that are business-related, rather than the share of total *average* benefits. The tax law literature has also addressed the related issue of deductibility for expenditures like commuting that shape the disutility of labor supply; see, e.g., Klein (1968), Bittker (1973), and Hemel and Weisbach (2021).

2.2 Technology

An individual with skill w and labor supply l generates $x = wl$ units of labor income. Total income, $y = y(wl, d)$, depends on labor income and the hybrid good. The income production function $y : \mathbb{R}^2_+ \to \mathbb{R}_+$ is continuously differentiable with $y_x > 0$ and $y_d \geq 0$, where the subscripts represent partial derivatives with respect to the first and second argument. The cross derivative $y_{x,d}$ can be positive or negative. The income production function y is constant across individuals.

Income can be converted one-to-one into regular consumption. Income (or regular consumption) can be converted into the hybrid good d at rate $p > 0$.

2.3 Connection to Prior Literature

To clarify the relationship between our setting and the models that prior work has considered, we can define a virtual skill parameter

$$\tilde{w} := \phi(w, d, l) := \frac{y(wl, d)}{l}, \tag{2}$$

which yields an income process that appears multiplicative in labor effort: $y(wl, d) = \tilde{w}l$. In this transformation, the virtual skill level $\tilde{w}$ depends on the exogenous parameter w, labor effort l and consumption of the hybrid good d. In contrast, prior research in optimal taxation, including the work studying optimal commodity taxation with preference non-separability, generally takes skill to be exogenous (e.g., Atkinson and Stiglitz, 1976; Mirrlees, 1976; Christiansen, 1984).

As highlighted by the transformation in Equation (2), mixed-purpose goods can be reinterpreted as a non-standard model of endogenous skill formation. In this regard, our framework shares some similarities with models of optimal human capital policy. However, those models typically abstract from a consumption value of skills (e.g., Bohacek and Kapička, 2008; Findeisen and Sachs, 2016; Stantcheva, 2017) and often focus on specific skill formation processes, such as the multiplicative case $\tilde{w} = w \cdot \phi(d)$ explored by Bovenberg and Jacobs (2005).

Alternatively, one could define taxpayers' preferences over the consumption goods (c and d) and pre-tax income (y), rather than over the consumption goods and labor effort (l). Viewed through this lens, we can interpret the presence of a mixed-purpose good as giving rise to a special form of preference non-separability, in that an individual's (dis)utility associated with obtaining a given level of pre-tax income depends on the individual's choice of d. For this interpretation, we invert the income-generating process and express labor effort as $l(d, y, w) = g(y, d)/w$, where $g(y, d)$ is the inverse of $y(x, d)$ with respect to x for any given level of d. Then, we can express utilities as

$$\tilde{U}(c,d,y,w) = u\left(v(c,d), \frac{g(y,d)}{w}, w\right). \tag{3}$$

This interpretation shifts the effect of the mixed-purpose good d from the budget constraint to preferences. A model that is general enough to accommodate this approach is Koehne and Sachs (2022), if one is willing to impose the additional restriction that utility takes the additive form $v(c,d) + g(y,d,w)$. In addition to the difference in assumptions, that paper focuses on time-saving household services rather than mixed-purpose goods, and they do not derive any of our results. More generally, when mixed-purpose goods are incorporated into the model via preferences rather than the budget constraint, preferences obtain a specific form of skill dependence, which tends to limit the transparency and interpretability of the results. In contrast, our representation of mixed-purpose goods in terms of the income-generating function yields a particularly simple and intuitive condition for Pareto optimality, as we demonstrate below.

2.4 Allocations

An *allocation* is a triple $(c,d,x) : \mathcal{W} \to \mathbb{R}^3_+$. Allocation $(c(w), d(w), x(w))_{w\in\mathcal{W}}$ is *feasible* if it satisfies the following resource constraint:

$$\int_{\underline{w}}^{\overline{w}} \left[c(w) + p \cdot d(w) - y(x(w), d(w))\right] f(w)\,\mathrm{d}w \leq 0. \tag{4}$$

Our analysis will focus on allocations generated by tax systems that depend on gross income, y, and additionally on labor income, $x = wl$, or deductible expenditures on the hybrid good, pd. Moreover, it will sometimes be helpful to employ general methods of mechanism design, without resorting to specific implementations of the allocation. Assuming that the variables c, d, y and $x = wl$ are observable,[4] while skills w and labor supply l are not, incentive compatibility requires that individuals do not gain from choosing the commodity bundle (c,d) and labor income x of any other type. Thus, allocation $(c(w), d(w), x(w))_{w\in\mathcal{W}}$ is *incentive compatible* if it satisfies

$$u\left(v(c(w), d(w)), \frac{x(w)}{w}, w\right) \geq u\left(v(c(w'), d(w')), \frac{x(w')}{w}, w\right) \quad \text{for all } w, w' \in \mathcal{W}. \tag{5}$$

The allocation is *incentive feasible* if it is feasible and incentive compatible.

[4]If d is observable, then knowledge of the income production function $y(x,d)$ allows one to infer labor input wl based on total income y, or vice-versa. Therefore, it is sufficient to assume that at least one of the two income variables is observable.

Importantly, due to the separability of the utility function, the incentive compatibility condition (5) depends on commodity bundles $(c(w), d(w))$ only via the sub-utilities $v(c(w), d(w))$. Thus, any reform that leaves the sub-utilities unchanged will not affect the incentives to report truthfully.

2.5 Pareto Optima

An allocation is *constrained Pareto optimal* if it is incentive feasible and not Pareto dominated by any other incentive feasible allocation.

We assume throughout the analysis that the resource constraint (4) is binding in every constrained Pareto optimum. Intuitively, we are assuming that if resources are left over, the planner can design a reform that increases consumption without violating the incentive compatibility condition. In appendix A, we verify this claim for two benchmark utility functions. More broadly it is plausible to assume that the planner can distribute additional resources in some lump-sum way to make individuals better off without endangering incentive compatibility.

2.6 Decentralization

Our primary focus is on an income tax system (T, α) of the form

$$T(y(wl, d) - \alpha(wl)\,pd)\,. \tag{6}$$

In this specification, the income tax schedule $T : \mathbb{R} \to \mathbb{R}$ maps taxable incomes to tax levels and is continuously differentiable. Taxable income is defined as total income y minus the deduction for expenditures on the hybrid good, $\alpha(wl)\,pd$, where the deduction rate $\alpha : \mathbb{R}_+ \to \mathbb{R}$ is a continuously differentiable function of labor income. Later, we will consider alternative specifications where the deduction rate is a function of expenditures on the hybrid good, pd, or of total income, y.

Faced with a tax system of the form in (6), the individual problem is given by

$$\max_{c,d,l} u(v(c,d), l, w) \ \text{ s.t. } \ c + pd = y(wl, d) - T(y(wl, d) - \alpha(wl)\,pd). \tag{7}$$

This problem can be decomposed into two stages.

Stage 2: Commodity Choice. Given labor income wl, choose a combination (c, d) that maximizes sub-utility subject to the budget constraint:

$$v^*(wl; T, \alpha) = \max_{c,d} v(c, d) \text{ s.t. } c + pd = y(wl, d) - T(y(wl, d) - \alpha(wl)pd). \tag{8}$$

Equivalently (by substituting the budget constraint):

$$v^*(wl; T, \alpha) = \max_{d} v(y(wl, d) - T(y(wl, d) - \alpha(wl)pd) - pd, d). \tag{9}$$

This problem generates the following first-order condition:

$$0 = [y_d - (y_d - \alpha p)T' - p] v_c + v_d. \tag{10}$$

Note that stage 2 does not depend on the skill type w other than through the level of labor income wl.

Stage 1: Labor Supply. Given sub-utility $v^*(wl; T, \alpha)$, choose labor supply l to maximize utility

$$U^*(w; T, \alpha) = \max_{l} u(v^*(wl; T, \alpha), l, w). \tag{11}$$

Note that the labor supply problem depends on taxes, deductions and commodity choices only indirectly through the sub-utility $v^*(wl; T, \alpha)$. This property mirrors the structure of the incentive compatibility constraint (5) in the mechanism design formulation.

3 Main Result

In this section, we characterize optimal deduction rules for expenditures on the hybrid good. Our first step presents a necessary condition for Pareto optimality based on a mechanism design approach. We then explore the consequences of this condition for a specific tax system.

3.1 First-Best Rule

We consider two-dimensional allocation perturbations that change regular consumption c and the hybrid good d, while holding fixed the sub-utility derived from these commodities.

For a given allocation $(c(w), d(w), x(w))_{w\in\mathcal{W}}$, we pick arbitrary types $w_0, w_1 \in \mathcal{W}$ with $w_0 < w_1$ and an arbitrary function $\delta : [w_0, w_1] \to \mathbb{R}$ with $\delta(w_0) = \delta(w_1) = 0$. Using the scaling factor $\varepsilon \in \mathbb{R}$, we perturb the commodity levels as follows:

$$\hat{d}(w) = d(w) + \varepsilon \cdot \delta(w), \tag{12}$$

$$\hat{c}(w) = c(w) - \gamma(w, \varepsilon \cdot \delta(w)), \tag{13}$$

where $\gamma(w, \varepsilon \cdot \delta(w))$ is defined such that the sub-utility remains unchanged,[5] i.e.,

$$v(c(w) - \gamma(w, \varepsilon \cdot \delta(w)), d(w) + \varepsilon \cdot \delta(w)) = v(c(w), d(w)) =: \bar{v}(w). \tag{14}$$

Equation (14) ensures

$$v\left(\hat{c}(w), \hat{d}(w)\right) = v(c(w), d(w)) \text{ for all } w \in \mathcal{W}. \tag{15}$$

Therefore, if the original allocation $(c(w), d(w), x(w))_{w\in\mathcal{W}}$ is incentive compatible, so is the perturbed allocation $\left(\hat{c}(w), \hat{d}(w), x(w)\right)_{w\in\mathcal{W}}$.

A necessary condition for Pareto optimality is that no perturbation can reduce the resources needed for a given pattern of sub-utilities. Because the incentive structure is unaffected by such perturbations, this insight holds in the first-best world as well as in the second-best world of constrained optimality. The first-order condition for resource minimization is given in the following result.

Proposition 1 (First-Best Rule)**.** *If an allocation $(c(w), d(w), x(w))_{w\in\mathcal{W}}$ is constrained Pareto optimal, commodity choices satisfy the first-best rule*

$$p = \frac{v_d(c(w), d(w))}{v_c(c(w), d(w))} + y_d(x(w), d(w)) \tag{16}$$

for almost all skill types w.

Proof. First, we show by contradiction that constrained Pareto optimality implies resource minimization for a given pattern of sub-utilities $v(c, d)$ and labor incomes x. Suppose, to the contrary, that a given allocation is constrained Pareto optimal and does not minimize the resources for the given pattern of sub-utilities and labor incomes. Then, there exists an alternative allocation with identical sub-utilities and labor incomes that requires fewer resources. Since the sub-utilities and labor supplies are the same as in the original allocation,

[5] By inverting this condition, we can state the reduction of regular consumption alternatively as

$$\gamma(w, \varepsilon \cdot \delta(w)) = c(w) - v^{-1}(\bar{v}(w); d(w) + \varepsilon \cdot \delta(w)),$$

where v^{-1} is the inverse of sub-utility v with respect to its first argument.

the alternative allocation is also incentive compatible. It also implements the same utility levels, so it is also constrained Pareto optimal. However, since the original allocation was feasible, the resource constraint for the alternative allocation must be slack, which contradicts our assumption that the resource constraint is binding in every constrained Pareto optimum.

Now, we derive the first-order condition for resource minimization. (This step provides a direct proof of a basic version of the Euler-Lagrange equation.) If the given allocation minimizes the resources needed for a given pattern of sub-utilities, then $\varepsilon = 0$ is a solution of the following resource problem:

$$\min_{\varepsilon} \int_{w_0}^{w_1} \left[p\left(d\left(w\right) + \varepsilon\delta\left(w\right)\right) + c\left(w\right) - \gamma\left(w, \varepsilon\delta\left(w\right)\right) - y\left(x\left(w\right), d\left(w\right) + \varepsilon\delta\left(w\right)\right)\right] f\left(w\right) \mathrm{d}w.$$

The FOC evaluated at $\varepsilon = 0$ is

$$\int_{w_0}^{w_1} \left[p\delta\left(w\right) - \delta\left(w\right)\gamma_2\left(w, 0\right) - \delta\left(w\right) y_d\left(x\left(w\right), d\left(w\right)\right)\right] f\left(w\right) \mathrm{d}w = 0,$$

where γ_2 denotes the partial derivative of γ with respect to its second argument. Equivalently,

$$\int_{w_0}^{w_1} \left[p - \gamma_2\left(w, 0\right) - y_d\left(x\left(w\right), d\left(w\right)\right)\right] \delta\left(w\right) f\left(w\right) \mathrm{d}w = 0.$$

Importantly, the perturbation δ was arbitrary and the FOC is valid for every perturbation δ. Therefore, we must have

$$p - \gamma_2\left(w, 0\right) - y_d\left(x\left(w\right), d\left(w\right)\right) = 0 \text{ for almost all } w \in \left[w_0, w_1\right].$$

Differentiating Equation (14) with respect to $\varepsilon \cdot \delta\left(w\right)$, we obtain

$$-\gamma_2\left(w, \varepsilon \cdot \delta\left(w\right)\right) v_c\left(\hat{c}\left(w\right), \hat{d}\left(w\right)\right) + v_d\left(\hat{c}\left(w\right), \hat{d}\left(w\right)\right) = 0.$$

Hence,

$$\gamma_2\left(w, 0\right) = \frac{v_d\left(c\left(w\right), d\left(w\right)\right)}{v_c\left(c\left(w\right), d\left(w\right)\right)}.$$

Hence, we can express the optimality condition as

$$p - \frac{v_d\left(c\left(w\right), d\left(w\right)\right)}{v_c\left(c\left(w\right), d\left(w\right)\right)} - y_d\left(x\left(w\right), d\left(w\right)\right) = 0 \text{ for almost all } w \in \left[w_0, w_1\right].$$

Because w_0 and w_1 were arbitrary, the condition extends to the entire type space $[\underline{w}, \overline{w}]$. $\square$

The first-best rule of Proposition 1 is intuitive. Since the choice of goods is separated from the supply of labor by the weak separability of preferences, the trade-off between the hybrid good and regular consumption should not be distorted. This requires that the relative price of the hybrid good, p, corrected for its marginal income benefit, y_d, equals the marginal rate of substitution between these goods, v_d/v_c. Importantly, the correction of the relative price reflects the *gross* marginal income benefit of the hybrid good, y_d. Thus, to the extent that income is taxed, the tax code must provide compensating measures to support the demand for the hybrid good.

3.2 Consequence for the Optimal Deduction Rate

In the next step, we explore the implications of the first-best rule of Proposition 1 for specific tax systems. As discussed above, the first-best rule for commodity choice is compatible with income taxation only if the tax system subsidizes the hybrid good relative to general consumption.

We initially focus on tax functions for which the deduction rate is a function of labor income; further below, we consider additional tax functions for which the deduction rate depends on expenditures on the hybrid good or on total income. Suppose that an allocation $(c(w), d(w), x(w))_{w\in\mathcal{W}}$ is implemented by a tax function of the form $T\left(y - \alpha\left(x\right)pd\right)$, where $x = wl$ represents labor income. By rearranging Equation (10), the first-order condition for the individual commodity choice problem equals

$$y_d\left(x(w), d(w)\right) + \frac{v_d\left(c(w), d(w)\right)}{v_c\left(c(w), d(w)\right)} = p + \left(y_d\left(x(w), d(w)\right) - \alpha\left(x(w)\right)p\right)T'. \tag{17}$$

Equating this condition with the result of Proposition 1, we obtain

$$p = y_d\left(x(w), d(w)\right) + \frac{v_d\left(c(w), d(w)\right)}{v_c\left(c(w), d(w)\right)} = p + \left(y_d\left(x(w), d(w)\right) - \alpha\left(x(w)\right)p\right)T'. \tag{18}$$

Hence, if $T' \neq 0$, we obtain $y_d - \alpha p = 0$. Solving this condition for α, we obtain the following result.

Proposition 2 (Optimal Deduction Rate)**.** *If a tax function of the form $T\left(y - \alpha\left(x\right)pd\right)$ implements a constrained Pareto optimum and T' is never zero, then the deduction rate satisfies*

$$\alpha\left(x(w)\right) = \frac{y_d\left(x(w), d(w)\right)}{y_d\left(x(w), d(w)\right) + \frac{v_d(c(w),d(w))}{v_c(c(w),d(w))}} \tag{19}$$

for almost all skill types w.

Proof. Equation (18) implies $(y_d - \alpha p) \cdot T' = 0$. Since $T' \neq 0$, we obtain $y_d = \alpha p$. Equivalently, $\alpha = y_d/p$. Now the result follows from Proposition 1. $\square$

According to Proposition 2, the Pareto optimal deduction rate relates the marginal income benefit y_d to the marginal consumption benefit v_d/v_c of the good. The larger the income benefit relative to the consumption benefit, the larger is the rate of deduction. More specifically, Proposition 2 implies that under the Pareto optimal tax function, a marginal change in the hybrid good d does not affect taxable income:

$$\frac{\partial \left(y\left(x,d\right) - \alpha\left(x\right) pd\right)}{\partial d} = y_d - \alpha p = y_d - \frac{y_d}{y_d + \frac{v_d}{v_c}} p = 0, \tag{20}$$

where the last equality follows from Proposition 1. Put differently, the optimal deduction rate excludes income from good d from the income tax base at the margin.

Another way to understand Proposition 2 is in relation to the two polar settings considered by Atkinson and Stiglitz (1976) and Diamond and Mirrlees (1971). In particular, Atkinson and Stiglitz (1976) focus on weakly separable preferences without any hybrid good. In their framework, income is determined solely by skills and labor supply, i.e., $y\left(wl, d\right) = wl$ for all $\left(wl, d\right) \in \mathbb{R}^2_+$. They show that within this model, the optimal combination of income and commodity taxes does not distort the relative prices of the available commodities. Because an income tax deduction for a commodity lowers the relative price of that commodity, their result implies that any (non-uniform) deduction would be sub-optimal. When we apply our Proposition 2 to the special case in which the available commodities are pure consumption goods—i.e., d does not enter into the income production function so that $y_d = 0$—we obtain the same result: Pareto optimality implies that good d should not be deductible ($\alpha = 0$).

At the other extreme, Diamond and Mirrlees (1971) consider the taxation of pure income-generating goods, which do not directly enter into utility, i.e., $v(c, d) = c$ for all $(c, d) \in \mathbb{R}^2_+$. Within this model, they show that Pareto optimality requires that the real cost to the taxpayer of consuming intermediate goods like d should not be distorted by taxation. Because income taxation raises the cost of purchasing any commodity with after-tax dollars, their result implies that expenditures on intermediate goods should be fully deductible from the income tax base. When our Proposition 2 is applied to this special case, in which $v_d = 0$, it yields the same result, i.e., $\alpha = 1$.

An important practical consideration in implementing Proposition 2 is that the optimal deduction rate depends on the relative importance of consumption versus income-generating motives, but policymakers cannot directly observe the strength of these alternative channels. For settings in which the shape of the income-generating function $y\left(\cdot\right)$ is known, this challenge can be alleviated by expressing the optimal deduction rate from Proposition 2 as

$$\alpha\left(x(w)\right) = \frac{y_d\left(x(w), d(w)\right)}{p}, \tag{21}$$

where the equivalence follows from taxpayers' first order conditions under the optimal policy, reflected in Equations (10) and (18).

Finally, although our focus has been on settings in which the mixed-purpose good positively contributes to income, Proposition 2 also sheds light on the optimal tax treatment of goods that generate consumption utility but that negatively affect labor income, i.e., for which $y_d < 0$. For expenditures on goods of this type, Proposition 2 suggests that the optimal tax treatment consists of a negative deduction—an extra inclusion—into taxable income. For example, consider a recreational drug that reduces a user's labor output; our framework suggests a rationale for adding a percentage of taxpayers' expenditures on this drug into their taxable income.

4 Extensions

This section extends our results to alternative tax functions and scenarios where the planner faces additional constraints on the instruments that can be used.

4.1 Alternative Tax Functions

In the prior section, we considered tax functions for which the deduction rate is a function of labor income. Here, we extend our analysis to two additional parameters on which the deduction rate could potentially depend.

4.1.1 Deduction Rate Depending on Expenditures

Deduction rates that depend on the expenditure level pd give rise to a different relevant class of tax functions. The individual commodity choice problem for a tax function of the form $T\left(y - \alpha\left(pd\right)pd\right)$ is given by

$$\max_d v\left(y\left(x,d\right) - T\left(y\left(x,d\right) - \alpha\left(pd\right)pd\right) - pd, d\right). \tag{22}$$

By equating the associated individual first-order condition with Proposition 1, we obtain

$$p = y_d\left(x(w), d(w)\right) + \frac{v_d\left(c(w), d(w)\right)}{v_c\left(c(w), d(w)\right)} = p + \left(y_d\left(x(w), d(w)\right) - p\alpha\left(pd(w)\right) - p^2 d(w)\alpha'\left(pd(w)\right)\right)T'. \tag{23}$$

Hence, if $T' \neq 0$, the deduction rate is characterized by the following equation:

$$p\alpha\left(pd(w)\right) + p^2 d(w)\alpha'\left(pd(w)\right) = y_d\left(x(w), d(w)\right). \tag{24}$$

Integration of this equation yields the following result.

Proposition 3. *If a tax function of the form $T\left(y - \alpha\left(pd\right)pd\right)$ implements a constrained Pareto optimum in which d is differentiable and T' is never zero, then the deduction rate satisfies*

$$\alpha\left(pd\left(w\right)\right) = \frac{\int y_d\left(x(w), d(w)\right)d'(w)\mathrm{d}w}{pd\left(w\right)} \tag{25}$$

for almost all skill types w.

Proof. Assuming that d is differentiable with respect to skill, multiplication of Equation (24) by $d'(w)$ yields

$$p\alpha\left(pd(w)\right)d'(w) + p^2 d(w)\alpha'\left(pd(w)\right)d'(w) = y_d\left(x(w), d(w)\right)d'(w).$$

Differentiating the product $pd\left(w\right)\alpha\left(pd\left(w\right)\right)$ with respect to w yields

$$\frac{\mathrm{d}\left[pd\left(w\right)\alpha\left(pd\left(w\right)\right)\right]}{\mathrm{d}w} = p\alpha\left(pd\left(w\right)\right)d'(w) + p^2 d\left(w\right)\alpha'\left(pd\left(w\right)\right)d'(w).$$

Combining the last two results, we obtain

$$\frac{\mathrm{d}\left[pd\left(w\right)\alpha\left(pd\left(w\right)\right)\right]}{\mathrm{d}w} = y_d\left(x(w), d(w)\right)d'(w).$$

Integration yields

$$pd\left(w\right)\alpha\left(pd\left(w\right)\right) - C = \int y_d\left(x(w), d(w)\right)d'(w)\mathrm{d}w$$

where C is an arbitrary constant. Hence,

$$\alpha\left(pd\left(w\right)\right) = \frac{\int y_d\left(x(w), d(w)\right)d'(w)\mathrm{d}w + C}{pd\left(w\right)}.$$

Finally, note that the constant C represents a level shift of taxable income:

$$y\left(x(w), d(w)\right) - \alpha\left(pd(w)\right) pd(w) = y\left(x(w), d(w)\right) - \int y_d\left(x(w), d(w)\right) d'(w)\mathrm{d}w - C.$$

Hence, C can be set to zero if the outer tax function T is suitably adjusted. □

In Proposition 3, the deduction rate relates the *cumulative* marginal income benefit of the hybrid good to the expenditure on that good. Using a change of variables, the deduction rate in Proposition 3 can be equivalently stated as:

$$\alpha\left(pd\left(w\right)\right) = \frac{\int_{d(\underline{w})}^{d(w)} y_d\left(x \circ d^{-1}(d), d\right) \mathrm{d}d}{pd\left(w\right)}, \tag{26}$$

where d^{-1} represents the inverse of d with respect to w. Again, the hybrid good is removed from the tax base at the margin. Note that a small increase in the hybrid good yields y_d units of additional income and simultaneously increases the deduction level, $\alpha\left(pd\right)pd$, by y_d units so that taxable income remains unchanged.

To relate this result to Proposition 2, we can use the first-best rule (16) to express the deduction rate in Proposition 3 as follows:

$$\alpha\left(pd\left(w\right)\right) = \frac{\frac{1}{d(w)} \int y_d\left(x(w), d(w)\right) d'(w)\mathrm{d}w}{y_d\left(x\left(w\right), d\left(w\right)\right) + \frac{v_d(c(w), d(w))}{v_c(c(w), d(w))}}. \tag{27}$$

Hence, Proposition 3 differs from Proposition 2 only in the sense that the local level of the marginal income benefit y_d is replaced by its average level, where the average is taken over all taxpayers with lower incomes.

Additive Separability. If the income-generating process is additively separable, i.e., if $y(x, d) = g(x) + h(d)$ for some functions g and h, Proposition 3 takes a particularly intuitive form and removes the deductible good from the tax base in both the marginal and non-marginal sense. Specifically, we have

$$\int y_d\left(x(w), d(w)\right) d'(w)\mathrm{d}w = \int h'\left(d(w)\right) d'(w)\mathrm{d}w = h\left(d(w)\right), \tag{28}$$

and hence

$$y\left(x(w), d(w)\right) - \alpha\left(pd(w)\right) pd(w) = g\left(x(w)\right), \tag{29}$$

which shows that taxable income in this case is completely determined by the level of labor income $x(w)$.[6]

[6] Analogous to the proof of Proposition 3, the additive constant in the indefinite integral can be set to zero without loss of generality.

4.1.2 Deduction Rate Depending on Total Income.

Finally, by inverting the relationship $y\left(x(w), d(w)\right)$ with respect to w, we can extend Proposition 3 to deduction rates defined in terms of incomes rather than expenditures. Specifically, let y^{-1} be the inverse of y with respect to w, i.e., let $y^{-1}\left(y\left(x(w), d(w)\right)\right) = w$. Then, implementation of a constrained Pareto optimum implies a deduction rate that satisfies

$$\alpha\left(y\left(x(w), d(w)\right)\right) = \frac{1}{pd\left(w\right)} \int_{d(\underline{w})}^{d\left(y^{-1}\left(y\left(x(w), d(w)\right)\right)\right)} y_d\left(x \circ d^{-1}(d), d\right) \mathrm{d}d. \tag{30}$$

This construction mimics Proposition 3 and ensures once more that the hybrid good is removed from the tax base at the margin. See Appendix B for further details on the derivation of this result.

4.2 Uniform Deduction Rate

This subsection considers the optimal deduction rule when the deduction rate, α, is constrained to be uniform across taxpayers – i.e., when all taxpayers can deduct a constant percentage of their expenditures on the hybrid good. Under this restriction, the tax system is defined by

$$T\left(y\left(wl, d\right) - \alpha \cdot pd\right). \tag{31}$$

As above, T is a mapping $T : \mathbb{R} \to \mathbb{R}$, but we now assume $\alpha \in \mathbb{R}$. This restriction may be applicable when factors outside of our model, such as societal fairness norms or a preference for simplicity, constrain the degree to which the tax system can provide different deduction rates across taxpayers.

To shed light on the optimal deduction in this setting, we adapt the proof strategy employed by Kaplow (2006, 2012, 2020). In particular, we will derive necessary conditions for Pareto optimality by considering a set of reforms to the tax system, and characterizing conditions under which such reforms yield a Pareto improvement. By construction, the reforms leave taxpayers' well-being unchanged, but may increase or decrease the amount of tax revenue collected. When a reform results in no change in well-being but an increase in tax revenue, the additional revenue can be refunded by a uniform lump-sum transfer, yielding a Pareto improvement, and vice-versa when a reform results in a tax revenue reduction.[7] We will therefore identify necessary conditions for Pareto optimal tax systems by characterizing when welfare-neutral but revenue-increasing reforms are possible. Different from our previous approach, the reforms are devised at

[7] See Kaplow (2006) for a formal proof.

the tax system level rather than at the allocation level.

Applying this approach allows us to establish the optimal uniform deduction rate as a weighted average of the optimal type-specific rates. The weights for each type are based on the marginal tax rate and sensitivity to the deduction rate; the larger the weight, the larger the fiscal externality from additional tax revenue generated by the reform.[8]

Proposition 4. *If the tax function is of the form $T(y - \alpha \cdot pd)$ with a constant deduction rate α, then the Pareto optimal deduction rate satisfies*

$$\alpha = \int_{\underline{w}}^{\overline{w}} \gamma_w \frac{y_d\left(x(w), d(w)\right)}{p} \mathrm{d}F(w) \tag{32}$$

where the weights γ_w are given by

$$\gamma_w = \frac{T'\left(z(w)\right) \frac{\mathrm{d}d(w)}{\mathrm{d}\alpha}}{\int_{\underline{w}}^{\overline{w}} T'\left(z(w)\right) \frac{\mathrm{d}d(w)}{\mathrm{d}\alpha} \mathrm{d}F(w)}, \tag{33}$$

with $z(w) = y\left(x(w), d(w)\right) - \alpha \cdot pd(w)$ representing taxable income.

Proof. See appendix C. □

This result is an intuitive extension of our main result from Proposition 2. The formula for the optimal individual-specific deduction rate, which equals

$$\alpha = \frac{y_d}{p} = \frac{y_d}{y_d + \frac{v_d}{v_c}} \tag{34}$$

under the conditions of Proposition 2, is now weighted by the γ_w term and integrated across the distribution. Thus, the optimal constant deduction rate corresponds to a weighted average of the optimal deduction rates across the entire distribution.

4.3 Fixed Income Tax Schedule

In our main analysis, we consider perturbations that hold the sub-utility from c and d fixed by considering joint changes in the deduction parameter α that can be offset by a suitable change in the income tax schedule. However, in some real-world settings, policymakers may not be able to adjust the income tax schedule due

[8] In the definition of the weights γ_w, the derivative $\mathrm{d}d(w)/\mathrm{d}\alpha$ represents the joint marginal effect on consumption of the hybrid good induced by a marginal increase of the deduction rate accompanied by welfare-neutralizing change of the tax level.

to institutional or political constraints. In this subsection, we consider the optimal $\alpha(wl)$ function for some fixed and potentially sub-optimal income tax schedule T.

Rather than considering two-dimensional perturbations as we did throughout the previous analysis, we now consider a perturbation to the α function while holding the tax schedule fixed. For this reason, similar to the scenario with a uniform deduction rate presented above, we will work directly with the tax function rather than with allocations. Our approach adapts the popular perturbation method by Saez (2001) to the case of deductions.

Specifically, we start from a given deduction rate $\alpha(wl)$ and suppose that α increases by $\Delta\alpha$ for a small interval of types $[w^*, w^* + \Delta w)$. We consider the impact of this perturbation on social welfare defined as

$$\int_w \pi\left(w\right) U\left(c\left(w\right), d\left(w\right), l\left(w\right), w\right) \, \mathrm{d}F\left(w\right),$$

where $\pi(w) > 0$, for $w \in \mathcal{W}$, are exogenous welfare weights, and $l\left(w\right) = x\left(w\right)/w$ denotes labor effort. Different from our previous analysis, we do not impose the weak separability condition specified in Equation (1).

The welfare-maximizing deduction rate is described in the following proposition.

Proposition 5. *If the tax function is of the form $T\left(y - \alpha\left(wl\right) pd\right)$ and T' is never zero, then the welfare-maximizing deduction rule for an exogenously fixed (and potentially sub-optimal) function T satisfies*

$$\alpha\left(wl(w)\right) = \frac{y_d \frac{\mathrm{d}d}{\mathrm{d}\alpha} + y_x w \frac{\mathrm{d}l}{\mathrm{d}\alpha} + \left(\frac{\pi U_c}{\lambda} - 1\right) pd}{p \frac{\mathrm{d}d}{\mathrm{d}\alpha}}. \tag{35}$$

for almost all skill types w. In this expression, λ is the value of a dollar to the government.

Proof. See appendix D. □

Proposition 5 provides an intuitive extension of Proposition 2 to settings where labor supply and welfare effects of deduction reforms are not neutralized. The optimal deduction rate is obtained as the ratio of marginal social benefit and cost. The benefit measures the positive fiscal externality of the deduction rate through income levels (via labor supply and use of the hybrid good) plus a (possibly negative) distributional term that depends on the social welfare weight $\pi u_v v_c/\lambda$ of the individuals affected by the rate. The cost measures the negative fiscal effect through the amount deductible. Note that, if the deduction reform were accompanied by a neutralizing change in the tax level, the labor supply effects and the distributional effects would vanish and Equation (35) would condense to the simple form of Equation (19).

5 Conclusion

We have characterized the Pareto optimal income tax deduction for mixed-purpose expenditures. The optimal formula reflects a mix of the familiar Atkinson-Stiglitz and Diamond-Mirlees intuitions in proportion to the degree to which marginal expenditures on the mixed-purpose good are motivated by consumption or income-production purposes. It is striking that the optimal formula has the simple form that it does—based on the relative magnitudes of these two motivations—given the lack of structure we impose regarding how the mixed-purpose good generates income, such as allowing for flexible interactions between the mixed-purpose good and labor income. The result confirms some existing policy intuitions while also highlighting some new ones, such as a novel rationale for including in taxable income some portion of taxpayers' expenditures on commodities that they purchase for consumption but that suppress their labor income.

Throughout, we have maintained several implicit assumptions that bear emphasizing. First, we have assumed that the income generated by the mixed-purpose good enters into the income tax base. If the mixed purpose good generates nontaxable income, the rationale for providing the deduction would disappear, since there would no longer be any fiscal externality from taxpayers' private consumption decisions.[9] Second, we have focused on mixed-purpose expenditures that self-employed individuals purchase to generate their own income, but a related set of questions arise regarding the optimal tax treatment of fringe benefits provided by an employer to an employee (e.g., Butler and Calcott, 2018; Leite, 2024). Third, we have assumed that the mixed-purpose good generates new income rather than simply reallocating income from one taxpayer to another. If instead the main function of a mixed-purpose expenditure was to reallocate rents, the case for a deduction would be weakened, and the fiscal externality would depend in part on the relative marginal tax rates of the taxpayers with respect to whom the income was being redistributed.

Finally, our goal has been to provide analytical clarity regarding the optimal design of deductions for mixed-purpose expenditures. However, there is a sense in which our results correspond to an idealized version of this policy. In particular, the optimal deduction rule depends on the relative importance of the different considerations that motivate taxpayers, but policymakers cannot peer into taxpayers' minds to directly observe the strength of these various motivations. When the shape of $y(\cdot)$ is known, one potential solution is to set the deduction rate in Proposition 2 to $\alpha = y_d/p$. Alternatively, the tax authority may address ambiguity of this form through case-by-case determinations based on observable signals of taxpayers' internal

[9] To illustrate, in the United States, capital gains from the sale of one's primary residence are non-taxable (up to some limit). A taxpayer who engages in an enjoyable wood-working project on her house, expecting to sell the house at a higher value because of the improvements, would not be entitled to any deduction for the purchase of her tools. See Givati (2020) for a discussion of this point.

motivations—an approach that is administratively costly but common in some areas of tax law (Bankman et al., 2023). A third approach, also common in practice, is for the tax law to take the form of a broad rule that provides the same treatment to all taxpayers, along the lines of the deduction rule in Proposition 4. Extending our results to account for multi-dimensional heterogeneity among taxpayers of the same ability with respect to preferences for the mixed-purpose good, or with respect to how the mixed-purpose good enters the income production function, is a promising avenue for future research.

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.

Bankman, J., D. N. Shaviro, K. J. Stark, and E. A. Scharff (2023): *Federal income taxation*, Aspen Publishing.

Bastani, S., S. Blomquist, and L. Micheletto (2019): "Nonlinear and piecewise linear income taxation, and the subsidization of work-related goods," *International Tax and Public Finance*, 26, 806–834.

——— (2020): "Child care subsidies, quality, and optimal income taxation," *American Economic Journal: Economic Policy*, 12, 1–37.

Bittker, B. I. (1973): "Income tax deductions, credits, and subsidies for personal expenditures," *The Journal of Law and Economics*, 16, 193–213.

Bohacek, R. and M. Kapička (2008): "Optimal human capital policies," *Journal of Monetary Economics*, 55, 1–16.

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

Butler, C. and P. Calcott (2018): "Optimal fringe benefit taxes: the implications of business use," *International Tax and Public Finance*, 25, 654–672.

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

Diamond, P. A. and J. A. Mirrlees (1971): "Optimal Taxation and Public Production I: Production Efficiency," *American Economic Review*, 61, 8–27.

Findeisen, S. and D. Sachs (2016): "Education and optimal dynamic taxation: The role of income-contingent student loans," *Journal of Public Economics*, 138, 1–21.

Givati, Y. (2020): "Theories of Tax Deductions: Income Measurement versus Efficiency," *Journal of Law, Finance, and Accounting*, 5, 107–136.

Griffith, T. D. (1993): "Efficient taxation of mixed personal and business expenses," *UCLA L. Rev.*, 41, 1769.

Halperin, D. I. (1973): "Business Deduction for Personal Living Expenses: A Uniform Approach to an Unsolved Problem," *U. Pa. L. Rev.*, 122, 859.

Hemel, D. J. and D. A. Weisbach (2021): "The Behavioral Elasticity of Tax Revenue," *Journal of Legal Analysis*, 13, 381–438.

Ho, C. and N. Pavoni (2020): "Efficient Child Care Subsidies," *American Economic Review*, 110, 162–99.

Kaplow, B. L. (2012): "Optimal control of externalities in the presence of income taxation," *International Economic Review*, 53, 487–509.

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

——— (2010): "Taxing leisure complements," *Economic Inquiry*, 48, 1065–1071.

——— (2020): "A unified perspective on efficiency, redistribution, and public policy," *National Tax Journal*, 73, 429–472.

Klein, W. A. (1968): "Income Taxation and Commuting Expenses Tax Policy and the Need for Nonsimplistic Analysis of Simple Problems," *Cornell L. Rev.*, 54, 871.

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

Koehne, S. and D. Sachs (2022): "Pareto-improving reforms of tax deductions," *European Economic Review*, 148, 104214.

Leite, D. (2024): "The Firm as Tax Shelter: Micro Evidence and Aggregate Implications of Consumption Through the Firm," Unpublished Paper, Paris School of Economics.

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

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

Richter, W. F. (2006): "Efficiency effects of tax deductions for work-related expenses," *International Tax and Public Finance*, 13, 685–699.

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

——— (2002): "The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes," *Journal of Public Economics*, 83, 217–230.

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

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

Appendix

A Binding Resource Constraint

Lemma 1 (Binding Resource Constraint)**.** *If preferences are additively separable, i.e., if there exists continuously differentiable and strictly increasing functions g, h with*

$$u\left(v\left(c,d\right),l,w\right) = g\left(v\left(c,d\right)\right) - h\left(l,w\right) \text{ for all } (c,d,l),$$

then the resource constraint (4) *is binding in every constrained Pareto optimum. The same result holds for multiplicatively separable preferences.*

Proof. First, consider the additive specification

$$u\left(v\left(c,d\right),l,w\right)=g\left(v\left(c,d\right)\right)-h\left(l,w\right).$$

For a given allocation $(c(w),d(w),x(w))_{w\in\mathcal{W}}$, consider the modified allocation $(c(w)+\gamma(w),d(w),x(w))_{w\in\mathcal{W}}$ that increases the consumption of every type by a small amount $\gamma\left(w\right)$ determined by the condition

$$g\left(v\left(c\left(w\right)+\gamma\left(w\right),d\left(w\right)\right)\right)=g\left(v\left(c\left(w\right),d\left(w\right)\right)\right)+\delta$$

for some $\delta>0$. If the resource constraint is slack at the given allocation, the modified allocation is feasible provided that δ is sufficiently small. By the monotonicity of the utility function, the modified allocation Pareto improves the given allocation. Moreover, if the given allocation is incentive compatible, then so is the modified allocation. To see this, note that incentive compatibility holds if and only if

$$g\left(v\left(c\left(w\right),d\left(w\right)\right)\right)-g\left(v\left(c\left(w'\right),d\left(w'\right)\right)\right)\geq h\left(\frac{x\left(w\right)}{w},w\right)-h\left(\frac{x\left(w'\right)}{w},w\right)\text{ for all }w,w'\in\mathcal{W}.$$

Equivalently, incentive compatibility holds if and only if

$$g\left(v\left(c\left(w\right),d\left(w\right)\right)\right)+\delta-g\left(v\left(c\left(w'\right),d\left(w'\right)\right)\right)-\delta\geq h\left(\frac{x\left(w\right)}{w},w\right)-h\left(\frac{x\left(w'\right)}{w},w\right)\text{ for all }w,w'\in\mathcal{W}.$$

Hence, if the given allocation is incentive compatible and has a non-binding resource constraint, the modified allocation is incentive feasible and Pareto superior. This step establishes the first claim of the proposition.

For the second claim, consider the multiplicative specification

$$u\left(v\left(c,d\right),l,w\right)=g\left(v\left(c,d\right)\right)h\left(l,w\right).$$

Similar to the previous argument, introduce a consumption increment that Pareto improves the given allocation without affecting the incentive problem. In this case, this can be achieved by an increment $\gamma\left(w\right)$ that proportionally increases the commodity part of the utility function:

$$g\left(v\left(c\left(w\right)+\gamma\left(w\right),d\left(w\right)\right)\right)=\left(1+\delta\right)g\left(v\left(c\left(w\right),d\left(w\right)\right)\right)$$

for $\delta>0$. It is easy to see that this modification Pareto improves the given allocation, and does so while preserving the incentive pattern of the allocation. Moreover, if the given allocation has a non-binding resource constraint, the modified allocation is feasible if δ is sufficiently small. $\square$

B Deduction Rate Depending on Total Income

This appendix provides the formula for the optimal deduction rate as a function of total income, as given in Equation (30).

Consider deduction rates that depend on total income $y(x,d)$. In this case, the individual commodity choice problem is given by

$$\max_{d} v\left(y\left(x,d\right)-T\left(y\left(x,d\right)-\alpha\left(y\left(x,d\right)\right)pd\right)-pd,d\right). \tag{36}$$

The first-order condition is

$$\left[y_d-\left(y_d-p\alpha-pdy_d\alpha'\right)T'-p\right]v_c+v_d=0. \tag{37}$$

Equivalently,

$$y_d+\frac{v_d}{v_c}=p+T'\left(y_d-p\alpha-pdy_d\alpha'\right). \tag{38}$$

By equating the individual first-order condition with Proposition 1, we obtain

$$p=y_d+\frac{v_d}{v_c}=p+\left(y_d-p\alpha-pdy_d\alpha'\right)T'. \tag{39}$$

Hence, if $T'\neq 0$, the deduction rate is characterized by the following equation:

$$p\alpha\left(y\left(x\left(w\right),d\left(w\right)\right)\right)+pd\left(w\right)y_d\left(x\left(w\right),d\left(w\right)\right)\alpha'\left(y\left(x\left(w\right),d\left(w\right)\right)\right)=y_d\left(x\left(w\right),d\left(w\right)\right). \tag{40}$$

Differentiating the product $pd\left(w\right)\alpha\left(y\left(x\left(w\right),d\left(w\right)\right)\right)$ with respect to $d\left(w\right)$ yields

$$p\alpha\left(y\left(x\left(w\right),d\left(w\right)\right)\right)+pd\left(w\right)y_d\left(x\left(w\right),d\left(w\right)\right)\alpha'\left(y\left(x\left(w\right),d\left(w\right)\right)\right), \tag{41}$$

which coincides with the left-hand side of (40). Hence, integration of (40) with respect to $d\left(w\right)$ yields

$$pd\left(w\right)\alpha\left(y\left(x\left(w\right),d\left(w\right)\right)\right)-C=\int y_d\left(x\left(d^{-1}\left(d\left(w\right)\right)\right),d\left(w\right)\right)\mathrm{d}d\left(w\right), \tag{42}$$

where C is an arbitrary constant. Alternatively, using a change of variables, we obtain

$$pd(w)\,\alpha(y(x(w),d(w))) - C = \int y_d(x(w),d(w))\,d'(w)\,\mathrm{d}w. \tag{43}$$

Now, Equation (30) follows after changing the integration variable and setting $C = 0$. $\square$

C Proof of Proposition 4

This appendix presents the proof of Proposition 4 for the optimal constant deduction rate α. We abandon our previous mechanism design approach and adapt the tax reform approach of Kaplow (2006, 2012, 2020).

Let v^* denote the indirect sub-utility over c and d one would obtain as a function of the tax system and a given choice of labor income wl:

$$v^*(wl;T,\alpha) = \max_{c,d} v(c,d) \text{ s.t. } c + pd = y(wl,d) - T(y(wl,d) - \alpha pd).$$

Consider an increase of the deduction rate α by a small amount $\Delta\alpha$. By the envelope theorem, the increased deduction rate changes individual sub-utilities by

$$v_c \cdot p \cdot d \cdot T' \cdot \Delta\alpha.$$

To offset the welfare effect of the perturbation, we adjust the individual income tax. Note that, by the envelope theorem, increasing the income tax from T to $T + \Delta T$ affects individual sub-utilities by $-v_c \cdot \Delta T$. Hence, the welfare effect of the deduction perturbation is neutralized by setting:

$$\Delta T = p \cdot d \cdot T' \cdot \Delta\alpha.$$

By construction, the joint reform is neutral with respect to individual sub-utilities, i.e., we have

$$v^*(wl;T+\Delta T,\alpha+\Delta\alpha) = v^*(wl;T,\alpha) \quad \text{for all } wl.$$

The next step shows that the reform is also neutral with respect to labor effort and, therefore, neutral with respect to individual welfare $U^*(w;T,\alpha)$. Prior to the reform, an individual with type w chooses labor effort l to maximize $u(v^*(wl;T,\alpha),l)$. Since $v^*(wl;T,\alpha) = v^*(wl;T+\Delta T,\alpha+\Delta\alpha)$ for every possible combination wl, the mapping of utilities associated with each labor effort option is the same before and after the reform. Therefore, the optimal labor effort choice does not change.

Because individual welfare is unaffected by the reform, a necessary condition for Pareto optimality of the original tax system (T, α) is that the tax revenue is maximized. Tax revenue after the reform is given by

$$\int_w \left[T\left(y\left(wl\left(w\right), d\left(w\right)\right) - \left(\alpha + \Delta\alpha\right) pd\left(w\right)\right) + \Delta T\left(w\right)\right] \mathrm{d}F\left(w\right).$$

Differentiating the tax revenue with respect to $\Delta\alpha$, evaluated at $\Delta\alpha = 0$, yields the following necessary condition:

$$\int_w \left[\left(y_d \frac{\mathrm{d}d}{\mathrm{d}\alpha} - pd - \alpha p \frac{\mathrm{d}d}{\mathrm{d}\alpha}\right) T' + pdT'\right] \mathrm{d}F\left(w\right) = 0,$$

where we have used the result that labor effort is not affected by the reform as shown above. Equivalently,

$$\int_w \left[\left(y_d - \alpha p\right) \frac{\mathrm{d}d}{\mathrm{d}\alpha} T'\right] \mathrm{d}F\left(w\right) = 0.$$

Equivalently,

$$\int_w \frac{y_d}{p} \frac{\mathrm{d}d}{\mathrm{d}\alpha} T' \,\mathrm{d}F\left(w\right) - \alpha \int_w \frac{\mathrm{d}d}{\mathrm{d}\alpha} T' \,\mathrm{d}F\left(w\right) = 0.$$

Equivalently,

$$\alpha = \frac{\int_w \frac{y_d}{p} \frac{\mathrm{d}d}{\mathrm{d}\alpha} T' \,\mathrm{d}F\left(w\right)}{\int_w \frac{\mathrm{d}d}{\mathrm{d}\alpha} T' \,\mathrm{d}F\left(w\right)} = \int_w \frac{y_d}{p} \frac{\frac{\mathrm{d}d}{\mathrm{d}\alpha} T'}{\int_w \frac{\mathrm{d}d}{\mathrm{d}\alpha} T' \,\mathrm{d}F\left(w\right)} \mathrm{d}F\left(w\right).$$

Relabelling the terms yields the result in Proposition 4. □

D Proof of Proposition 5

Consider an increase of the deduction rate $\alpha\left(wl\right)$ by a small amount $\Delta\alpha$ on a small interval of types $[w, w + \Delta w)$. By the envelope theorem, the increased deduction rate changes individual utilities by

$$U_c \, p \, d \, T' \Delta\alpha.$$

Since the mass of affected individuals is $f\left(w\right) \Delta w$, the effect on social welfare is given by

$$W := \pi\left(w\right) U_c \, p \, d \, T' f\left(w\right) \Delta\alpha\Delta w.$$

Mechanically, the reform reduces the level of taxable income of the affected types by $p \cdot d \cdot \Delta\alpha$. The welfare effect of this mechanical fiscal cost is given by

$$M := -\lambda \, p \, d \, T' \, f(w) \, \Delta\alpha\Delta w,$$

where λ represents the value of a dollar to the government. Finally, the affected types may change their consumption and labor supply. The welfare consequence of this behavioral fiscal cost is given by

$$B := \lambda \left[y_x w \frac{\mathrm{d}l}{\mathrm{d}\alpha}\Delta\alpha + y_d \frac{\mathrm{d}d}{\mathrm{d}\alpha}\Delta\alpha - \alpha p \frac{\mathrm{d}d}{\mathrm{d}\alpha}\Delta\alpha \right] T' f(w) \, \Delta w.$$

If the original tax system maximizes social welfare, the welfare effect of the perturbation must be zero. Setting $W + M + B = 0$ and eliminating the common factors yields

$$\left(\frac{\pi U_c}{\lambda} - 1 \right) pd + y_x w \frac{\mathrm{d}l}{\mathrm{d}\alpha} + y_d \frac{\mathrm{d}d}{\mathrm{d}\alpha} - \alpha p \frac{\mathrm{d}d}{\mathrm{d}\alpha} = 0$$

Solving for α, we obtain

$$\alpha = \frac{\left(\frac{\pi U_c}{\lambda} - 1 \right) pd + y_x w \frac{\mathrm{d}l}{\mathrm{d}\alpha} + y_d \frac{\mathrm{d}d}{\mathrm{d}\alpha}}{p \frac{\mathrm{d}d}{\mathrm{d}\alpha}},$$

which establishes Proposition 5. □