HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

ON THE UNDESIRABILITY
OF COMMODITY TAXATION
EVEN WHEN INCOME TAXATION
IS NOT OPTIMAL

Louis Kaplow

Discussion Paper No. 470

03/2004

Harvard Law School
Cambridge, MA 02138

On the Undesirability of Commodity Taxation Even When Income Taxation Is Not Optimal

Louis Kaplow[*]

Abstract

An important result due to Atkinson and Stiglitz (1976) is that differential commodity taxation is not optimal in the presence of an optimal nonlinear income tax (given weak separability of utility between labor and all consumption goods). This article demonstrates that their conclusion holds regardless of whether the income tax is optimal. In particular, given any commodity tax and income tax system, differential commodity taxation can be eliminated in a manner that results in a Pareto improvement. Also, differential commodity taxation can be proportionally reduced so as to generate a Pareto improvement. In addition, for commodity tax reforms that do not eliminate or proportionally reduce differential taxation, a simple efficiency condition is offered for determining whether a Pareto improvement is possible.

On the Undesirability of Commodity Taxation Even When Income Taxation Is Not Optimal

Louis Kaplow

1. Introduction

Atkinson and Stiglitz's (1976) important article considers the relationship between income taxation and commodity taxation.[1] Most notably, they prove that the social welfare optimum entails no differential commodity taxation, assuming that a nonlinear income tax is optimally employed and that individuals' utility functions are weakly separably between labor and commodities (taken together). This result is of significance because, among other things, it implies that Ramsey (1927) tax principles – and many implications derived from them, such as conclusions regarding the taxation of capital – are displaced by qualitatively different guidelines. See, for example, Atkinson and Stiglitz (1976, 1980) and Stiglitz (1987).

It is therefore of interest to determine whether the result of Atkinson and Stiglitz holds even when the income tax may not be optimal. Standard methods of proof used to characterize optimal schemes – including those in Atkinson and Stiglitz and related papers – rely on arguments that are unavailable when not at an optimum, notably, the first-order conditions and the associated fact that certain effects are second-order when at an optimum. Furthermore, commodity tax reforms may well be contemplated in settings in which one does not know whether the existing income tax is optimal or in which the income tax in fact is not optimal and reform is infeasible, perhaps for political reasons that prevent changing the extent of redistribution.

This article demonstrates that the Atkinson and Stiglitz result holds whatever the income tax is. Also, a notable feature of the analysis is that the method of proof illuminates the intuition behind the conclusions in a very direct way, as outlined in section 2. This method is used in section 3 to demonstrate that, without regard to the nature of the preexisting income tax, under weak separability it is possible to eliminate differential commodity taxation in a manner that produces a Pareto improvement. That is, for any differential commodity tax and arbitrary income tax, there exists an alternative regime with no differential commodity taxation and a different income tax under which everyone is better off.

In section 4, the same technique is used to prove that one can generate a Pareto improvement for commodity tax reforms that do not eliminate differential taxation but only reduce it proportionally. In order to facilitate the analysis, this section introduces a manner of normalizing differential commodity taxation that allows presentation of a clear, unambiguous

---

[1] Historically, income taxation is often referred to as direct taxation and commodity taxation as indirect taxation because of the assumption that only the former can be tailored to individuals' circumstances.

notion of a reform that proportionally reduces (but need not eliminate) commodity tax differentials.

In section 5, commodity tax reforms that do not proportionally reduce or fully eliminate differential taxation are considered. It is demonstrated that a simple efficiency condition – one that depends only on efficiency in consumption, without regard to labor supply – determines whether a Pareto improvement is possible (again assuming weak separability). Concluding remarks are offered in section 6.

The present article differs from most prior work on optimal commodity taxation in the presence of a nonlinear income tax because, as noted, prior work focuses on the full optimum, in which the income tax is taken to be optimal; also, such work does not examine global reforms but instead considers only whether introducing some differential commodity taxation would raise welfare. See, in addition to Atkinson and Stiglitz (1976, 1980), Mirrlees (1976) and Christiansen (1984). (Additionally, Deaton (1979) characterizes the restrictions on utility functions necessary for the Atkinson-Stiglitz result to hold when the optimal income tax is linear.[2]) Some attention has also been devoted to commodity taxation with an arbitrary initial income tax.[3] Most relevant is Konishi (1995), whose technical exploration of local commodity tax reforms under restrictive assumptions does not set forth the results derived here or offer an intuitively accessible understanding of the problem.[4]

2. Intuitive Explanation of the Result

It is useful to begin with Atkinson and Stiglitz's (1976) result that no differentiated commodity taxation is optimal in the presence of an optimal nonlinear income tax, given weak separability. The income tax can be used both to raise needed revenue and to redistribute income. Why, then, might one want to employ differential commodity taxation? The motivation would seem to be to raise revenue or redistribute income with less distortion, specifically of the labor/leisure decision, than is caused by the income tax. However, given weak separability, differential commodity taxation cannot lessen this distortion, but it does introduce distortions in commodity choices.

---

[2]A substantial literature, not relevant to the present investigation, extends Ramsey (1927) by examining optimal commodity taxation and commodity tax reforms when there is no income tax.

[3]Yang and Haller (1993), in contrast to the present investigation, assume that the nonlinear income tax remains fixed; in addition, they use a surplus measure for welfare rather than examining Pareto improvements or invoking a social welfare function to assess whether there is a welfare gain.

[4]Konishi (1995) introduces numerous assumptions about utility and the preexisting income tax to guarantee that the optimization problem is locally well-behaved. He further assumes (implicitly) that any marginal change in a commodity tax rate can be accompanied by a change in the income tax schedule that satisfies the revenue constraint and everyone's incentive compatibility constraint and also simultaneously redistributes income so that either everyone gains or everyone loses from the reform. Using these important assumptions, he shows (as one might expect) that either a marginal increase or decrease in a commodity tax rate raises everyone's welfare, or else that no change is optimal – where the condition for a local optimum requires that the integrals over the population of all pertinent effects on each type of individual balance to zero.

Upon examination, this basic intuition does not depend on the assumption that the preexisting income tax is set optimally. To be sure, as noted in the introduction, conditions true at an optimum often do not hold when one deviates from the optimum. Nevertheless, in the present context it can be shown that the two problems – the optimal nonlinear income tax problem, which trades off redistribution and labor supply distortion, and the optimal commodity tax problem, which involves relative prices of different commodities – are in an important sense orthogonal to each other. In particular, commodity tax reform can be isolated from the income tax, so there is no reason to refrain from maximizing efficiency with regard to the former on account of shortcomings regarding the latter.

This isolation is accomplished in the proofs in sections 3, 4, and 5 by combining a reform in commodity taxes that reduces or eliminates the targeted inefficiency – here, distortion in the relative prices of commodities – with an adjustment of the income tax schedule that keeps everything else constant – notably, the extent of redistribution and of labor supply distortion. Specifically, when one reduces distortionary commodity taxation, one can simultaneously imagine adjusting the income tax schedule to offset any effects on individuals' utility at every level of income.[5] For example, if individuals at a given income level gain from commodity tax reform because they pay less in commodity taxes and also benefit from adjusting their pattern of consumption, the income tax at that level of income can be raised by just enough to offset this utility gain. When such adjustments are made at every level of income, it turns out that not only is utility (and thus the distribution of utility) the same – which is true by construction – but also labor effort is unaffected (if one makes the weak separability assumption used by Atkinson and Stiglitz). It follows that the only net effect of this reform is on revenue.

It remains to be shown that combining seemingly efficient commodity tax reform with the hypothesized adjustment to the income tax schedule increases revenue. An increase in fact occurs because the reduction in distortion of consumption is not a mere transfer but a real savings in resources that serves to raise individuals' utilities. Therefore, the income tax adjustment, which is defined as that which keeps everyone's utility constant, must be taxing away this increase in utility that otherwise would obtain. Accordingly, the regime that reduces inefficiencies in commodity taxation and employs an income tax adjustment that keeps utility constant necessarily yields a surplus. One can, therefore, construct an actual income tax adjustment that distributes this surplus to all individuals, generating a Pareto improvement.

On reflection, the feasibility of the aforementioned reform strategy should not be surprising. At its core, the Atkinson-Stiglitz result does not depend on whether the existing income tax schedule is determined optimally, for such optimality only insures that society is making the appropriate tradeoff of redistribution and labor supply distortion. If one could hold the extent of redistribution and labor supply distortion fixed, one expects that reducing distortions in commodity taxation would raise social welfare. And if one indeed holds constant the overall distribution of well-being, then the gain in total welfare is being allocated in part to

---

[5]It appears that this technique is first employed in Hylland and Zeckhauser (1979), in the context of examining the propriety of distributive adjustments to cost-benefit analysis. For further discussion and applications, see Kaplow (1996, 2004).

each individual, which makes everyone better off. The analysis to follow formalizes this basic insight.

3. Elimination of Differential Commodity Taxation

Individuals choose levels of consumption of n commodities, x_1, ..., x_n, and of labor effort l to maximize their utility functions $u(v(x_1, ..., x_n),l)$, where v is a subutility function. The utility functions are assumed to be continuously differentiable, strictly concave, increasing in commodities, and decreasing in labor effort. This form of the utility function entails what is referred to as weak separability of labor (or leisure): For a given level of after-income-tax income, individuals will allocate their disposable income among commodities in the same manner regardless of the level of labor effort required to earn that level of income.

Individuals earn income wl that depends on their wage (type) w, which has density $f(w)$. Commodity prices for good x_i (which equal production costs measured in units of income) are p_i, assumed to be greater than zero. There is a (nonlinear) income tax schedule $T(wl)$ and commodity taxes on each good x_i of τ_i (which may be subsidies, in which case they are negative, but we restrict them so that $\tau_i > -p_i$, so all net prices are positive). Individuals thus face net prices of $p_i + \tau_i$.

An individual of type w's budget constraint can be written as

$$(1) \quad \sum (p_i + \tau_i)x_i(wl) = wl - T(wl),$$

where summations throughout are from i equals 1 to n and the notation $x_i(wl)$ denotes the level of x_i chosen by an individual of type w and thus income of wl, where l implicitly refers to the labor effort of an individual of type w. The government's budget constraint is

$$(2) \quad \int \left[T(wl) + \sum \tau_i x_i(wl) \right] f(w)dw = R,$$

where R is a given revenue requirement.

The approach is to begin with a differentiated tax system and attempt to construct an undifferentiated tax system that makes everyone better off. First, it is useful to define terms.

Differentiated tax system. A differentiated tax system $\{\tau_1, ..., \tau_n\}$, $T(wl)$ is one for which there exists i, j such that $(p_i+\tau_i)/(p_j+\tau_j) \neq p_i/p_j$.

In other words, the taxes and subsidies much be such that the price ratio of at least one pair of goods does not equal its production-cost ratio.

Undifferentiated tax system. An undifferentiated tax system $\{\tau_1^*, ..., \tau_n^*\}$, $T^*(wl)$ is one that is not differentiated, i.e., one for which $(p_i+\tau_i^*)/(p_j+\tau_j^*) = p_i/p_j$, for all i, j.

As discussed further in section 4, there are an infinite variety of equivalent ways to describe any commodity tax - income tax system. For example, an undifferentiated system can involve $\tau_i^* = 0$, for all i, or instead one could have $\tau_i^* = \alpha\, p_i$, for all i, with a corresponding adjustment to the income tax schedule. For example, if $\alpha > 0$, everyone pays proportionally more for any commodity vector; hence, the income tax schedule can be reduced so that everyone's after-income-tax income is greater by the same proportion. Note that this composite adjustment to the commodity tax - income tax system is revenue neutral; the additional commodity tax revenue just offsets the reduced income tax revenue.

To begin the construction of a Pareto-improving tax reform, start with an initial, differentiated regime $\{\tau_1, ..., \tau_n\}$, $T(wl)$. For simplicity, choose from among the multitude of equivalent undifferentiated tax systems the one for which $\tau_i^* = 0$, for all i. Moving to this commodity tax vector will tend to change individuals' utility because they no longer pay commodity taxes (or receive subsidies) and because, with a new relative price vector, they will change their consumption vector. Whatever is the net effect on utility for any type w and given labor effort $l(w)$, we can now define an intermediate income tax schedule $T^\circ(wl)$ at each income level so as to offset the net effect on utility.[6] That is, we will examine an income tax schedule $T^\circ(wl)$ that has the property that, if all individuals (of every type w) continue to choose the same level of labor effort $l(w)$ as under the initial tax system, then their utility will be unchanged.[7] (Whether individuals will choose the same labor effort under this intermediate regime is the subject of Lemma 1, below.) That it is possible to construct such an income tax schedule for each income level is due to the continuity of utility in income.[8]

Observe that this intermediate income tax schedule $T^\circ(wl)$ has the property that it leaves subutility v unaffected for all levels of income (again, continuing to assume that labor effort is unchanged). That is, stated in reduced form, $v(wl) = v^\circ(wl)$ for all wl. This result about the subutility functions must be true because the income tax schedule is constructed such that $u(v(wl),l) = u^\circ(v^\circ(wl), l)$; because the u function does not change and l is assumed to be constant, it must be that the levels of subutility are unchanged. (Changing each of commodity taxes and the income tax schedule in general alters the level of subutility v produced by a given level of

---

[6]If two types initially earn the same income (a possibility that could be ruled out with appropriate assumptions), the same adjustment to the income tax schedule would work for both types due to the weak separability assumption.

[7]It is not asserted at this point that the tax schedule $T^\circ(wl)$ is feasible; in Lemma 2, it will in fact be shown to generate a surplus. For purposes of the analysis, it is simply a hypothetical, intermediate construct. Only the final schedule, $T^*(wl)$, needs to be feasible.

[8]It is straightforward that one could define an intermediate income tax sufficiently high that an individual's utility would be below what it is under the initial regime; consider a level of income tax that leaves the individual with insufficient funds to purchase the lowest of the initial x_i's if the price were the lowest of the p_i's. Likewise, one could set an intermediate income tax sufficiently low to guarantee that the individual's utility would be above what it is under the initial regime; consider a level of tax that leaves the individual with sufficiently great funds to purchase each commodity at a level above the highest of the initial x_i's if the prices on all commodities equalled the highest of the p_i's. Since both levels of income tax, guaranteeing a lower and a higher utility level than in the initial regime, are possible, by continuity there will exist an intermediate level of income tax that can be imposed at the given income level such that the individual's utility will be the same as it is in the initial regime. And this is possible for all income levels, $wl(w)$.

income wl; however, under the stated assumptions, these effects must be offsetting.) This characteristic of the intermediate income tax schedule $T°(wl)$ can be employed to establish the following result.

Lemma 1: Every type of individual w chooses the same level of labor effort $l(w)$ under $\{\tau_1*, ..., \tau_n*\}$, $T°(wl)$ as under $\{\tau_1, ..., \tau_n\}$, $T(wl)$.

Proof: Consider an individual of any type w. Under the initial regime, the level of labor effort, $l(w)$, selected is that which maximizes utility. Under the intermediate regime with income tax schedule $T°(wl)$, it was just shown that, for any level of income wl, and thus for any choice of l, the subutility $v°$ under the intermediate regime precisely equals the subutility v under the initial regime. This in turn implies that $u(l(w)) = u°(l(w))$ for all $l(w)$, where the reduced form $u(l(w))$ refers to the level of utility achieved for any choice of l by the given type w. Now, since utility as a function of labor effort is precisely the same under the new, intermediate regime as it is under the initial regime, it follows that whatever $l(w)$ maximizes $u(l(w))$ also maximizes $u°(l(w))$. Q.E.D.

The next question is how revenue compares between the initial regime and the intermediate regime.

Lemma 2: Regime $\{\tau_1*, ..., \tau_n*\}$, $T°(wl)$ (with undifferentiated taxes $\tau_i* = 0$, for all i) raises more revenue than regime $\{\tau_1, ..., \tau_n\}$, $T(wl)$.

Proof: The strategy will be to show that no individuals under the intermediate regime $\{\tau_1*, ..., \tau_n*\}$, $T°(wl)$ can still afford the consumption vector purchased under the initial regime $\{\tau_1, ..., \tau_n\}$, $T(wl)$, and that the only way this can be true is if each pays more tax (from commodity taxes and the income tax combined) under the intermediate regime than under the initial regime. (Throughout, labor effort of all types will be taken to be the same under the two regimes, as established in Lemma 1.)

First, suppose that, under the intermediate regime, an individual of some type w can afford the consumption vector from the initial regime. Observe that such an individual would not in fact choose the same consumption vector but rather will choose a different one under the intermediate regime because of the change in relative prices. This follows because the individual's optimal consumption vector is determined by standard first-order conditions, $v_i/v_j = (p_i+\tau_i)/(p_j+\tau_j)$, for all i, j. Given the definition of a differentiated tax system, that there exists i, j such that $(p_i+\tau_i)/(p_j+\tau_j) \neq p_i/p_j$, and the fact that the move to an undifferentiated system (as exists under the intermediate regime) eliminates the discrepancy between the tax-inclusive price ratio and the ratio of production costs, at least one of these first-order conditions no longer holds. As a consequence, the consumption vector that was optimal initially cannot be optimal under the hypothesized intermediate regime. It follows that utility must be higher under the intermediate regime. But this is a contradiction because the intermediate regime's income tax schedule $T°(wl)$ is constructed to keep every individual's utility constant.

Second, using the budget constraint (1), the conclusion that no individual in the

intermediate regime can afford the consumption bundle from the original regime means that

$$(3) \sum (p_i + \tau_i^*) x_i(wl) > wl - T^\circ(wl), \text{ for all } wl.$$

Recalling that $\tau_i^* = 0$, for all i, and using the budget constraint (1) for the initial regime to substitute for $\sum p_i x_i(wl)$ on the left side of expression (3), we obtain

$$(4) \ wl - T(wl) - \sum \tau_i x_i(wl) > wl - T^\circ(wl), \text{ or}$$

$$T^\circ(wl) > T(wl) + \sum \tau_i x_i(wl), \text{ for all } wl.$$

In the second line of expression (4), the left side is an individual's total tax payments under the intermediate regime and the right side is total payments under the initial regime. Because every type of individual pays more under the intermediate regime, total revenue is higher. Q.E.D.

To complete the argument, construct $T^*(wl)$ from $T^\circ(wl)$ as follows:

$$(5) \ T^*(wl) = T^\circ(wl) - c,$$

where c is the positive constant such that the government's budget constraint (2) is satisfied using $T^*(wl)$. That is, beginning from $T^\circ(wl)$, we can reduce everyone's income tax by the same dollar amount until the budget balances. Since there is a surplus when the income tax schedule is set at $T^\circ(wl)$, there are funds from which everyone's income tax payments can be reduced. As this reduction is made, individuals may change their purchases of commodities and labor effort, but however they choose to do so, their utility will increase (since they have more income and relative prices are unchanged). As one continuously increases the rebate, beginning from zero, at some $c > 0$ the government's budget will just balance. (Note that, as the income tax schedule is reduced, individuals may reduce labor supply and thus tax revenue may fall, but as long as there initially is a surplus and behavior is continuous, some net reduction in everyone's income tax payments will be possible.[9])

At this point, we have constructed a new income tax schedule, to accompany an undifferentiated commodity tax schedule, such that every type of individual is strictly better off, thereby establishing the following result:

Proposition 1: For any differentiated tax system $\{\tau_1, ..., \tau_n\}$, $T(wl)$, there exists an

---

[9]This result necessarily holds if individuals' behavior is continuous. However, depending on the shape of the income tax schedule and of individuals' utility functions, discontinuities are possible (notably, with a convex tax schedule, some individuals might discontinuously reduce their labor supply). Nevertheless, minimal assumptions on the distribution of types is sufficient to guarantee that, at any level of rebate, such individuals comprise a set of measure zero, so tax revenue will be continuous and the stated adjustment will be feasible.

undifferentiated tax system $\{\tau_1{}^*, ..., \tau_n{}^*\}$, $T^*(wl)$ that is strictly Pareto superior – i.e., $u^*(w) > u(w)$, for all w.

4. Reduction of Differential Commodity Taxation

Most explorations of optimal taxation do not explore the optimality of partial reforms. If one relies on properties of the optimum (and in a neighborhood thereof), it may be difficult to make statements about changes away from the optimum. The present approach, however, can readily be adapted to consider partial reforms.

4.1. Normalization

To deal with partial commodity tax reform, it is helpful to introduce a normalization that clarifies the relationship among commodity tax vectors and enables a notationally simple definition of partial reform and proof of the result. The issue arises because, for any commodity tax vector and associated income tax schedule, there exists an infinite variety of equivalent systems. For example, suppose that one changes all commodity taxes τ_i to $\tau_i{}' = \lambda(p_i+\tau_i)-p_i$, for some positive constant λ. This leaves all relative prices unchanged – i.e., the prices individuals face are all changed proportionally. Suppose further that one adjusts the income tax schedule by an amount that offsets at every income level the changes in commodity tax payments or receipts. For example, if $\lambda = 0.9$, income taxes would be raised so as to leave individuals at every income level with 10% less disposable income. Such transformations leave individuals' choices of commodities and of labor effort unaffected, produce the same level of utility, and generate the same total tax payments from each individual (and thus in aggregate).

Observe that the use in the proof in section 3 of a scheme in which $\tau_i{}^* = 0$, for all i, as the undifferentiated system was thus merely a matter of convenience; any system with rates in equal proportion to commodity prices would have sufficed. However, when attempting even to define unambiguously a partial reduction of differentiation in commodity taxation, the matter is somewhat more complicated. To assist in this process, it is helpful to introduce a particular normalization, one that maintains the price ratios faced by individuals but results in the commodity tax system raising no net revenue (implying that there will be both taxes and subsidies).

Normalized tax system. A normalized version of the tax system $\{\tau_1, ..., \tau_n\}$, $T(wl)$, denoted $\{\tau_1{}^N, ..., \tau_n{}^N\}$, $T^N(wl)$, is a tax system with the following characteristics:

1. $\sum \tau_i{}^N \int x_i(wl) f(w)dw = 0.$

2. $\dfrac{p_i + \tau_i}{p_j + \tau_j} = \dfrac{p_i + \tau_i{}^N}{p_j + \tau_j{}^N}$, *for all i, j.*

3. $T^N(wl) = T(wl) + \sum \tau_i x_i(wl) - \sum \tau_i{}^N x_i(wl)$, *for all wl.*

The first requirement is that the commodity taxes raise no net revenue. The second states that price ratios remain the same. The third adjusts the income tax schedule, at each level of income, by the difference in commodity tax payments (subsidy receipts); the result is that, at every level of income, total tax payments are the same under the normalized tax system as they are under the original tax system.

If one wishes to construct this normalized system, observe that requirement 2 provides n-1 independent equations and requirement 1 adds another, allowing one to solve uniquely for the τ_i^N's.[10] Once this is done, the income tax adjustment specified by requirement 3 can be computed. (Note that, once a tax system is normalized, requirement 1 implies that if any τ_i^N is nonzero, the system must be differentiated.)

It should be emphasized that normalizing a tax system has no real effects. Because individuals face the same price ratios at any level of income they choose to earn (requirement 2) and have enough after-income-tax income to just afford their original consumption vector for any level of earnings (requirement 3), they will all continue to choose the same consumption vector and level of labor effort. Furthermore, requirement 3 guarantees that each individual will make the same tax payments under the normalized system, so total revenue is unaffected.

4.2. Analysis

Any differentiated tax system can be normalized, as just defined, to produce an equivalent tax system. It will be convenient to assume, without loss of generality, in the analysis that follows that the initial differentiated tax system under consideration is a normalized one (and the "*N*" superscripts will be dropped). The partial tax reform to be considered will be one under which $\tau_i^* = \alpha\tau_i$, for all *i*, where $\alpha \in (0,1)$.

If one reviews the derivation in section 3 (implicitly for the special case in which $\alpha = 0$), it will be apparent that all of the analysis is applicable, mutatis mutandis, to this proportional partial move toward an undifferentiated tax system except for the proof of Lemma 2. Specifically, one can define the intermediate income tax regime $T°(wl)$ that holds utility constant if labor effort is constant in the same manner. Lemma 1 still holds because it only depends on the properties of $T°(wl)$. Finally, if Lemma 2 could be established for this more general case allowing for partial reforms, then the existence of a surplus can be used to complete the proof as before.

---

[10]Solving these equations yields

$$\tau_i^N = -\frac{\sum_j \left\{\left[p_j - p_j\left(\frac{p_j + \tau_j}{p_i + \tau_i}\right)\right]\int x_j(wl)f(w)dw\right\}}{\sum_j \left\{\left(\frac{p_j + \tau_j}{p_i + \tau_i}\right)\int x_j(wl)f(w)dw\right\}}.$$

Accordingly, let us reconsider Lemma 2, which states that more revenue is raised in the intermediate regime with $T°(wl)$ than under the initial regime. The proof used the fact (in moving from expression (3) to (4) and in interpreting the latter) that individuals' changes in their consumption vectors would have no effect on commodity tax revenue because the reform under consideration had $\tau_i^* = 0$, for all i. For a partial reform, individuals' changes in consumption will affect commodity tax revenue, so additional analysis is necessary to demonstrate the claim, which may be restated as follows:

Lemma 2': Regime $\{\tau_1^*, ..., \tau_n^*\}$, $T°(wl)$ (with commodity taxes $\tau_i^* = \alpha\tau_i$, for all i) raises more revenue than regime $\{\tau_1, ..., \tau_n\}$, $T(wl)$.

Proof: The proof of Lemma 2 remains valid, mutatis mutandis, through the derivation of expression (3). However, because we can no longer use the fact that $\tau_i^* = 0$, for all i, expression (4) becomes

$$(4') \quad T°(wl) + \sum \tau_i^* x_i(wl) > T(wl) + \sum \tau_i x_i(wl), \quad \text{for all } wl.$$

Because the second term on the left side of $(4')$ has x_i instead of $x_i°$, the left side cannot be interpreted as total revenue under the intermediate regime. However, this expression can still yield results regarding income tax revenues. In expression $(4')$, substitute $\alpha\tau_i$ for τ_i^* and integrate each side over the population of types to yield

$$(6) \quad \int T°(wl) f(w)dw + \alpha \sum \tau_i \int x_i(wl) f(w)dw >$$
$$\int T(wl) f(w)dw + \sum \tau_i \int x_i(wl) f(w)dw,$$

$$\text{or}$$

$$\int T°(wl) f(w)dw > \int T(wl) f(w)dw,$$

where the latter version follows because the summations on each side both equal zero due to the assumption that the τ_i's are normalized (recall requirement 1). This establishes that income tax revenue is greater in the intermediate regime than in the initial regime.

It can also be demonstrated that commodity tax revenues are higher in the intermediate regime – i.e., that they are positive rather than zero. To prove this, first observe that, for any type earning a given level of income, the $x_i°$'s produce the same level of total utility as the x_i's because the intermediate regime was constructed to yield equal utility. Next, note that, under the initial tax regime, it must be that the $x_i°$'s cannot be afforded. (Suppose they could be. We know, using the same sort of argument as in the original Lemma 2, that the first-order conditions for utility maximization would be violated, so an individual would choose a different commodity vector and achieve higher utility. But this contradicts the construction of the intermediate regime, which ensures that utility must be the same.) From the individual's budget constraint (1), this conclusion implies

(7) $\sum (p_i + \tau_i)x_i^{\circ} > \sum (p_i + \tau_i)x_i$, for all wl.

Solving for wl using the budget constraint (1) for each regime and equating the two yields

(8) $T^{\circ}(wl) + \sum (p_i + \tau_i^{*})x_i^{\circ}(wl) = T(wl) + \sum (p_i + \tau_i)x_i(wl)$, for all wl.

That is, for a given level of income, all of it must be spent on income tax payments and commodities (producer prices and commodity taxes) in each regime. Solving (8) for $\sum p_i x_i^{\circ}$ and substituting this into the left side of (7) gives us

(9) $\sum (p_i + \tau_i)x_i(wl) + T(wl) - T^{\circ}(wl) + \sum \tau_i x_i^{\circ}(wl) - \sum \tau_i^{*}x_i^{\circ}(wl) >$
$$\sum (p_i + \tau_i)x_i(wl), \text{ for all } wl.$$

Simplifying and rearranging terms, we have

(10) $\sum (\tau_i - \tau_i^{*})x_i^{\circ}(wl) > T^{\circ}(wl) - T(wl)$, for all wl.

The intuition behind expression (10) is as follows: The left side is the additional commodity tax burden in the initial regime (versus the intermediate regime) if one were to purchase the intermediate regime's commodity tax vector instead. This added burden must exceed the amount by which income taxes are lower in the initial regime in order for this commodity vector to be unaffordable in the initial regime.

To complete the argument, on the left side of (10) we can substitute for τ_i using the fact that $\tau_i^{*} = \alpha\tau_i$ and then integrate each side over the population of types to yield

(11) $\left(\frac{1}{\alpha} - 1\right)\sum \tau_i^{*}\int x_i^{\circ}(wl)f(w)dw > \int T^{\circ}(wl)f(w)dw - \int T(wl)f(w)dw$.

The right side, the difference in income tax revenues, was already proved to be positive (see expression 6). Because we are assuming that $\alpha \in (0,1)$, the first component on the left side is positive; hence, the summation component must be positive. That component, in turn, is the total commodity tax revenue raised in the intermediate regime. Therefore, it must exceed commodity tax revenue in the initial regime, which equals zero (because a normalized commodity tax vector is employed).

The result from expression (6) that income tax revenue is greater in the intermediate regime combined with the just-obtained result that commodity tax revenue is also greater in that regime establishes that total revenue is greater in the intermediate regime than in the initial

regime. Q.E.D.

As stated above, the remainder of the proof of Proposition 1 follows from this conclusion as before, so now we can state:

Proposition 2: Beginning with any normalized differentiated tax system $\{\tau_1, ..., \tau_n\}$, $T(wl)$, for any tax reform such that $\tau_i^* = \alpha\tau_i$, for all i, where $\alpha \in (0,1)$, there exists $T^*(wl)$ such that the reform regime is strictly Pareto superior – i.e., $u^*(w) > u(w)$, for all w.

Furthermore, for any differentiated tax system that is not normalized one can define an equivalent normalized tax system, so it follows that one can proportionally reduce the degree of differentiation in any differentiated tax system in a manner that results in a Pareto improvement.[11]

On reflection, it should not be surprising that Lemma 2′ could be established and, accordingly, the result in Proposition 1 could be extended to partial reform of commodity taxation. When one begins from a normalized commodity tax system that raises no revenue in the aggregate and implements a uniform proportional move toward zero, consumption changes will tend to involve shifts from commodities that were subsidized (but now are subsidized less) to commodities that were taxed (but now are taxed less), and from lower to higher taxed commodities (because the tax differential is now less), and from highly to less highly subsidized commodities (because the subsidy differential is now less). All such shifts raise additional revenue. The argument is not quite so simple because, not having greatly restricted the form of utility functions, this need not be true with regard to every commodity. (For example, purchasing more of a commodity that is taxed might result in a reduction of purchases of some substitute that is even more heavily taxed.) In addition, the argument is not immediate because disposable income differs due to the reform, and again without further restricting utility functions one cannot rule out shifts in the commodity vector that may have revenue-reducing effects. The analysis of Lemma 2′ establishes, however, that the intuitively expected tendency is indeed correct in the aggregate, which is all that is necessary to prove the more general proposition.

5. Other Reforms of Differential Commodity Taxation

The partial commodity tax reforms encompassed by Proposition 2 have a proportional, uniform character. Obviously, many other partial reforms could also result in Pareto improvements. The reason for the restriction in the proposition is that otherwise it is difficult even to define unambiguously what a reduction in differentiation means. For example, if there where three commodities, two taxed at 10% and the third untaxed, would moving one of the 10% rates to 8% be a reduction in differentiation? Distortion would be reduced between the

---

[11]Using the method of the proof of Proposition 2, one could first normalize the tax system and then undertake the proportional reduction in differentiation. Afterward, one could add a further step, to reverse the initial normalization. Thus, one could simply reduce the differentiation without normalizing. (It should be apparent that the normalization was not necessary to prove the general claim but rather was only for expositional convenience.)

commodity for which the tax rate is reduced and the untaxed commodity, but distortion would be introduced between the two taxed commodities. Accordingly, one cannot say a priori whether overall distortion would be reduced. Suppose, however, that one could determine for a given commodity tax reform whether distortion in consumption – in a simple, traditional sense – was reduced. Would that be sufficient for a Pareto improvement to be possible given an arbitrary income tax and the existence of labor supply distortion? With weak separability, it turns out that the answer is affirmative.

Begin with any differentiated tax system $\{\tau_1, ..., \tau_n\}$, $T(wl)$ and consider any commodity tax reform τ_i^*. As in the proofs of Propositions 1 and 2, define an intermediate regime with income tax schedule $T°(wl)$ that has the property that, if all individuals (of every type w) continue to choose the same level of labor effort $l(w)$ as under the initial tax system, then their utility will be unchanged.

First observe that Lemma 1, stating that labor effort is indeed unchanged, once again holds because the analysis depends only on the manner in which $T°(wl)$ is constructed (and weak separability). Let us now define a broader set of commodity tax reforms that are of interest.

Efficiency-increasing commodity tax reform. For any tax system $\{\tau_1, ..., \tau_n\}$, $T(wl)$, a commodity tax reform $\{\tau_1^*, ..., \tau_n^*\}$ is efficiency increasing if, when combined with the income tax schedule $T°(wl)$:

$$(12) \quad \sum p_i \int x_i°(wl) f(w)dw < \sum p_i \int x_i(wl) f(w)dw.$$

Expression (12) states that the total real resource cost of everyone's consumption vectors in the intermediate regime is less than the total real resource cost in the initial regime. Because everyone's utility is the same in these two regimes, this condition indicates that the intermediate regime is more efficient with regard to consumption choices in a narrow, conventional sense – i.e., when concerns with the labor/leisure distortion and distortionary income taxation are ignored.

Given an efficiency-increasing commodity tax reform, it is straightforward to show that a version of Lemma 2 holds.

Lemma 2″: Any efficiency-increasing commodity tax reform $\{\tau_1^*, ..., \tau_n^*\}$, $T°(wl)$ (with taxes τ_i^* such that (12) holds) raises more revenue than regime $\{\tau_1, ..., \tau_n\}$, $T(wl)$.

Proof: For each of the budget constraints (1) for these two regimes, integrate them over the population of types, subtract one from the other, and rearrange terms, to yield:

$$(13) \left[\int T^\circ(wl)f(w)dw + \sum \tau_i^* \int x_i^\circ(wl)f(w)dw \right]$$
$$- \left[\int T(wl)f(w)dw + \sum \tau_i \int x_i(wl)f(w)dw \right] =$$
$$\sum p_i \int x_i(wl)f(w)dw - \sum p_i \int x_i^\circ(wl)f(w)dw.$$

The first bracketed term on the left side of expression (13) is total revenue under the intermediate regime and the second bracketed term is total revenue under the initial regime. The right side is positive according to expression (12), the definition of an efficiency-increasing commodity tax reform. Q.E.D.

Accordingly, we can complete the argument as before, which establishes:

Proposition 3: Beginning with any tax system $\{\tau_1, ..., \tau_n\}$, $T(wl)$, for any efficiency-increasing commodity tax reform τ_i^*, there exists $T^*(wl)$ such that the reform regime is strictly Pareto superior – i.e., $u^*(w) > u(w)$, for all w.

In essence, Proposition 3 states that if a commodity tax reform increases efficiency in a traditional sense – i.e., if it increases surplus in a world in which labor supply is constant, tantamount to a world with fixed labor supply or simply one in which initial wealth endowments are given – then the reform will be desirable, indeed strictly Pareto improving, when combined with an appropriate income tax adjustment, even in a world in which labor supply is not constant and there exists a distortionary labor income tax.

The intuition behind this result parallels the analysis just presented: If a commodity tax reform increases efficiency, this means that fewer resources are needed for individuals to achieve their initial levels of utility. Because individuals thus do not need to spend as much (aside from commodity taxes) in the hypothesized intermediate regime, total tax collections must be greater for given income levels, and the resulting surplus can be distributed in a manner that yields a Pareto improvement. Furthermore, when the income tax is adjusted in a manner that accomplishes this, labor supply effects do not interfere with the argument. This final Proposition therefore reinforces the sense in which the commodity tax problem and the income tax problem can be viewed as independent (given the assumption of weak separability).[12]

---

[12]It might appear that Proposition 3 is more general than the first two propositions and therefore subsumes them; however, this is not the case. An analogue to Lemma 2 is indeed easy to establish for an efficiency-increasing commodity tax reform, but the full analysis of Lemma 2 or Lemma 2′, as the case may be, is necessary to show that expression (12) holds for the elimination or reduction of differential commodity taxation. In essence, demonstrating that total tax revenue is greater under the intermediate regime and demonstrating that fewer productive resources are utilized amount to the same thing. (One might describe expression (13) as a sort of accounting identity when a regime is constructed with $T^\circ(wl)$, as in the proof of all three propositions.)

6. Conclusion

Differential commodity taxation distorts individuals' consumption choices and thus is presumptively inefficient. However, in second-best settings, such presumptions may be overcome. Nevertheless, Atkinson and Stiglitz (1976) show that, when there is present a nonlinear income tax that is set optimally, the presumption holds: Differential commodity taxation is indeed inefficient. The explanation for the result is that the other relevant distortion involves labor supply, and when weak separability is assumed, commodity taxes offer no leverage for lessening that distortion.

The present analysis extends Atkinson and Stiglitz's result in three ways. First, it shows that, without regard to the optimality of the preexisting income tax, one can always eliminate differential commodity taxation in a manner that produces a Pareto improvement. Second, it shows that a Pareto improvement is likewise possible for any partial reform of commodity taxation that proportionally reduces but does not eliminate differential taxation. Third, with regard to commodity tax reforms that do not proportionally reduce or eliminate differential taxation, it offers a simple efficiency condition for determining whether a Pareto improvement is possible.

The framework for these results, like that of Atkinson and Stiglitz (1976), employs weak separability. In addition, the approach makes the intuition transparent by using an income tax adjustment that holds the extent of redistribution and the labor/leisure distortion constant, so in essence the only remaining effects are the narrow efficiency consequences of commodity tax reform. Finally, because the income tax adjustment employed in demonstrating the possibility of a Pareto improvement – in combination with commodity tax reform – has the characteristic of being distribution neutral, such a reform strategy may possess some political plausibility as well.[13]

Numerous qualifications to Atkinson and Stiglitz's (1976) result are familiar from their article and from other literature on optimal commodity taxation in the presence of an optimally determined nonlinear income tax.[14] Most notably, the assumption of leisure separability rules out the possibility that differential commodity taxation might help to offset the labor/leisure distortion from the income tax. Without separability, it might be possible to tax complements to leisure and to subsidize complements to labor, improving efficiency by reducing the labor/leisure distortion. (Formally, relaxing this assumption would affect Lemma 1: If, instead of being unaffected by the intermediate reform, labor supply were increased (reduced), the reform would be more (less) favorable than otherwise because of the additional positive (negative) effect on tax revenue.) Because this qualification and most others associated with the Atkinson and Stiglitz

---

[13]To verify the distributional characteristics of the reform as a whole, observe first that $T°(wl)$ is constructed such that, in combination with the reform of differential commodity taxation, everyone's utility is held constant, and the final income tax schedule $T*(wl)$ is constructed from $T°(wl)$ by subtracting a constant, as indicated by expression (5).

[14]See, for example, Cremer, Pestieau, and Rochet (2001), Marchand, Pestieau, and Racionero (2003), Mirrlees (1976), Naito (1999), and Saez (2002, 2004).

result would seem to have a similar effect on the present analysis as on that which assumes an optimal income tax, it does not appear useful to pursue such matters further here.

References

Atkinson, A.B., Stiglitz, J.E., 1976. The design of tax structure: Direct versus indirect taxation. Journal of Public Economics 6, 55-75.

Atkinson, A.B., Stiglitz, J.E., 1980. Lectures on Public Economics. McGraw-Hill, New York.

Christiansen, V., 1984. Which commodity taxes should supplement the income tax? Journal of Public Economics 24, 195-220.

Cremer, H., Pestieau, P., Rochet, J.-C., 2001. Direct versus indirect taxation: The design of the tax structure revisited. International Economic Review 42, 781-799.

Deaton, A., 1979. Optimally uniform commodity taxes. Economics Letters 2, 357-361.

Hylland, A., Zeckhauser, R., 1979. Distributional objectives should affect taxes but not program choice or design. Scandinavian Journal of Economics 81, 264-284.

Kaplow, L., 1996. The optimal supply of public goods and the distortionary cost of taxation. National Tax Journal 49, 513-533.

Kaplow, L. 2004. On the (ir)relevance of distribution and labor supply distortion to government policy. Journal of Economic Perspectives (forthcoming).

Konishi, H., 1995. A Pareto-improving commodity tax reform under a smooth nonlinear income tax. Journal of Public Economics 56, 413-446.

Marchand, M., Pestieau, P., Racionero, M., 2003. Optimal redistribution when different workers are indistinguishable. Canadian Journal of Economics 36, 911-922.

Mirrlees, J.A. 1976. Optimal tax theory: A synthesis. Journal of Public Economics 6, 327-358.

Naito, H., 1999. Re-examination of uniform commodity taxes under a non-linear income tax system and its implication for production efficiency. Journal of Public Economics 71, 165-188.

Ramsey, F.P., 1927. A contribution to the theory of taxation. Economic Journal 37, 47-61.

Saez, E., 2002. The desirability of commodity taxation under non-linear income taxation and heterogeneous tastes. Journal of Public Economics 83, 217-230.

Saez, E. 2004. Direct or indirect tax instruments for redistribution: short-run versus long-run. Journal of Public Economics 88, 503-518.

Stiglitz, J.E., 1987. Pareto efficient and optimal taxation and the new new welfare economics.
In Auerbach, A.J., Feldstein, M. (Eds.), Handbook of Public Economics, vol. 2. North-Holland, Amsterdam, pp. 991-1042.

Yang, C.C., Haller, H., 1993. On directions of commodity tax reform in the presence of a given non-linear income tax schedule. Canadian Journal of Economics 26, 469-480.