Optimal Income Tax Deductions for Mixed Business and Personal Expenditures*

Jacob Goldin Sebastian Koehne Nicholas Lawson

November 18, 2025

Abstract

We study the optimal taxation of expenditures that generate income while also serving a consumption function. We characterize the Pareto optimal income tax deduction for such mixed-purpose 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.

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, Spencer Bastani, Antoine Ferey, Louis Kaplow, Dominik Sachs, David Weisbach, seminar participants at Cornell and the University of Chicago, as well as conference audiences at the National Tax Association Annual Meeting, the AMT Tax Conference, and the CESifo Conference on Public Economics. We also thank the editor, Ben Lockwood, and three anonymous referees for valuable comments and suggestions.

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 aim to provide a principled basis for policy decisions about deductibility by studying 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 phone camera). Under these facts, our result suggests that phone plan expenditures should not be deductible, or at most deductible at a low rate. 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 labor 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, the optimal deduction rate we derive is a function of labor income, but the tax authority might not be able to separately observe pure labor income as distinct from income generated by the mixed-purpose good. We therefore derive alternative formulations for the optimal deduction rate that are functions of expenditures on the mixed-purpose good or of total income. Under mild regularity conditions, these alternative formulations are equivalent to the labor-income-based approach.

A second potential administrative constraint is that 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. For example, a uniform 50% deduction for business meals would be optimal if, on average, marginal expenditures on business meals contribute equally to consumption and income-generating purposes. This average is weighted more heavily for taxpayers in higher brackets and for those whose responses are more elastic.

A third 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.[1] 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 in Appendix A, 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.[2]

Distinct from this literature, Baake et al. (2004) also study an issue that is related to our focus: the in-

[1]Some contributions to this literature model the non-separability between consumption and leisure using a needs constraint tied to a specific consumption good, such as child care. In contrast to our focus, that specific good may be publicly provided and affect the individual budget constraint through post-tax rather than pre-tax income (e.g., Bastani et al., 2015).

[2]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. In Appendix A, we illustrate how our approach extends the optimal policy analysis relative to the preference-based approach.

ability 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.[3]

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).[4] 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 first-best 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 in 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 $F : \mathcal{W} \to [0, 1]$.

[3] 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.

[4] 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).

2.1 Preferences

Individuals have weakly separable preferences of the form

$$U(c, d, l, w) = u(v(c, d), l, w), \quad (1)$$

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

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.[5] The income production function $y : \mathbb{R}_+^2 \to \mathbb{R}_+$ is continuously differentiable with $y_x > 0$ and $y_d \geq 0$,[6] 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 identical across individuals.

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

2.3 Allocations

An *allocation* is a function $(c, d, x) : \mathcal{W} \to \mathbb{R}_+^3$ specifying regular consumption c, hybrid consumption d and labor income x for every type w.

[5] As shown in Appendix A, our model can equivalently be represented as a model with a specific form of non-separable preferences over consumption and *pre-tax income*.

[6] The latter restriction is for ease of interpretation; our analysis extends to the case in which $y_d < 0$, which we discuss below.

2.4 Tax Systems

Our primary focus is on allocations generated by an income tax system (T, α, γ) of the following form:[7]

$$T\left(y\left(wl, d\right) - \alpha\left(wl\right) pd + \gamma\left(wl\right)\right). \tag{2}$$

In this specification, the income tax schedule $T : \mathbb{R} \rightarrow \mathbb{R}$ maps taxable incomes to tax liabilities. Taxable income is defined as total income y minus the deduction for the hybrid good, which is given by the product of the deduction rate, α, and expenditures on the hybrid good, pd. We initially focus on the case in which the deduction rate $\alpha : \mathbb{R}_+ \rightarrow \mathbb{R}$ is a function of labor income; later, we will consider alternative specifications in which the deduction rate α is uniform across individuals or is a function of expenditures on the hybrid good. It will also be convenient to include an auxiliary tax component $\gamma\left(wl\right)$ in the definition of taxable income, $\gamma : \mathbb{R}_+ \rightarrow \mathbb{R}$. This additional term serves as a flexible device for offsetting the welfare effects of reforms to other components of the tax system. Importantly, our characterization of the optimal deduction rate α will not depend on the shape of γ. In particular, if a tax system defined by the schedule $T\left(y\left(wl, d\right) - \alpha\left(wl\right) pd\right)$ is already Pareto optimal, then no welfare-neutral, revenue-increasing reform exists and the auxiliary tax component γ can be set to zero. Throughout, we assume the functions T, α, and γ are continuously differentiable in their respective arguments.

2.5 Individual Problem

Faced with a tax system of the form in (2), the individual problem is given by

$$\max_{c,d,l} u\left(v\left(c, d\right), l, w\right) \text{ s.t. } c + pd = y\left(wl, d\right) - T\left(y\left(wl, d\right) - \alpha\left(wl\right) pd + \gamma\left(wl\right)\right). \tag{3}$$

The weak separability of the utility function allows us to decompose this problem into two stages.

Stage 2: Commodity Choice. For any labor income wl, the individual chooses a combination (c, d) that maximizes subutility subject to the budget constraint. Indirect subutility for this problem is given by:

$$v^*\left(wl; T, \alpha\left(wl\right), \gamma\left(wl\right)\right) = \max_{c,d} v\left(c, d\right) \text{ s.t. } c + pd = y\left(wl, d\right) - T\left(y\left(wl, d\right) - \alpha\left(wl\right) pd + \gamma\left(wl\right)\right). \tag{4}$$

[7] An alternative mechanism design approach to characterizing the optimal policy would be to identify the feasible allocations that satisfy incentive compatibility conditions arising from the non-observability of labor effort and skill. In this paper, our approach is to instead focus on allocations induced by an explicit tax system with a parsimonious functional form, drawing on the tax reform framework developed in Kaplow (2006, 2012, 2020). A complementary mechanism design analysis is sketched in Appendix B.

This problem generates the following first-order condition:

$$0 = \left[y_d - (y_d - \alpha p) T' - p\right] v_c + v_d. \tag{5}$$

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

Stage 1: Labor Supply. Anticipating the subutilities $v^* (wl; T, \alpha (wl), \gamma (wl))$, the individual chooses labor supply l to maximize utility

$$U^* (w; T, \alpha, \gamma) = \max_l u (v^* (wl; T, \alpha (wl), \gamma (wl)), l, w). \tag{6}$$

Note that the labor supply problem depends on the tax system and commodity choices only indirectly through the subutilities $v^* (wl; T, \alpha (wl), \gamma (wl))$.

3 Optimal Deduction Rate

In this section, we present our main result: the characterization of the Pareto optimal deduction rate. We find that the optimal deduction rate depends solely on the relative marginal income and consumption benefits provided by the hybrid good—independent of distributional concerns or labor supply considerations. The key intuition is that, due to the weak separability of preferences, the choice between the hybrid good and regular consumption is decoupled from the individual labor supply decision. As a result, the relative prices of the two consumption goods should not be distorted. However, the true price of the hybrid good reflects the income it generates—this additional income effectively constitutes a discount to its price. Thus, when the income produced by the hybrid good is taxed, the tax system must include some extra component to achieve an undistorted trade-off between the hybrid good and other consumption goods. The optimal hybrid good deduction plays this role by inducing taxpayers to consume the hybrid good according to a first-best choice rule notwithstanding the presence of the income tax.

The following proposition formalizes this intuition by deriving the optimal deduction rate.

Proposition 1 (Optimal Deduction Rate)**.** *Consider tax functions of the form $T (y (x, d) - \alpha (x) pd + \gamma (x))$, where $\alpha (x)$ denotes the deduction rate for expenditures on the hybrid good. Assume $T' \neq 0$ for all levels of* *taxable income. Then, the Pareto optimal deduction rate satisfies*

$$\alpha\left(x\left(w\right)\right)=\frac{y_{d}\left(x\left(w\right),d\left(w\right)\right)}{y_{d}\left(x\left(w\right),d\left(w\right)\right)+\frac{v_{d}\left(c\left(w\right),d\left(w\right)\right)}{v_{c}\left(c\left(w\right),d\left(w\right)\right)}}\tag{7}$$

for every skill type w.

Proof. We adapt the tax reform approach of Kaplow (2006, 2012, 2020) to the case of marginal deduction reforms.

Recall from Equation (4) that $v^{*}\left(wl;T,\alpha\left(wl\right),\gamma\left(wl\right)\right)$ denotes the indirect subutility over c and d that individuals obtain as a function of the tax system and a given choice of labor income $x=wl$. Consider a reform that perturbs the deduction rate for each individual by a small amount $\Delta\alpha(x)$, where the dependence on x reflects that the perturbation may vary with the individual's labor income level. As $\Delta\alpha\left(x\right)\to 0$, the envelope theorem implies that the effect of the reform on the subutility of an individual with labor income x is given by

$$v_{c}\cdot p\cdot d\cdot T'\cdot\Delta\alpha\left(x\right).\tag{8}$$

To compensate for this welfare effect, we adjust the component $\gamma\left(x\right)$ in the tax function. By the envelope theorem, increasing $\gamma(x)$ by a small amount $\Delta\gamma(x)$ reduces subutility for an individual with labor income x by $v_{c}\cdot T'\cdot\Delta\gamma(x)$. Thus, to neutralize the effect of the deduction reform, we set

$$\Delta\gamma\left(x\right)=p\cdot d\cdot\Delta\alpha\left(x\right).\tag{9}$$

Therefore, the *combined reform* $(\Delta\alpha,\Delta\gamma)$, for which $\Delta\gamma(x)$ is determined according to (9), leaves individual subutility unchanged for every value of labor income x:

$$v^{*}\left(x;T,\alpha\left(x\right),\gamma\left(x\right)\right)=v^{*}\left(x;T,\alpha\left(x\right)+\Delta\alpha\left(x\right),\gamma\left(x\right)+\Delta\gamma\left(x\right)\right)\tag{10}$$

The next step is to show that this combined reform is also neutral with respect to labor effort, and therefore neutral with respect to total individual welfare $U^{*}\left(w;T,\alpha,\gamma\right)$. Prior to the reform, an individual of type w chooses the labor effort l to maximize $u\left(v^{*}\left(wl;T,\alpha\left(wl\right),\gamma\left(wl\right)\right),l,w\right)$. Since (10) holds for every value of $x=wl$, we have, for any given w and l:

$$u\left(v^{*}\left(wl;T,\alpha\left(wl\right),\gamma\left(wl\right)\right),l,w\right)=u\left(v^{*}\left(x;T,\alpha\left(x\right)+\Delta\alpha\left(x\right),\gamma\left(x\right)+\Delta\gamma\left(x\right)\right),l,w\right)\tag{11}$$

where the left and right sides of the equality respectively denote the utilities associated with choosing labor effort l for an individual of type w. Because these levels are equal, the utility associated with each labor effort option is the same before and after the combined reform. Because this is true for every type w, each individual's labor effort choice remains unchanged. Consequently, each individual's overall welfare U^* is also unaffected by the combined reform.

Because the combined reform does not affect individual welfare, a necessary condition for the Pareto optimality of the original tax system (T, α, γ) is that no small reform that satisfies (9) can increase tax revenue. Otherwise, such a reform would increase tax revenue without affecting individual welfare. The additional revenue collected from the reform could then be distributed to individuals lump-sum, leading to a strict welfare improvement for each individual relative to the original tax system, and establishing that the original tax system was not Pareto optimal. For this reason, Pareto optimality requires that the function $\Delta\alpha = 0$ solve the following tax revenue maximization problem:

$$\max_{\Delta\alpha} \int_{x(\underline{w})}^{x(\overline{w})} T\left(y\left(x, \hat{d}\left(x; \Delta\alpha(x)\right)\right) - \left[\alpha\left(x\right) + \Delta\alpha\left(x\right)\right] p\hat{d}\left(x; \Delta\alpha(x)\right) + \gamma\left(x\right) + pd\left(x\right)\Delta\alpha\left(x\right)\right) \mathrm{d}G\left(x\right), \quad (12)$$

where we have used $\hat{d}\left(x; \Delta\alpha(x)\right)$ to denote the demand for the hybrid good under the combined reform, G to denote the distribution of labor income x (which is unaffected by the combined reform), and where we have used the compensation condition (9) to express $\Delta\gamma\left(x\right)$ as a function of $\Delta\alpha\left(x\right)$.

The objective in (12) is separable in x; that is, each contribution to the integral depends only on x and $\Delta\alpha(x)$, but not on its derivative, $\Delta\alpha'(x)$. Hence, by the Euler–Lagrange condition, optimality can be imposed point-wise. Evaluating this condition at $\Delta\alpha = 0$ gives

$$T'\left(\hat{z}\left(x\right)\right) \cdot g\left(x\right) \cdot \left(y_d \cdot \frac{\mathrm{d}\hat{d}\left(x;0\right)}{\mathrm{d}\Delta\alpha(x)} - p\hat{d}\left(x;0\right) - \alpha\left(x\right) p \frac{\mathrm{d}\hat{d}\left(x;0\right)}{\mathrm{d}\Delta\alpha(x)} + pd\left(x\right)\right) = 0 \quad (13)$$

for all x. Here, $\hat{z}\left(x\right) = y\left(x, \hat{d}\left(x;0\right)\right) - \alpha\left(x\right) p\hat{d}\left(x;0\right) + \gamma\left(x\right)$ denotes taxable income, $g = G'$ denotes the density function of G, and $\mathrm{d}\hat{d}\left(x; \Delta\alpha(x)\right)/\mathrm{d}\Delta\alpha(x)$ denotes the total derivative of demand for the hybrid good with respect to the combined reform $\left(\Delta\alpha\left(x\right), \Delta\gamma\left(x\right)\right) = \left(\Delta\alpha\left(x\right), pd\left(x\right)\Delta\alpha\left(x\right)\right)$.[8]

Conceptually, the left-hand side of (13) captures the change in revenue from a single labor income group

[8] More explicitly, if $\widetilde{d}\left(x; \Delta\alpha(x), \Delta\gamma(x)\right)$ specifies the demand for the hybrid good under any reform (not necessarily one satisfying (9)) as a separate function of $\Delta\alpha(x)$ and $\Delta\gamma(x)$, we have

$$\frac{\mathrm{d}\hat{d}\left(x; \Delta\alpha(x)\right)}{\mathrm{d}\Delta\alpha(x)} = \frac{\partial\widetilde{d}\left(x; \Delta\alpha(x), \Delta\gamma(x)\right)}{\partial\Delta\alpha(x)} + pd\left(x\right)\frac{\partial\widetilde{d}\left(x; \Delta\alpha(x), \Delta\gamma(x)\right)}{\partial\Delta\gamma(x)}.$$

The second term arises because $\Delta\gamma(x)$ is adjusted alongside $\Delta\alpha(x)$ under the combined reform satisfying (9).

in response to a reform that modifies the deduction rate pertaining to that group. If this revenue change is non-zero for some value of x, one could adopt a revenue-increasing reform by modifying $\alpha(x)$ in a small neighborhood of x. Hence, for a given $\alpha(x)$ schedule to be revenue-maximizing, it must be the case that (13) holds for all x.

Using the identity $\hat{d}\,(x;0) = d\,(x)$, the optimality condition simplifies to

$$T'\left(\hat{z}\,(x)\right) \cdot g\,(x) \cdot \left(y_d - \alpha\,(x)\,p\right) \cdot \frac{\mathrm{d}\hat{d}\,(x;0)}{\mathrm{d}\Delta\alpha(x)} = 0 \tag{14}$$

for all x. Because the reform creates a welfare-neutral price change of the hybrid good, we have $\frac{\mathrm{d}\hat{d}(x;0)}{\mathrm{d}\Delta\alpha(x)} \neq 0$. (More specifically, if $T' > 0$, the reform corresponds to a compensated price decrease of the hybrid good, leading to an increase in its demand; if $T' < 0$, it corresponds to a compensated price increase, thereby reducing its demand.) For any x for which $T'\left(z\,(x)\right) \neq 0$ and $g\,(x) \neq 0$, Equation (14) thus implies

$$y_d = \alpha\,(x)\,p. \tag{15}$$

Finally, we relate the price p to individual preferences using the individual optimization problem. Substituting Equation (15) into the first-order condition given in Equation (5), we obtain

$$y_d + \frac{v_d}{v_c} = p. \tag{16}$$

Hence, we can express Equation (15) as

$$\alpha\,(x) = \frac{y_d}{p} = \frac{y_d}{y_d + \frac{v_d}{v_c}}. \tag{17}$$

□

According to Proposition 1, the Pareto optimal deduction rate relates the marginal income-production benefit y_d to the marginal consumption benefit v_d/v_c of the good. The larger the income-production benefit relative to the consumption benefit, the larger is the optimal deduction rate. More specifically, Proposition 1 implies that under any Pareto optimal tax system, a marginal change in the hybrid good d does not affect taxable income:

$$\frac{\partial\left(y\,(x,d) - \alpha\,(x)\,pd\right)}{\partial d} = y_d - \alpha p = y_d - \frac{y_d}{y_d + \frac{v_d}{v_c}}p = 0, \tag{18}$$

where the last equality follows from the individual first-order condition under the optimal policy. Put differently, the optimal deduction rate excludes income from good d from the income tax base at the margin.

Another way to understand Proposition 1 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., $y(wl, d) = wl$ for all $(wl, d) \in \mathbb{R}^2_+$. 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 1 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., $v(c, d) = v(c)$ 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 1 is applied to this special case, in which $v_d = 0$, it yields the same result, i.e., $\alpha = 1$.[9]

An important practical consideration in implementing Proposition 1 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 individual objectives. For settings in which the shape of the income-generating function $y(\cdot)$ is known, this challenge can be alleviated by expressing the optimal deduction rate from Proposition 1 as

$$\alpha(x(w)) = \frac{y_d(x(w), d(w))}{p}, \tag{19}$$

where the equivalence follows from taxpayers' first-order conditions under the optimal policy, reflected in Equation (16). Intuitively, this formulation highlights that the optimal deduction rate depends on the share

[9] A related special case occurs when $v_d = 0$ and d is interpreted to be a productivity-enhancing good, such as education that does not carry any consumption benefits. As in Bovenberg and Jacobs (2005), Proposition 1 then implies that expenditures on such a good should be fully deductible. Conversely, our results suggest that productivity-enhancing expenditures should be less than fully deductible to the extent they also generate consumption value.

of expenditures on d that the taxpayer gets back in the form of income.

Finally, although our focus has been on settings in which the hybrid good positively contributes to income, Proposition 1 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 1 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. More generally, one can understand hybrid goods as generating a (positive or negative) fiscal externality with magnitude that depends on the taxpayer's marginal tax rate; hence the corresponding Pigouvian corrective tax must also vary based on the taxpayer's marginal tax rate. An income tax deduction characterized by Proposition 1 is a convenient instrument for accomplishing this objective.

4 Extensions

This section extends our results in a number of directions. First, we consider settings in which the deduction rate is constrained to be uniform. Second, to address potential challenges in isolating "pure" labor income wl from total income y, we explore deduction rates that are functions of expenditures or total income rather than labor income. For both of these extensions, we show that the optimal deduction rate is shaped by forces similar to those that govern Proposition 1. Finally, we consider the design of deduction rates when the income tax schedule cannot be adjusted; with that additional constraint, we show that the optimal deduction rule incorporates both distributional and labor supply considerations.

4.1 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. 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.

Under this restriction, the tax system is defined by

$$T\left(y\left(wl, d\right) - \alpha \cdot pd + \gamma(wl)\right). \tag{20}$$

As above, $T : \mathbb{R} \to \mathbb{R}$ and $\gamma : \mathbb{R}_+ \to \mathbb{R}$ denote the tax function and the auxiliary tax component, respectively, but we now assume $\alpha \in \mathbb{R}$.

Applying the tax reform approach to this setting yields 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.

Proposition 2 (Optimal Uniform Deduction Rate)**.** *Consider tax functions of the form $T\left(y\left(x,d\right) - \alpha \cdot pd + \gamma\left(x\right)\right)$ with a constant deduction rate α. Then, the Pareto optimal deduction rate satisfies*

$$\alpha = \int_{\underline{w}}^{\overline{w}} \theta_w \frac{y_d\left(x\left(w\right), d\left(w\right)\right)}{p} \mathrm{d}F(w) \tag{21}$$

where the weights θ_w are given by

$$\theta_w = \frac{T'\left(z(w)\right) \frac{\mathrm{d}\hat{d}(w)}{\mathrm{d}\alpha}}{\int_{\underline{w}}^{\overline{w}} T'\left(z(\omega)\right) \frac{\mathrm{d}\hat{d}(\omega)}{\mathrm{d}\alpha} \mathrm{d}F(\omega)}, \tag{22}$$

with $z\left(w\right) = y\left(x\left(w\right), d\left(w\right)\right) - \alpha \cdot pd\left(w\right) + \gamma\left(x\left(w\right)\right)$ representing taxable income.[10] *Here, the derivative $\mathrm{d}\hat{d}(w)/\mathrm{d}\alpha$ denotes the total derivative of demand for the hybrid good with respect to a combined reform to the deduction rate α and auxiliary tax component $\gamma(\cdot)$.*

Proof. See Appendix C.1. □

This result is an intuitive extension of our main result from Proposition 1. The formula for the optimal individual-specific deduction rate, which equals $\alpha = \frac{y_d}{p}$ under the conditions of Proposition 1, is now weighted by the θ_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.

4.2 Expenditure-Based Deduction Rate

In the main analysis, we considered tax functions where the deduction rate is a function of labor income. We now extend the framework to allow deduction rates that are functions of the level of expenditure on the hybrid good. Under plausible conditions, we demonstrate that this alternative formulation is equivalent

[10] The notation switches from w to ω in the integral defining θ_w to distinguish the running variable of integration from the evaluation point w of the weight θ_w.

to the labor-income-based approach. We also briefly sketch another equivalent formulation in which the deduction rate depends on total income.

Our focus in this subsection is on tax systems of the form

$$T\left(y\left(wl,d\right)-\alpha\left(pd\right)pd+\gamma\left(wl\right)\right), \tag{23}$$

where the deduction rate α is a function of total spending on the hybrid good, pd. The individual optimization problem is then given by

$$\max_{c,d,l} u\left(v\left(c,d\right),l,w\right) \text{ s.t. } c+pd=y\left(wl,d\right)-T\left(y\left(wl,d\right)-\alpha\left(pd\right)pd+\gamma\left(wl\right)\right). \tag{24}$$

As in the previous analysis, the weak separability of preferences allows us to decompose the problem into two stages: a commodity choice stage and a labor supply stage. This structure implies that any two tax systems yielding the same commodity choices at each level of labor income will also induce identical labor supply decisions.

Hence, to compare the tax systems (2) and (23), it suffices to examine the commodity choice problem conditional on labor income $x=wl$. Under the expenditure-based deduction rule (23) and associated budget constraint from (24), this stage of the individual's problem reduces to

$$\max_{d} v\left(y\left(x,d\right)-T\left(y\left(x,d\right)-\alpha\left(pd\right)pd+\gamma\left(x\right)\right)-pd,d\right). \tag{25}$$

To construct an expenditure-based deduction rule that replicates the Pareto optimal policy characterized in Proposition 1, we proceed in two steps. First, we derive a candidate deduction rule that yields the same first-order condition for commodity choice as the labor-income-based tax system from the main analysis. Second, we establish conditions under which the individual's commodity choice problem is globally concave, thereby ensuring that the first-order condition is sufficient to determine a unique solution to the individual choice problem.

4.2.1 Candidate Deduction Rule

To derive the candidate rule, observe that the first-order condition of problem (25) can be written as

$$y_d+\frac{v_d}{v_c}=p+\left(y_d-p\alpha-p^2d\alpha'\right)T'. \tag{26}$$

This condition equates the total marginal benefit of consuming the hybrid good—consisting of both its income-generating and consumption value—to its effective marginal cost. The latter includes the direct out-of-pocket price p as well as the marginal tax impact arising from income generation and the expenditure-based deduction.

In comparison, the first-order condition under the optimal labor-income-based deduction rule, as characterized in Proposition 1, is given by

$$p = y_d + \frac{v_d}{v_c}. \tag{27}$$

Assuming $T' \neq 0$, the two first-order conditions are equivalent if and only if the following condition holds for every w:

$$p\alpha\left(pd(w)\right) + p^2 d(w)\alpha'\left(pd(w)\right) = y_d\left(x(w), d(w)\right). \tag{28}$$

Integration of this equation yields the candidate deduction rule for the expenditure-based system.

Proposition 3. *Let* $\left(c(w), d(w), x(w)\right)_{w\in\mathcal{W}}$ *be the allocation induced by a tax system with the optimal labor-income-based deduction rule given in Proposition 1. Assume that* d *is differentiable and strictly monotonic. Moreover, assume* $T' \neq 0$ *for all levels of taxable income. Then, this allocation satisfies the first-order condition of the individual commodity choice problem under the expenditure-based deduction rule defined by*

$$\alpha\left(pd(w)\right) := \frac{\int y_d\left(x(w), d(w)\right) d'(w)\,\mathrm{d}w}{pd(w)}. \tag{29}$$

Proof. See Appendix C.2. □

In Proposition 3, the deduction rate relates the *cumulative* marginal income benefit of the hybrid good to the expenditure on that good. The integral in Equation (29) is indefinite, representing the accumulation of the marginal income benefit y_d, weighted by d', up to type w; without loss of generality, we normalize the associated integration constant to zero.[11] Using a change of variables, the deduction rate in Equation (29) can be equivalently stated as:

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

[11] More precisely, the indefinite integral in Equation (29) be written as

$$\int y_d\left(x(w), d(w)\right) d'(w)\,\mathrm{d}w = \int_{\underline{w}}^{w} y_d\left(x\left(\omega\right), d\left(\omega\right)\right) d'\left(\omega\right)\,\mathrm{d}\omega + C,$$

where $C \in \mathbb{R}$ is an arbitrary constant of integration. This constant can be normalized to zero without loss of generality because a constant level shift of the deduction leaves marginal incentives regarding d unchanged, and the effect of such a shift on overall tax liability can be neutralized by an offsetting adjustment to the auxiliary tax component in Equation (36).

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, $\alpha(pd)\,pd$, by y_d units so that taxable income remains unchanged.

To relate this result to Proposition 1, we can use the individual first-order condition (27) to express the deduction rate in Proposition 3 as follows:

$$\alpha\left(pd\left(w\right)\right)=\frac{\frac{1}{d(w)}\int y_d\left(x(w),d(w)\right)d'(w)\mathrm{d}w}{y_d\left(x\left(w\right),d\left(w\right)\right)+\frac{v_d(c(w),d(w))}{v_c(c(w),d(w))}}. \tag{31}$$

Hence, Proposition 3 differs from Proposition 1 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 $y(x,d)=m(x)+n(d)$ for some functions m and n, 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

$$\int y_d\left(x(w),d(w)\right)d'(w)\mathrm{d}w=\int n'\left(d(w)\right)d'(w)\mathrm{d}w=n\left(d(w)\right) \tag{32}$$

where, as above, we have normalized the integration constant to zero. We obtain

$$y\left(x(w),d(w)\right)-\alpha\left(pd(w)\right)pd(w)=m\left(x(w)\right), \tag{33}$$

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

Budget Alignment. In addition to replicating the marginal incentives for the choice of the hybrid good, the expenditure-based deduction rule must also replicate the same consumption levels as the labor-income-based formulation. For this purpose, we use the auxiliary tax component γ to fine-tune the individual's budget.

For any labor income $x\left(w\right)$, the individual's tax liability under the expenditure-based deduction rule in Proposition 3 is given by

$$T\left(y\left(x\left(w\right),d\left(w\right)\right)-\int_{\underline{w}}^{w}y_d\left(x\left(\omega\right),d\left(\omega\right)\right)d'\left(\omega\right)\,\mathrm{d}\omega+\gamma\left(x\left(w\right)\right)\right). \tag{34}$$

By contrast, the tax liability under the labor-income-based rule given in Proposition 1 takes the form

$$T\Big(y\left(x\left(w\right),d\left(w\right)\right)-y_d\left(x\left(w\right),d\left(w\right)\right)d\left(w\right)+\gamma_0\left(x\left(w\right)\right)\Big), \tag{35}$$

where γ_0 denotes the auxiliary tax component in this baseline system. To ensure that the same consumption bundle $(c(w), d(w))$ remains budget feasible under both systems, we set the auxiliary tax component γ in the expenditure-based system to the following level:

$$\gamma\left(x\left(w\right)\right):=\gamma_0\left(x\left(w\right)\right)+\int_{\underline{w}}^{w} y_d\left(x\left(\omega\right),d\left(\omega\right)\right)d'\left(\omega\right)\,\mathrm{d}\omega-y_d\left(x\left(w\right),d\left(w\right)\right)d\left(w\right), \tag{36}$$

where we assume strict monotonicity of x to ensure that the argument w on the right-hand side of (36) is well-defined in terms of $x(w)$. Under this construction, the marginal incentives for commodity choice are aligned by the design of the deduction rule, as demonstrated in Proposition 3, while budget feasibility is restored through the adjusted auxiliary tax component.[12]

Income-Based Deduction Rate. 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 total incomes rather than expenditures.[13] 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, the expenditure-based rule (30) is equivalent with the income-based rule defined by

$$\alpha\left(y\left(x(w),d(w)\right)\right):=\frac{1}{pd\left(w\right)}\int_{d(\underline{w})}^{d\left(y^{-1}\left(y\left(x(w),d(w)\right)\right)\right)} y_d\left(x\circ d^{-1}(d),d\right)\mathrm{d}d. \tag{37}$$

This construction mirrors Proposition 3 and ensures once more that the hybrid good is removed from the tax base at the margin.

4.2.2 Concavity of Commodity Choice Problem

We now examine the concavity of the commodity choice problem under the deduction rules characterized in Proposition 1 and Proposition 3. The following result formalizes a set of sufficient conditions.

[12] Provided that income after deductions, defined as $z(w) = y(x(w), d(w)) - \int_{\underline{w}}^{w} y_d(x(\omega), d(\omega))d'(\omega)\,\mathrm{d}\omega$, is strictly monotonic in w, the auxiliary term can be absorbed into the outer tax function by redefining the tax schedule as $\tilde{T}(z(w)) := \tilde{T}(w) := T(z(w) + \gamma(x(w)))$. This reparametrization renders the auxiliary term irrelevant, so it can be set to zero without loss of generality. In practice this means that the reformed tax system need not directly depend on the labor component of income (which may be difficult to isolate).

[13] This inverse exists if both d and x are strictly monotonic in w.

Proposition 4. *The commodity choice problems associated with both the labor-income-based deduction rule* (7) *and the expenditure-based rule* (29) *are globally strictly concave under the following sufficient conditions:*

(i) the marginal tax rate satisfies $0 < T' < 1$ *for all levels of taxable income,*

(ii) the tax function is weakly progressive: $T'' \geq 0$*,*

(iii) the income-generating function $y(x, d)$ *is strictly concave in* d*,*

(iv) the cross-partial $y_{x,d}$ *is sufficiently small (possibly negative).*

Condition (iv) is satisfied, for instance, if $y_{x,d} = 0$*, or alternatively if* $y_{x,d} \leq 0$ *and both* $x(w)$ *and* $d(w)$ *are increasing in type* w*.*

Proof. See Appendix C.3. □

The conditions stated in Proposition 4 are relatively mild. Condition (i) is both realistic and standard in optimal tax theory. Condition (ii), which postulates weak progressivity of the tax function, is also empirically supported across a wide range of tax systems. Condition (iii) assumes decreasing returns and is standard in many production or human capital settings. Condition (iv) is somewhat more restrictive, as it imposes structure on the interaction between the labor input and the hybrid good. However, its main role is to ensure that each individual term in the objective function contributes to concavity. In practice, strong curvature arising from the income process and the tax schedule often dominates, making global concavity likely even if the cross-partial $y_{x,d}$ does not strictly satisfy the stated condition. Thus, condition (iv) should be seen as a technical device to facilitate the concavity result rather than a central modeling restriction.

Under the conditions stated in Proposition 4, the individual commodity choice problem is globally concave and admits a unique solution. This solution can be flexibly implemented using either the labor-income-based deduction rule (7), the expenditure-based rule (29), or the income-based deduction rule in (37). Since these rules yield identical commodity choices, they also induce the same labor supply behavior and, consequently, the same overall allocation. In summary, the functional parameter on which the nonlinear deduction rate is based can be chosen freely, provided that certain mild regularity conditions are satisfied. This flexibility allows policymakers to tailor the design of deduction rules to administrative considerations without compromising distributional goals or economic efficiency.

4.3 Fixed Income Tax Schedule

In our main analysis, we considered welfare-neutral reforms of the deduction rate, offsetting changes to α with adjustments to the auxiliary tax component γ. However, in some real-world settings, policymakers may not be able to modify the income tax schedule due to institutional or political constraints (e.g., Fennell and McAdams, 2015). In this subsection, we consider the optimal uniform deduction rate $\alpha \in \mathbb{R}$ for some fixed and potentially suboptimal income tax schedule $T\left(y(x,d) - \alpha \cdot pd + \gamma(x)\right)$, where $\gamma(\cdot)$ and $T(\cdot)$ are exogenous.

Throughout this analysis, we define social welfare as

$$W = \int_{\underline{w}}^{\overline{w}} \pi\left(w\right) U\left(c\left(w\right), d\left(w\right), l\left(w\right), w\right) \,\mathrm{d}F\left(w\right), \tag{38}$$

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.

The government's budget constraint is given by

$$b = \int_{\underline{w}}^{\overline{w}} T\left(y\left(wl\left(w\right), d\left(w\right)\right) - \alpha pd\left(w\right) + \gamma\left(wl\right)\right) \mathrm{d}F\left(w\right) - G, \tag{39}$$

where G denotes exogenous government spending and b is a residual lump-sum demogrant that balances the budget as α is reformed.

The welfare-maximizing deduction rate α is described in the following proposition. Following Diamond (1975), our characterization of the optimal α employs a weighting parameter β that captures both the direct welfare benefit to the individual as well as any changes in tax revenue from changes to the individual's behavior:

$$\beta := \frac{\pi U_c}{\lambda} + h, \tag{40}$$

where h captures the change in an individual's taxes from a marginal increase in disposable income and $\lambda = \frac{\mathbb{E}[\pi U_c]}{1 - \mathbb{E}[h]}$ captures the social value of additional government revenue.

Proposition 5. *If the tax function is of the form $T\left(y(x,d) - \alpha pd + \gamma(x)\right)$, then the welfare-maximizing deduction rate for exogenously fixed (and potentially suboptimal) functions γ and T is characterized by*

$$\alpha = \frac{\mathbb{E}\left[y_d \frac{\partial d^c}{\partial \alpha} T'\right] + \mathbb{E}\left[\left(y_x + \gamma'\right) w \frac{\partial l^c}{\partial \alpha} T'\right] + \mathrm{cov}\left(\beta, pdT'\right)}{p\mathbb{E}\left[\frac{\partial d^c}{\partial \alpha} T'\right]}, \tag{41}$$

where $\frac{\partial d^c}{\partial \alpha}$ and $\frac{\partial l^c}{\partial \alpha}$ denote the respective compensated (Hicksian) derivatives of d and l with respect to α, and where $\mathbb{E}$ and cov denote the expectation and covariance operators.

Proof. See Appendix C.4. □

Proposition 5 provides an intuitive extension of Proposition 2 to settings where labor supply and welfare effects of deduction reforms are not neutralized by adjustments to the income tax schedule. When only the first term in the numerator of (41) is non-zero, the optimal deduction rate in Proposition 5 coincides with the prior result from Proposition 2 with the behavioral response component of the weights modified to reflect the difference in the nature of the reforms.[14] Without an income tax adjustment, however, the second and third terms may be non-zero and must be accounted for when determining the optimal α. In particular, the second term reflects the effect of the reform on tax revenue collected from labor income, whereas the third term captures the distributional properties of the reform.

In the special case that the tax schedule (T, γ) is fixed at its optimum, the marginal social welfare benefits from additional redistribution offset the marginal efficiency costs from additional distortions to labor supply. In this case, the condition in Equation (41) reverts to the condition described in Equation (21), as described in the following extension to Proposition 5.

Proposition 6. *Suppose the tax function is of the form $T\left(y(x,d) - \alpha p d + \gamma(x)\right)$, and the functions γ and T are fixed at their respective optimums. Then, a marginal budget-neutral increase in the uniform deduction rate α increases social welfare if and only if*

$$\alpha < \int_{\underline{w}}^{\overline{w}} \theta_w \frac{y_d\left(x\left(w\right), d\left(w\right)\right)}{p} \mathrm{d}F(w) \tag{42}$$

where the weights θ_w are given by

$$\theta_w = \frac{T'\left(z(w)\right) \frac{\mathrm{d}\hat{d}(w)}{\mathrm{d}\alpha}}{\int_{\underline{w}}^{\overline{w}} T'\left(z(\omega)\right) \frac{\mathrm{d}\hat{d}(\omega)}{\mathrm{d}\alpha} \mathrm{d}F(\omega)}. \tag{43}$$

Here, the derivative $\mathrm{d}\hat{d}(w)/\mathrm{d}\alpha$ denotes the total derivative of demand for the hybrid good with respect to a combined reform to the deduction rate α and auxiliary tax component $\gamma(\cdot)$.

Proof. See Appendix C.5. □

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–Mirrlees intuitions in proportion to the degree to which marginal expenditures on a good are motivated by consumption or income-production

[14] The relevant behavioral response in Proposition 2 is to a reform of α and $\gamma(\cdot)$, whereas the relevant behavioral response for Proposition 5 is to a reform of α and the lump-sum demogrant b.

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 with 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 (see Givati (2020) for a discussion of this point).[15] 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. Fourth, by assuming that taxpayers have identical subutility functions, we have ruled out the possibility of tagging via heterogeneity in commodity preferences. If instead high and low skill types differed in their preferences for the mixed-purpose good, policymakers could potentially reduce the efficiency cost of redistribution by accounting for such heterogeneity in setting the deduction rate (e.g., Saez, 2002; Ferey et al., 2024). Fifth, for some mixed-purpose goods, the consumption decision is effectively binary, such as a potential trip to meet with business clients and also visit friends. For binary goods, the marginal benefits of purchasing the good coincide with the total benefits of doing so; hence, what matters for the optimal deduction rate is the relative importance of the *total* consumption and income-producing benefits, evaluated for the taxpayer type that is marginal with respect to the purchase decision.

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

[15] One can apply our framework to a hybrid good of this form by modeling the non-taxable income as additional units of regular consumption c, since this good does not affect tax liability under any of the tax systems we consider.

directly observe the strength of these various motivations. When the shape of the income function $y(\cdot)$ is known, one potential solution is to set the deduction rate in Proposition 1 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 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 2. 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.

Bastani, S., S. Blomquist, and J. Pirttilä (2015): "How should commodities be taxed? A counter-argument to the recommendation in the Mirrlees Review," *Oxford Economic Papers*, 67, 455–478.

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

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

Fennell, L. A. and R. H. McAdams (2015): "The distributive deficit in law and economics," *Minn. L. Rev.*, 100, 1051.

Ferey, A., B. B. Lockwood, and D. Taubinsky (2024): "Sufficient statistics for nonlinear tax systems with general across-income heterogeneity," *American Economic Review*, 114, 3206–3249.

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.

Gauthier, S. and G. Laroque (2009): "Separability and public finance," *Journal of Public Economics*, 93, 1168–1174.

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. (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.

Online Appendix (Not For Publication)

A 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

$$\widetilde{w} := \phi(w, d, l) := \frac{y(wl, d)}{l}, \tag{44}$$

which yields an income process that appears multiplicative in labor effort: $y(wl, d) = \widetilde{w}l$. In this transformation, the virtual skill level $\widetilde{w}$ 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 (44), 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 $\widetilde{w} = w \cdot \phi(d)$ 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

$$\widetilde{U}(c, d, y, w) = u\left(v(c, d), \frac{g(y, d)}{w}, w\right). \tag{45}$$

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 $v(c, d) + h(y, d, w)$.[16] In addition

[16] Gauthier and Laroque (2009) examine a related framework of individual production sets, where cost functions take the

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 income-generating function, 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.

To illustrate how our approach facilitates the optimal policy analysis relative to the preference-based approach, we apply our Pareto condition (54) to the utility specification in (45). Differentiating the identity $g\left(y\left(x,d\right),d\right)=x$, we obtain $y_d=-g_d/g_y$. Setting $p=1$ as in the model of Koehne and Sachs (2022), our condition becomes

$$1=\frac{v_d}{v_c}-\frac{g_d}{g_y}. \tag{46}$$

By contrast, the Pareto optimality condition derived by Koehne and Sachs (2022) involves cross-partials of the utility function and takes the form

$$\frac{\widetilde{U}_c+\widetilde{U}_y}{\widetilde{U}_{wy}}=\frac{\widetilde{U}_c-\widetilde{U}_d}{-\widetilde{U}_{wd}}. \tag{47}$$

In what follows, we demonstrate that our condition (46) implies the Koehne and Sachs (2022) condition (47).

Under the additive specification of Koehne and Sachs (2022), we have $u_v=1$ and $u_{vl}=u_{vw}=0$. Therefore, the first derivatives of the utility function (45) are given by

$$\begin{aligned}
\widetilde{U}_c&=v_c,\\
\widetilde{U}_d&=v_d+\frac{1}{w}u_l g_d,\\
\widetilde{U}_y&=\frac{1}{w}u_l g_y,
\end{aligned}$$

and the cross-partials with respect to skill w are

$$\begin{aligned}
\widetilde{U}_{wy}&=g_y\cdot\left(-\frac{1}{w^2}u_l-\frac{g}{w^3}u_{ll}+\frac{1}{w}u_{lw}\right),\\
\widetilde{U}_{wd}&=g_d\cdot\left(-\frac{1}{w^2}u_l-\frac{g}{w^3}u_{ll}+\frac{1}{w}u_{lw}\right).
\end{aligned}$$

separable form $h\left(H\left(y,d\right),w\right)$, and derive a generalized Atkinson–Stiglitz result within that setting.

Hence, the left-hand side of (47) can be written as

$$\frac{\widetilde{U}_c+\widetilde{U}_y}{\widetilde{U}_{wy}}=\frac{v_c+\frac{1}{w}u_l g_y}{g_y\cdot\left(-\frac{1}{w^2}u_l-\frac{g}{w^3}u_{ll}+\frac{1}{w}u_{lw}\right)}=-\frac{w^2v_c+wu_l g_y}{g_y\cdot\left(u_l+\frac{g}{w}u_{ll}-wu_{lw}\right)}.$$

The right-hand side of (47) can be written as

$$\frac{\widetilde{U}_c-\widetilde{U}_d}{-\widetilde{U}_{wd}}=-\frac{v_c-v_d-\frac{1}{w}u_l g_d}{g_d\cdot\left(-\frac{1}{w^2}u_l-\frac{g}{w^3}u_{ll}+\frac{1}{w}u_{lw}\right)}=\frac{w^2v_c-w^2v_d-wu_l g_d}{g_d\cdot\left(u_l+\frac{g}{w}u_{ll}-wu_{lw}\right)}.$$

After cross multiplication, we note that both sides of Equation (47) are indeed equal under the Pareto condition (46).

In summary, our approach recovers the preference-based optimality condition (47) as a special case. In addition, it provides a more transparent economic interpretation of Pareto optimality in environments with work-related goods, avoiding cross-partial utility terms involving unobservable skill.

B Mechanism Design Approach

This appendix offers an alternative derivation of Equation (7) using a mechanism design framework. The analysis proceeds in two steps. First, we characterize an optimal allocation rule under informational constraints. Second, we examine how this rule restricts the design of possible decentralizations.

B.1 Allocation constraints

An allocation $(c(w),d(w),x(w))_{w\in\mathcal{W}}$ is *feasible* if it satisfies the following resource constraint:

$$\int_{\underline{w}}^{\overline{w}}\left[c\left(w\right)+p\cdot d\left(w\right)-y\left(x\left(w\right),d\left(w\right)\right)\right]f\left(w\right)\mathrm{d}w\leq 0. \tag{48}$$

Assuming that the variables c, d, y and $x=wl$ are observable,[17] 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

$$u\left(v\left(c\left(w\right),d\left(w\right)\right),\frac{x\left(w\right)}{w},w\right)\geq u\left(v\left(c\left(w'\right),d\left(w'\right)\right),\frac{x\left(w'\right)}{w},w\right)\quad\text{for all }w,w'\in\mathcal{W}. \tag{49}$$

[17]If d is observable, then knowledge of the income production function $y(x,d)$ allows one to infer labor income $x=wl$ 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 weak separability of the utility function, the incentive compatibility condition (49) depends on commodity bundles $(c(w), d(w))$ only via the subutilities $v(c(w), d(w))$. Thus, any reform that leaves the subutilities unchanged will not affect the incentives to report truthfully.

B.2 Pareto optimal allocations

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

To support our resource minimization approach to deriving the optimal allocation rule, we assume throughout the analysis that the resource constraint (48) 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, such as through a lump-sum transfer. This assumption holds in many benchmark cases, including additively and multiplicatively separable utility functions.

B.3 First-Best Rule

Having established that the resource constraint binds, we now demonstrate that constrained Pareto optima satisfy a first-best rule for commodity allocation.

We consider two-dimensional allocation perturbations that change regular consumption c and the hybrid good d, while holding fixed the subutility 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:

$$\hat{d}(w) = d(w) + \varepsilon \cdot \delta(w), \tag{50}$$

$$\hat{c}(w) = c(w) - \psi(w, \varepsilon \cdot \delta(w)), \tag{51}$$

where $\psi(w, \varepsilon \cdot \delta(w))$ is defined such that the subutility remains unchanged,[18] i.e.,

$$v(c(w) - \psi(w, \varepsilon \cdot \delta(w)), d(w) + \varepsilon \cdot \delta(w)) = v(c(w), d(w)) =: \bar{v}(w). \tag{52}$$

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

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

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

Equation (52) ensures

$$v\left(\hat{c}\left(w\right),\hat{d}\left(w\right)\right)=v\left(c\left(w\right),d\left(w\right)\right)\text{ for all }w\in\mathcal{W}. \tag{53}$$

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 subutilities. 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 7 (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*

$$p=\frac{v_d\left(c\left(w\right),d\left(w\right)\right)}{v_c\left(c\left(w\right),d\left(w\right)\right)}+y_d\left(x\left(w\right),d\left(w\right)\right) \tag{54}$$

for almost all skill types w*.*

Proof. First, we show by contradiction that constrained Pareto optimality implies resource minimization for a given pattern of subutilities $v(c,d)$ 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 subutilities and labor incomes. Then, there exists an alternative allocation with identical subutilities and labor incomes that requires fewer resources. Since the subutilities and labor supplies are the same as in the original allocation, 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 subutilities, then $\varepsilon=0$ is a solution of the following resource problem:

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

The FOC evaluated at $\varepsilon = 0$ is

$$\int_{w_0}^{w_1} \left[p\delta(w) - \delta(w)\psi_2(w,0) - \delta(w)\, y_d(x(w), d(w))\right] f(w)\, \mathrm{d}w = 0,$$

where ψ_2 denotes the partial derivative of ψ with respect to its second argument. Equivalently,

$$\int_{w_0}^{w_1} \left[p - \psi_2(w,0) - y_d(x(w), d(w))\right] \delta(w) f(w)\, \mathrm{d}w = 0.$$

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

$$p - \psi_2(w,0) - y_d(x(w), d(w)) = 0 \text{ for almost all } w \in [w_0, w_1].$$

Differentiating Equation (52) with respect to $\varepsilon \cdot \delta(w)$, we obtain

$$-\psi_2(w, \varepsilon \cdot \delta(w))\, v_c\left(\hat{c}(w), \hat{d}(w)\right) + v_d\left(\hat{c}(w), \hat{d}(w)\right) = 0.$$

Hence,

$$\psi_2(w,0) = \frac{v_d(c(w), d(w))}{v_c(c(w), d(w))}.$$

Hence, we can express the optimality condition as

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

Because w_0 and w_1 were arbitrary, the condition extends to the entire type space $\mathcal{W}$. $\square$

The first-best rule derived above 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.

B.4 Consequence for the Optimal Deduction Rate

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 now demonstrate how this requirement constrains the design of a tax system with a simple deduction schedule.

Suppose that a constrained Pareto optimal allocation $(c(w), d(w), x(w))_{w \in \mathcal{W}}$ is implemented by a tax function of the form

$$T\left(y - \alpha\left(x\right) pd + \psi\left(x\right)\right),$$

where $x = wl$ represents labor income. By rearranging Equation (5), the first-order condition for the individual commodity choice problem equals

$$y_d\left(x(w), d(w)\right) + \frac{v_d\left(c(w), d(w)\right)}{v_c\left(c(w), d(w)\right)} = p + \left(y_d\left(x(w), d(w)\right) - \alpha\left(x(w)\right) p\right) T'. \tag{55}$$

Equating this condition with the first-best rule, we obtain

$$p = y_d\left(x(w), d(w)\right) + \frac{v_d\left(c(w), d(w)\right)}{v_c\left(c(w), d(w)\right)} = p + \left(y_d\left(x(w), d(w)\right) - \alpha\left(x(w)\right) p\right) T'. \tag{56}$$

Hence, if $T' \neq 0$, we obtain $y_d - \alpha p = 0$. Solving this condition for α, we obtain Equation (7).

C Supplementary Proofs

C.1 Proof of Proposition 2

This appendix presents the proof of Proposition 2 for the optimal constant deduction rate α. Similar to the proof of Proposition 1, we adapt the tax reform approach of Kaplow (2006, 2012, 2020) to the case of marginal deduction reforms.

Let v^* denote the indirect subutility over c and d that individuals obtain as a function of the tax system and a given choice of labor income $x = wl$:

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

Consider a small increase in the deduction rate α by an amount $\Delta\alpha$. By the envelope theorem, the increased deduction rate changes individual subutilities by

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

To compensate for this welfare effect, we adjust the component γ in the tax function. By the envelope theorem, increasing $\gamma(wl)$ by $\Delta\gamma(wl)$ reduces individual subutility by $v_c \cdot T' \cdot \Delta\gamma(wl)$. Thus, to neutralize the impact of the deduction perturbation on subutility, we set

$$\Delta\gamma(wl) = p \cdot d \cdot \Delta\alpha.$$

for each level of labor income wl. By construction, this combined reform leaves all individual subutilities unchanged.

The next step is to show that the reform is also neutral with respect to labor effort, and therefore neutral with respect to individual welfare $U^*(w; T, \alpha, \gamma)$. Prior to the reform, an individual of type w chooses the labor effort l to maximize $u(v^*(wl; T, \alpha, \gamma(wl)), l)$. Since

$$v^*(wl; T, \alpha, \gamma(wl)) = v^*(wl; T, \alpha + \Delta\alpha, \gamma(wl) + \Delta\gamma(wl))$$

for every possible value of wl, the utility associated with each labor effort option is the same before and after the reform. Therefore, the individual's labor effort choice remains unchanged. Consequently, each individual's overall utility U^* is also unaffected.

Since individual welfare is unchanged by the reform, a necessary condition for the Pareto optimality of the original tax system (T, α, γ) is that it maximizes tax revenue; otherwise welfare could be improved by implementing a welfare-neutral but revenue-increasing reform and distributing the additional revenue to individuals lump-sum. Tax revenue after the reform is given by

$$\int_{\underline{w}}^{\overline{w}} T\left(y\left(x(w), \hat{d}(x(w); \Delta\alpha)\right) - (\alpha + \Delta\alpha)\, p\hat{d}(x(w); \Delta\alpha) + \gamma(x(w)) + \Delta\gamma(x(w))\right) \mathrm{d}F(w),$$

where $\hat{d}(x(w); \Delta\alpha)$ denotes the demand for the hybrid good under the reform $(\Delta\alpha, \Delta\gamma) = (\Delta\alpha, pd\Delta\alpha)$. Differentiating this expression with respect to $\Delta\alpha$, and evaluating at $\Delta\alpha = 0$, yields the following necessary condition:

$$\int_{\underline{w}}^{\overline{w}} \left(y_d \frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha} - pd - \alpha p \frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha} + pd\right) T' \, \mathrm{d}F(w) = 0,$$

where $\mathrm{d}\hat{d}/\mathrm{d}\alpha$ denotes the the total derivative of demand for the hybrid good with respect to the combined reform $(\Delta\alpha, \Delta\gamma) = (\Delta\alpha, pd\Delta\alpha)$ and we have used the result that labor effort remains unchanged. Equivalently,

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

Equivalently,

$$\int_{\underline{w}}^{\overline{w}} \frac{y_d}{p} \frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha} T' \,\mathrm{d}F(w) - \alpha \int_{\underline{w}}^{\overline{w}} \frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha} T' \,\mathrm{d}F(w) = 0.$$

Equivalently,

$$\alpha = \frac{\int_{\underline{w}}^{\overline{w}} \frac{y_d}{p} \frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha} T' \,\mathrm{d}F(w)}{\int_{\underline{w}}^{\overline{w}} \frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha} T' \,\mathrm{d}F(w)} = \int_{\underline{w}}^{\overline{w}} \frac{y_d}{p} \frac{\frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha} T'}{\int_{\underline{w}}^{\overline{w}} \frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha} T' \,\mathrm{d}F(w)} \mathrm{d}F(w).$$

Relabelling the terms yields the result in Proposition 2. □

C.2 Proof of Proposition 3

This appendix presents the proof of Proposition 3 for the expenditure-based deduction rate $\alpha(pd)$.

As shown in section 4.2, the allocation $(c(w), d(w), x(w))_{w \in \mathcal{W}}$ induced by the optimal labor-income-based deduction rule satisfies the first-order condition of the commodity choice problem under the expenditure-based deduction rule $\alpha(pd)$ if and only if Equation (28) holds for every $w \in \mathcal{W}$. Setting $e(w) := pd(w)$ and noting $e'(w) = pd'(w)$, Equation (28) can be rewritten as

$$\frac{\mathrm{d}}{\mathrm{d}w} \left[e(w)\alpha(e(w)) \right] = y_d(x(w), d(w))d'(w). \tag{57}$$

Integrating both sides over w and solving for α yields

$$\alpha(pd(w)) = \frac{\int y_d(x(w), d(w))d'(w) \,\mathrm{d}w}{pd(w)}, \tag{58}$$

which establishes Equation (29). □

C.3 Proof of Proposition 4

This appendix presents the proof of Proposition 4 on the concavity of the commodity choice problems under the labor-income-based deduction rule (7) and the expenditure-based rule (29).

Given the concavity and monotonicity of the subutility function, a sufficient condition for the strict concavity of these problems is that disposable income $y - T$ varies strictly concavely with respect to the hybrid good d, holding labor income x fixed.

We begin with the expenditure-based rule (29). For notational convenience, we define the deduction level as

$$\mathcal{D}(pd) := \alpha(pd)\,pd.$$

The individual's disposable income can then be written as:

$$k(d) := y(x,d) - T\left(y(x,d) - \mathcal{D}(pd) + \gamma(x)\right).$$

The derivative of $k(d)$ with respect to d is given by:

$$k'(d) = y_d(x,d) - \left(y_d(x,d) - \mathcal{D}'(pd)\,p\right) T'\left(y(x,d) - \mathcal{D}(pd) + \gamma(x)\right).$$

Differentiating again, we obtain

$$k''(d) = y_{d,d}(x,d) - \left(y_{d,d}(x,d) - \mathcal{D}''(pd)\,p^2\right) T' - \left(y_d(x,d) - \mathcal{D}'(pd)\,p\right)^2 T'',$$

which can be rearranged as

$$k''(d) = (1-T')\,y_{d,d}(x,d) + \mathcal{D}''(pd)\,p^2 T' - \left(y_d(x,d) - \mathcal{D}'(pd)\,p\right)^2 T''.$$

By assumptions (i) and (ii), we have $1-T' > 0$ and $T'' \geq 0$. Moreover, we have $y_{d,d} < 0$ by assumption (iii). Hence,

$$k''(d) < \mathcal{D}''(pd)\,p^2 T'.$$

Using $p^2 > 0$ and $T' > 0$, a sufficient condition for the strict concavity of k is that $\mathcal{D}$ is concave in pd, i.e., $\mathcal{D}''(pd) \leq 0$ for every d.

Differentiating the definition of the deduction rule (29) with respect to type w yields

$$\mathcal{D}'(pd(w)) = \frac{1}{p} y_d(x(w), d(w)),$$

and

$$\mathcal{D}''(pd(w)) = \frac{y_{d,d}(x(w), d(w)) + y_{x,d}(x(w), d(w))\frac{x'(w)}{d'(w)}}{p^2}.$$

Since $y_{d,d} < 0$ by assumption (iii), we conclude that $\mathcal{D}'' \leq 0$ holds if the cross-partial term $y_{x,d}\frac{x'}{d'}$ is not too large. In particular, $\mathcal{D}'' \leq 0$ holds if $y_{x,d} = 0$, or if $y_{x,d} \leq 0$ and both $x', d' \geq 0$.

Finally, consider the labor-income-based deduction rule (7). Holding x fixed, the deduction level takes the form $\mathcal{D}(pd) := \alpha(x)pd$, which is linear in d. Consequently, $\mathcal{D}''(pd) = 0$ for all (d, x), and the term involving $\mathcal{D}''$ drops out of the second derivative. Since the remaining terms in $k''(d)$ are negative under assumptions (i), (ii), (iii), the objective remains strictly concave in d. □

C.4 Proof of Proposition 5

This appendix presents the proof of Proposition 5, which characterizes the optimal uniform deduction rate under a fixed income tax schedule. To keep the reform budget neutral, we pair the change in α with a change in the lump-sum demogrant b, which we denote by $\frac{\partial b}{\partial \alpha}$.

By the envelope theorem, the total derivative of the indirect utility function with respect to α equals:

$$u_v\left(\frac{\partial v^*}{\partial \alpha} + \frac{\partial v^*}{\partial b}\frac{\partial b}{\partial \alpha}\right) = U_c \cdot \left(pdT' + \frac{\partial b}{\partial \alpha}\right). \tag{59}$$

Hence, the government's first-order condition for welfare maximization with respect to α is:

$$\frac{\mathrm{d}W}{\mathrm{d}\alpha} = \int_{\underline{w}}^{\overline{w}} \pi \cdot U_c \cdot \left(pdT' + \frac{\partial b}{\partial \alpha}\right) \mathrm{d}F(w) = 0. \tag{60}$$

The government budget constraint is given by

$$b = R - G, \tag{61}$$

where R denotes the revenue raised by the income tax (not including the demogrant b),

$$R = \int_{\underline{w}}^{\overline{w}} T\left(y\left(wl, d\right) - \alpha pd + \gamma\left(wl\right)\right) \mathrm{d}F(w).$$

Below, we will use the facts that

$$R_\alpha := \frac{\partial R}{\partial \alpha} = \mathbb{E}\left[T' \cdot \left(\frac{\partial l}{\partial \alpha}\left(\gamma' + y_x\right) w + \frac{\partial d}{\partial \alpha}\left(y_d - \alpha p\right) - pd\right)\right], \tag{62}$$

and

$$R_b := \frac{\partial R}{\partial b} = \mathbb{E}\left[T' \cdot \left(\frac{\partial l}{\partial b}\,(\gamma' + y_x)\,w + \frac{\partial d}{\partial b}\,(y_d - \alpha p)\right)\right]. \tag{63}$$

It will also be helpful to define h as the change in the tax revenue collected from an individual in response to the individual having an additional dollar of income:

$$h = T' \cdot \left(\frac{\partial l}{\partial b}\,(\gamma' + y_x)\,w + \frac{\partial d}{\partial b}\,(y_d - \alpha p)\right), \tag{64}$$

so that $R_b = \mathbb{E}[h]$.

Totally differentiating the government budget constraint with respect to α, we obtain

$$\frac{\partial b}{\partial \alpha} = R_\alpha + R_b\,\frac{\partial b}{\partial \alpha}, \tag{65}$$

This in turn implies the identity

$$\frac{\partial b}{\partial \alpha} = \frac{R_\alpha}{1 - R_b}. \tag{66}$$

Substituting these results into the government's FOC and using expectations notation we have:

$$\begin{aligned}
0 &= \mathbb{E}\left[\pi \cdot U_c \cdot \left(pdT' + \frac{\partial b}{\partial \alpha}\right)\right] && (67)\\
&= \mathbb{E}[\pi \cdot U_c \cdot pdT'] + \frac{\mathbb{E}[\pi \cdot U_c]}{1 - R_b}\,\mathbb{E}\left[T' \cdot \left(\frac{\partial l}{\partial \alpha}\,(\gamma' + y_x)\,w + \frac{\partial d}{\partial \alpha}\,(y_d - \alpha p) - pd\right)\right] && (68)\\
&= \mathbb{E}[\pi \cdot U_c \cdot pdT'] + \frac{\mathbb{E}[\pi \cdot U_c]}{1 - R_b}\,\mathbb{E}\left[T' \cdot \left(\frac{\partial l}{\partial \alpha}\,(\gamma' + y_x)\,w + \frac{\partial d}{\partial \alpha}\,(y_d - \alpha p)\right)\right] - \frac{\mathbb{E}[\pi \cdot U_c]}{1 - R_b}\,\mathbb{E}[pdT'] && (69)
\end{aligned}$$

Because a marginal increase in α raises income by pdT', a Slutsky decomposition allows us to write:

$$\frac{\partial l}{\partial \alpha} = \frac{\partial l^c}{\partial \alpha} + pdT'\,\frac{\partial l}{\partial b} \tag{70}$$

where $\frac{\partial l^c}{\partial \alpha}$ denotes the compensated (Hicksian) derivative, and similarly for d:

$$\frac{\partial d}{\partial \alpha} = \frac{\partial d^c}{\partial \alpha} + pdT'\,\frac{\partial d}{\partial b}. \tag{71}$$

Substituting these results into the government FOC yields:

$$
\begin{aligned}
0 =& \mathbb{E}\left[\pi \cdot U_c \cdot pdT'\right] + \frac{\mathbb{E}\left[\pi \cdot U_c\right]}{1-R_b}\,\mathbb{E}\left[T' \cdot \left(\frac{\partial l}{\partial \alpha}\,\left(\gamma' + y_x\right) w + \frac{\partial d}{\partial \alpha}\left(y_d - \alpha p\right)\right)\right] - \frac{\mathbb{E}\left[\pi \cdot U_c\right]}{1-R_b}\,\mathbb{E}\left[pdT'\right] \\
=& \mathbb{E}\left[\pi \cdot U_c \cdot pdT'\right] + \frac{\mathbb{E}\left[\pi \cdot U_c\right]}{1-R_b}\,\mathbb{E}\left[T' \cdot \left(\frac{\partial l^c}{\partial \alpha}\,\left(\gamma' + y_x\right) w + \frac{\partial d^c}{\partial \alpha}\left(y_d - \alpha p\right)\right)\right] \\
& + \frac{\mathbb{E}\left[\pi \cdot U_c\right]}{1-R_b}\,\mathbb{E}\left[pdT' \cdot T' \cdot \left(\frac{\partial l}{\partial b}\,\left(\gamma' + y_x\right) w + \frac{\partial d}{\partial b}\left(y_d - \alpha p\right)\right)\right] - \frac{\mathbb{E}\left[\pi \cdot U_c\right]}{1-R_b}\,\mathbb{E}\left[pdT'\right] \\
=& \mathbb{E}\left[\pi \cdot U_c \cdot pdT'\right] + \frac{\mathbb{E}\left[\pi \cdot U_c\right]}{1-R_b}\,\mathbb{E}\left[T' \cdot \left(\frac{\partial l^c}{\partial \alpha}\,\left(\gamma' + y_x\right) w + \frac{\partial d^c}{\partial \alpha}\left(y_d - \alpha p\right)\right)\right] \\
& + \frac{\mathbb{E}\left[\pi \cdot U_c\right]}{1-R_b}\,\mathbb{E}\left[pdT' \cdot h\right] - \frac{\mathbb{E}\left[\pi \cdot U_c\right]}{1-R_b}\,\mathbb{E}\left[pdT'\right] \\
=& \mathbb{E}\left[\left(\pi\, U_c + h\,\frac{\mathbb{E}\left[\pi \cdot U_c\right]}{1-R_b}\right) \cdot pdT'\right] - \frac{\mathbb{E}\left[\pi \cdot U_c\right]}{1-R_b}\,\mathbb{E}\left[pdT'\right] \\
& + \frac{\mathbb{E}\left[\pi \cdot U_c\right]}{1-R_b}\,\mathbb{E}\left[T' \cdot \left(\frac{\partial l^c}{\partial \alpha}\,\left(\gamma' + y_x\right) w + \frac{\partial d^c}{\partial \alpha}\left(y_d - \alpha p\right)\right)\right].
\end{aligned}
\tag{72}
$$

We use the fact that $\mathbb{E}\left[h\right] = R_b$ implies

$$
\mathbb{E}\left[\pi\, U_c + h\,\frac{\mathbb{E}\left[\pi \cdot U_c\right]}{1-R_b}\right] = \frac{\mathbb{E}\left[\pi \cdot U_c\right]}{1-R_b}, \tag{73}
$$

and rewrite the final condition in (72) as

$$
0 = \operatorname{cov}\left(pdT'\,,\, \pi U_c + h\,\frac{\mathbb{E}\left[\pi \cdot U_c\right]}{1-R_b}\right) + \frac{\mathbb{E}\left[\pi \cdot U_c\right]}{1-R_b}\,\mathbb{E}\left[T' \cdot \left(\frac{\partial l^c}{\partial \alpha}\,\left(\gamma' + y_x\right) w + \frac{\partial d^c}{\partial \alpha}\left(y_d - \alpha p\right)\right)\right]. \tag{74}
$$

Dividing this equation by $\frac{\mathbb{E}[\pi \cdot U_c]}{1-R_b}$ and then solving for α yields the result. □

C.5 Proof of Proposition 6

If the tax schedule is optimal, no budget-neutral change to $\gamma(x)$ can be welfare-improving. So our first step will be to repeat the argument used in the proof of Proposition 5 with respect to α, now applied to $\gamma(x)$ for each value of x.

The government budget constraint is given by

$$
b = R - G, \tag{75}
$$

where R denotes the revenue raised by the income tax (not including the demogrant b),

$$R = \mathbb{E}\left[T\left(y\left(x,d\right) - \alpha pd + \gamma\left(x\right)\right)\right]. \tag{76}$$

Totally differentiating the government budget constraint with respect to α, we obtain

$$\frac{\partial b}{\partial \alpha} = R_\alpha + R_b \frac{\partial b}{\partial \alpha}, \tag{77}$$

where $R_\alpha = \frac{\partial R}{\partial \alpha}$ and $R_b = \frac{\partial R}{\partial b}$. This in turn implies the identity

$$\frac{\partial b}{\partial \alpha} = \frac{R_\alpha}{1 - R_b}. \tag{78}$$

By the same logic, we also obtain:

$$\frac{\partial b}{\partial \gamma(x)} = \frac{R_{\gamma(x)}}{1 - R_b}. \tag{79}$$

where $R_{\gamma(x)} = \frac{\partial R}{\partial \gamma(x)}$ denotes the derivative of R with respect to the tax function $\gamma(\cdot)$ at the point x, capturing the marginal change in revenue resulting from an infinitesimal, localized increase of γ at x.

Using this notation, we can express the change in social welfare from a marginal budget-neutral increase in α as:

$$\begin{aligned} \frac{\mathrm{d}W}{\mathrm{d}\alpha} &= \mathbb{E}\left[\pi U_c pdT'\right] + \frac{\partial b}{\partial \alpha}\mathbb{E}\left[\pi U_c\right] \\ &= \frac{1}{1 - R_b}\left(\left(1 - R_b\right)\mathbb{E}\left[\pi U_c pdT'\right] + R_\alpha \,\mathbb{E}\left[\pi U_c\right]\right). \end{aligned} \tag{80}$$

We now establish the FOC with respect to the optimal $\gamma(x)$ for each x. Similar as we did with α, we apply the envelope theorem to get the government's first-order condition for welfare maximization with respect to $\gamma(x)$. That is, for all x we have:

$$0 = -\pi U_c T' g(x) + \frac{\partial b}{\partial \gamma(x)}\mathbb{E}\left[\pi U_c\right], \tag{81}$$

where $g(x)$ denotes the density of x. After substitution of $\partial b/\partial \gamma(x)$ and multiplication, we obtain the equivalent condition:

$$0 = -\pi U_c T'(1 - R_b)pdg(x) + R_{\gamma(x)} pd\,\mathbb{E}\left[\pi U_c\right]. \tag{82}$$

Since the above holds for each x, it also holds when we integrate over x:

$$0 = -(1 - R_b)\mathbb{E}\left[\pi U_c T' pd\right] + \mathbb{E}\left[\pi U_c\right] \cdot \int_x pdR_{\gamma(x)} \mathrm{d}x, \tag{83}$$

where the expectation operator denotes integration with respect to x using the associated probability density $g(x)$, whereas the final integral term represents regular integration with respect to x without density weighting. Substituting this result into (80), we get:

$$\begin{aligned}\frac{\mathrm{d}W}{\mathrm{d}\alpha} &= \frac{1}{1-R_b}\left((1-R_b)\,\mathbb{E}\left[\pi U_c pdT'\right] + R_\alpha\,\mathbb{E}\left[\pi U_c\right]\right) \\ &= \frac{1}{1-R_b}\left(\mathbb{E}\left[\pi U_c\right]\int_x pdR_{\gamma(x)}\mathrm{d}x + R_\alpha\,\mathbb{E}\left[\pi U_c\right]\right) \\ &= \frac{\mathbb{E}\left[\pi U_c\right]}{1-R_b}\left(\int_x pdR_{\gamma(x)}\mathrm{d}x + R_\alpha\right).\end{aligned} \tag{84}$$

By differentiating the tax revenue definition, we now note

$$R_\alpha = \mathbb{E}\left[T' \cdot \left(\frac{\partial l}{\partial \alpha}\left(\gamma' + y_x\right)w + \frac{\partial d}{\partial \alpha}\left(y_d - \alpha p\right) - pd\right)\right], \tag{85}$$

and

$$R_{\gamma(x)} = T' \cdot \left(\frac{\partial l}{\partial \gamma(x)}\left(\gamma' + y_x\right)w + \frac{\partial d}{\partial \gamma(x)}\left(y_d - \alpha p\right) + 1\right)g(x). \tag{86}$$

Multiplying by $pd(x)$ and integrating with respect to x, the latter result implies:

$$\int_x pdR_{\gamma(x)}\mathrm{d}x = \mathbb{E}\left[pdT' \cdot \left(\frac{\partial l}{\partial \gamma(x)}\left(\gamma' + y_x\right)w + \frac{\partial d}{\partial \gamma(x)}\left(y_d - \alpha p\right) + 1\right)\right]. \tag{87}$$

We combine these results to rewrite the key term in (84) as follows:

$$\begin{aligned}\int_x pdR_{\gamma(x)}\mathrm{d}x + R_\alpha &= \mathbb{E}\left[pdT' \cdot \left(\frac{\partial l}{\partial \gamma(x)}\left(\gamma' + y_x\right)w + \frac{\partial d}{\partial \gamma(x)}\left(y_d - \alpha p\right) + 1\right)\right] \\ &\quad + \mathbb{E}\left[T' \cdot \left(\frac{\partial l}{\partial \alpha}\left(\gamma' + y_x\right)w + \frac{\partial d}{\partial \alpha}\left(y_d - \alpha p\right) - pd\right)\right] \\ &= \mathbb{E}\left[T' \cdot \left(\left(\frac{\partial l}{\partial \alpha} + pd\frac{\partial l}{\partial \gamma(x)}\right)\left(\gamma' + y_x\right)w + \left(\frac{\partial d}{\partial \alpha} + pd\frac{\partial d}{\partial \gamma(x)}\right)\left(y_d - \alpha p\right)\right)\right]\end{aligned} \tag{88}$$

We now want to argue that $\frac{\partial l}{\partial \alpha} + pd\frac{\partial l}{\partial \gamma(x)} = 0$. Because of weak separability, we know that

$$\frac{\partial l}{\partial \gamma(x)} = \frac{\partial l}{\partial v^*}\frac{\partial v^*}{\partial \gamma(x)} = \frac{\partial l}{\partial v^*} \cdot \left(-v_c T'(x)\right), \tag{89}$$

and similarly,

$$\frac{\partial l}{\partial \alpha} = \frac{\partial l}{\partial v^*}\frac{\partial v^*}{\partial \alpha} = \frac{\partial l}{\partial v^*} \cdot \left(v_c p d T'(x)\right). \tag{90}$$

Hence, we have

$$\frac{\partial l}{\partial \alpha} + pd\frac{\partial l}{\partial \gamma(x)} = 0. \tag{91}$$

Substituting this result into (88) implies

$$\int_x pdR_{\gamma(x)}\mathrm{d}x + R_\alpha = \mathbb{E}\left[T' \cdot \left(\left(\frac{\partial d}{\partial \alpha} + pd\frac{\partial d}{\partial \gamma(x)}\right)(y_d - \alpha p)\right)\right]. \tag{92}$$

By construction of the joint reform from our main analysis, we have $\frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha} = \frac{\partial d}{\partial \alpha} + pd\frac{\partial d}{\partial \gamma(x)}$. Substituting this result into the above yields:

$$\int_x pdR_{\gamma(x)}\mathrm{d}x + R_\alpha = \mathbb{E}\left[T' \cdot \frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha} \cdot (y_d - \alpha p)\right]. \tag{93}$$

Substituting this back into (84) yields:

$$\frac{\mathrm{d}W}{\mathrm{d}\alpha} = \frac{\mathbb{E}\left[\pi U_c\right]}{1 - R_b} \cdot \mathbb{E}\left[T' \cdot \frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha} \cdot (y_d - \alpha p)\right]. \tag{94}$$

Note that $\pi U_c \geq 0$, and it is strictly positive on a set of positive measure, so that $\mathbb{E}[\pi U_c] > 0$. Moreover, if $R_b > 1$, then the tax system could not be optimal because a lump-sum adjustment to T would be Pareto improving (i.e., giving everyone a dollar would raise more than a dollar). Hence, the pre-multiplying factor on the right-hand side of (94) is positive.

Therefore, $\mathrm{d}W/\mathrm{d}\alpha$ is positive if and only if

$$\mathbb{E}\left[T' \cdot \frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha} \cdot \frac{y_d}{p}\right] > \alpha \cdot \mathbb{E}\left[T' \cdot \frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha}\right]. \tag{95}$$

Equivalently,

$$\frac{\mathbb{E}\left[T' \cdot \frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha} \cdot \frac{y_d}{p}\right]}{\mathbb{E}\left[T' \cdot \frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha}\right]} = \mathbb{E}\left[\frac{T' \cdot \frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha}}{\mathbb{E}\left[T' \cdot \frac{\mathrm{d}\hat{d}}{\mathrm{d}\alpha}\right]} \cdot \frac{y_d}{p}\right] > \alpha. \tag{96}$$

Expressing expectations as integrals over the distribution of skills w, we rewrite this condition as

$$\int_{\underline{w}}^{\overline{w}} \frac{T'\left(z(w)\right)\frac{\mathrm{d}\hat{d}(w)}{\mathrm{d}\alpha}}{\int_{\underline{w}}^{\overline{w}} T'\left(z(\omega)\right)\frac{\mathrm{d}\hat{d}(\omega)}{\mathrm{d}\alpha}\mathrm{d}F(\omega)} \cdot \frac{y_d\left(x\left(w\right), d\left(w\right)\right)}{p}\mathrm{d}F(w) > \alpha. \tag{97}$$

Finally, defining the weights

$$\theta_w = \frac{T'\left(z(w)\right)\frac{\mathrm{d}\hat{d}(w)}{\mathrm{d}\alpha}}{\int_{\underline{w}}^{\overline{w}} T'\left(z(\omega)\right)\frac{\mathrm{d}\hat{d}(\omega)}{\mathrm{d}\alpha}\mathrm{d}F(\omega)}, \tag{98}$$

we express the welfare-improvement condition compactly as

$$\int_{\underline{w}}^{\overline{w}} \theta_w \frac{y_d\left(x\left(w\right), d\left(w\right)\right)}{p}\mathrm{d}F(w) > \alpha. \tag{99}$$

This establishes the result. □