



ELSEVIER

Journal of Public Economics 90 (2006) 1235–1250

JOURNAL OF
PUBLIC
ECONOMICS

www.elsevier.com/locate/econbase

On the undesirability of commodity taxation even when income taxation is not optimal

Louis Kaplow *

*Harvard University, Hauser Hall 322, Cambridge, MA 02138, USA
National Bureau of Economic Research, Cambridge, MA, USA*

Received 12 April 2004; received in revised form 1 May 2005; accepted 11 July 2005
Available online 16 September 2005

Abstract

An important result due to Atkinson and Stiglitz (1976) [Atkinson, A.B., Stiglitz, J.E., 1976. The design of tax structure: Direct versus indirect taxation. *Journal of Public Economics* 6, 55–75.] 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 this 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 neither eliminate nor proportionally reduce differential taxation, a simple efficiency condition is offered for determining whether a Pareto improvement is possible.

© 2005 Elsevier B.V. All rights reserved.

JEL classification: H21; H24

Keywords: Atkinson–Stiglitz; Commodity taxation; Income taxation; Indirect taxation; Optimal taxation; Ramsey taxation

* Tel.: +1 617 495 4101; fax: +1 617 496 4880.

E-mail address: moverholt@law.harvard.edu.

1. Introduction

Atkinson and Stiglitz (1976) analyze 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 separable between labor and commodities (taken together).

The Atkinson–Stiglitz result is acknowledged to be of great potential significance. A principal reason is that, if the Atkinson–Stiglitz theorem applies, widely-taught Ramsey (1927) tax principles are displaced by qualitatively different guidelines. Furthermore, Ramsey-based models underlie a variety of analyses and conclusions in public economics regarding issues ranging from capital taxation to environmental regulation to public sector pricing.²

It is therefore important to determine whether the result of Atkinson and Stiglitz extends to the case in which the income tax is not optimal. Standard methods of proof, including those in Atkinson and Stiglitz (1976) and related papers, rely on properties (first-order conditions) that do not hold when not at an optimum. In addition, 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. Boadway and Pestieau (2003, p. 400) conclude their retrospective on Atkinson and Stiglitz (1976) by emphasizing the conventional wisdom that “without optimal income taxation there is no A–S theorem.” Also, independent of the optimality of the income tax, it is useful to identify welfare-improving commodity tax reforms even when they involve nonmarginal changes or when commodity taxes cannot be set at ideal levels for various political or practical reasons.

This article demonstrates that the Atkinson–Stiglitz result indeed is true whatever the income tax is and, moreover, that the result can be extended to characterize a diverse array of commodity tax reforms. 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 then employed first 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. For commodity tax reforms that neither proportionally reduce nor fully eliminate differential taxation, it is demonstrated that a simple efficiency condition—one that depends only on efficiency in consumption, without regard to labor

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.

² This point is developed with regard to capital taxation in Atkinson and Stiglitz (1976, 1980) and Stiglitz (1987). See also Ordovery and Phelps (1979), who present results in an overlapping-generations model.

supply—determines whether a Pareto improvement is possible (again assuming weak separability). Concluding remarks are offered in Section 5.

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 a variety of 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, [Atkinson and Stiglitz \(1976\)](#) and [Deaton \(1979\)](#) characterize the restrictions on utility functions necessary for the Atkinson–Stiglitz result to hold when an optimal linear income tax is employed.)³ The few who have considered commodity taxation with an arbitrary initial income tax do not offer the analysis or obtain the results presented here.⁴

2. Intuitive explanation of the results

It is useful to begin with the result of [Atkinson and Stiglitz \(1976\)](#) 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? There are two complementary ways to state the motivation for doing so. One is that some differential commodity taxation would be optimal if it relaxed the incentive-compatibility constraints associated with the optimal income tax problem. Another is that the reason redistributive labor income taxation is not first best is on account of the labor–leisure distortion, so commodity taxation would be useful to the extent it can offset that distortion.⁵ Given weak separability, differential commodity taxation cannot relax the incentive-compatibility constraints and thereby mitigate the labor–leisure distortion. Differential commodity taxation does, however, introduce distortions in commodity choices; hence, it is not optimal.

³ 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.

⁴ Subsequent to this article's appearance as an NBER working paper and submission to this Journal, [Laroque \(2005\)](#) derived a result similar to that in the first of the three propositions below. [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. [Konishi \(1995\)](#) differs in that his method applies only to local reforms, whereas the present analysis and results apply to discrete changes — and in particular to discrete reforms of direct policy interest. Additionally, he uses more restrictive assumptions on utility and on the preexisting income tax (including that it be smooth, unlike actual income taxes) and, due to the technical nature of his approach, does not provide an intuitively accessible understanding of the problem.

⁵ The relationship between these two statements is best seen in models of optimal nonlinear income taxation when there are only two types of individuals. The binding incentive-compatibility constraint concerns the high-ability type mimicking the low-ability type. To prevent mimicking while achieving redistribution involves a positive marginal tax rate on the low-ability type, distorting that type's labor effort.

Upon examination, this result 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 and 4 by combining a reform of 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 the level of individuals' labor supply. 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.⁶ 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 noted by Atkinson and Stiglitz). It follows that the only net effect of this reform is on revenue.

It remains to be shown that combining a 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 constructed to be 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 revenue 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 set optimally, for such optimality only ensures that society is making the appropriate tradeoff of redistribution and labor supply distortion. If one could hold the extent of redistribution and individuals' labor supply fixed, one would expect that reducing distortions in commodity taxation would raise social welfare. And if

⁶ 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 analysis, applications, and discussion, see Kaplow (1996, 2004).

one indeed holds constant the overall distribution of well-being, then the gain in total welfare is being allocated in part to each individual, which makes everyone better off. The analysis to follow formalizes this basic idea.

3. Elimination of differential commodity taxation

Individuals choose levels of consumption of n commodities, x_1, \dots, x_n , and of labor effort l to maximize the utility function $u(v(x_1, \dots, x_n), l)$, where v is a subutility function. The utility function is 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 goods 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

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

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 .

In the analysis that follows, use will be made of the indirect subutility function $V(\tau, T, wl)$, which is defined to be the value of $v(x_1, \dots, x_n)$, maximized over the x_i 's, where the commodity tax vector τ , the income tax schedule T , and before-tax income wl are taken as given. That is, V is the maximized value of v subject to the budget constraint (1). Observe that since v depends only on the x_i 's, and since the constraint (1) depends only on the x_i 's, τ , T , and wl —and not on w or l independently—the indirect subutility function V is the same for all individuals, regardless of their type w . It will also be useful to use this indirect subutility function to define $U(\tau, T, w, l) = u(V(\tau, T, wl), l)$.

The government's budget constraint is

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

where R is a given revenue requirement.

The approach is to begin with a differentiated tax system and then 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 $\{\tau_1, \dots, \tau_n\}$, $T(wl)$ 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 must 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 $\{\tau_1^*, \dots, \tau_n^*\}$, $T^*(wl)$ 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 .

Observe that 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 $\{\tau_1, \dots, \tau_n\}$, $T(wl)$. 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’ utilities because they no longer pay commodity taxes (or receive subsidies) and because, with a new relative price vector, they will change their consumption vectors. Whatever is the net effect on utility for each type w and given labor effort for that type, 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.⁷ 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 as under the initial tax system, then their utility will be unchanged.⁸ (Whether individuals will choose the same labor effort under this intermediate regime is the subject of Lemma 1, below.)⁹

To be more precise, define $T^\circ(wl)$ such that $V(\tau, T, wl) = V(\tau^*, T^\circ, wl)$ for all wl . As just suggested, the reform from τ to τ^* will, for a given wl , change the value of subutility V . For each wl , one can set the tax schedule $T^\circ(wl)$ —directly changing after-

⁷ 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.

⁸ 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.

⁹ In the optimal tax literature, it is more common to focus on the related question whether incentive-compatibility constraints are satisfied, so as to avoid mimicking. As the text suggests and the statement and proof of Lemma 1 make explicit, the present approach is formulated more directly, in terms of individuals’ choices of labor effort. Specifically, it is assumed that individuals of each type choose l to maximize utility under any regime. In demonstrating that this choice does not change for each type, it follows that no type of individual chooses a different level of l —which, in turn, would yield a different level of income, mimicking some other type who earns that other income.

tax income for the stipulated level of before-tax income—at the level that restores the original level of subutility. This tax adjustment, $T^\circ(wl) - T(wl)$, is simply the utility-compensating change in disposable income. The ability to construct such an income tax schedule at each income level is due to the continuity of utility in income.¹⁰ Given how the intermediate income tax schedule is constructed, it is possible to establish the following result:

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

Proof. It is straightforward to establish that $U(\tau, T, w, l) = u(V(\tau, T, wl), l) = u(V(\tau^*, T^\circ, wl), l) = U(\tau^*, T^\circ, w, l)$, for all w, l . The first equality follows by the definition of U . The second equality follows because $T^\circ(wl)$ is constructed such that $V(\tau, T, wl) = V(\tau^*, T^\circ, wl)$ for all wl . And the third equality also follows from the definition of U . Therefore, $U(\tau, T, w, l) = U(\tau^*, T^\circ, w, l)$, for all w, l . Because, for any type w , this equality holds for all l , the level of utility an individual achieves for each possible choice of l is the same in each of the two regimes. Therefore, for each type w , whatever l maximizes U in the initial regime (τ, T) must be the l that maximizes U in the intermediate regime (τ^*, T°) . \square

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

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

Proof. The strategy will be to show that no individuals under the intermediate regime $\{\tau_1^*, \dots, \tau_n^*\}$, $T^\circ(wl)$ can still afford the consumption vector purchased under the initial regime $\{\tau_1, \dots, \tau_n\}$, $T(wl)$, 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 would choose a different one under the intermediate regime because of the change in relative prices. This follows

¹⁰ It is straightforward that one could set the income tax sufficiently high that an individual's subutility 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 the income tax sufficiently low to guarantee that the individual's subutility 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 subutility 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 subutility will be the same as it is in the initial regime. And this is possible for all income levels, wl .

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, it must be that 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 $T^\circ(wl)$ 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

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

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

$$wl - T(wl) - \sum \tau_i x_i(wl) > wl - T^\circ(wl), \text{ or} \\ T^\circ(wl) > T(wl) + \sum \tau_i x_i(wl), \text{ for all } wl. \quad (4)$$

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 under that regime. \square

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

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

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 gradually 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 (optimally) 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 aggregate behavior is continuous, some net reduction in everyone's income tax payments will be possible.)¹¹

¹¹ 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 income tax schedule, some individuals might discontinuously reduce their labor supply). Nevertheless, a minimal assumption 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.

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

4. Other reforms

Most analyses of commodity taxation with an income tax 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 employed to consider partial reforms.

4.1. Proportional reductions of differential commodity taxation

Consider a partial tax reform 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 $T^\circ(w)$ that holds utility constant if labor effort is constant. Lemma 1 still holds because it only depends on the properties of $T^\circ(w)$ (and weak separability). 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.

Accordingly, let us reconsider Lemma 2, which states that more revenue is raised in the intermediate regime with $T^\circ(w)$ than under the initial regime. In moving from expression (3) to (4) and in interpreting the latter, the proof used the fact 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 following claim:

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

The proof appears in the Appendix. Given this result, we can state:

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

On reflection, it should not be surprising that Lemma 2' can be established and, accordingly, the result in Proposition 1 can be extended to partial proportional reform of commodity taxation. When a commodity tax reform involves a uniform proportional move

of commodity taxes and subsidies 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 proof of Lemma 2' establishes, however, that the intuitively expected tendency is indeed correct in the aggregate, which is all that is necessary for the more general proposition.

4.2. Nonuniform 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 were 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 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—would decline. 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. To be more precise, consider the following sort of reform:

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

$$\sum p_i \int x_i^\circ(wl)f(w)dw < \sum p_i \int x_i(wl)f(w)dw. \quad (6)$$

Expression (6) 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 due to income taxation are ignored.

Just as in the prior subsection, all that is required to demonstrate our result is to prove the analogue to Lemma 2 because the rest of the analysis applies, *mutatis mutandis*.

Lemma 2''. *A commodity tax reform $\{\tau_1^*, \dots, \tau_n^*\}$, $T^\circ(wl)$ raises more revenue than does regime $\{\tau_1, \dots, \tau_n\}$, $T(wl)$ if and only if it is an efficiency-increasing commodity tax reform (i.e., with taxes τ_i^* such that (6) holds).*

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

$$\begin{aligned} & \left[\int T^\circ(wl)f(w)dw + \sum \tau_i^* \int x_i^\circ(wl)f(w)dw \right] \\ & - \left[\int T(wl)f(w)dw + \sum \tau_i \int x_i(wl)f(w)dw \right] \\ & = \sum p_i \int x_i(wl)f(w)dw - \sum p_i \int x_i^\circ(wl)f(w)dw. \end{aligned} \quad (7)$$

The first bracketed term on the left side of expression (7) is total revenue under the intermediate regime and the second bracketed term is total revenue under the initial regime. The right side is positive if and only if expression (6), the definition of an efficiency-increasing commodity tax reform, holds. \square

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

Proposition 3. *Beginning with any tax system $\{\tau_1, \dots, \tau_n\}$, $T(wl)$, 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. 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 (6) holds for the elimination or proportional 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 (7) as a sort of accounting identity when a regime is constructed with $T^\circ(wl)$, as in the proof of all three propositions.)

5. 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 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 neither proportionally reduce nor 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. The present approach makes the intuition transparent by using an income tax adjustment that holds the extent of redistribution and the level of individuals' labor supply constant, so in essence the only remaining effects are the narrow efficiency consequences of commodity tax reform.

Furthermore, the fact that 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 makes it a practically useful adjustment to consider.¹³ Political forces that oppose an optimally redistributive income tax according to some social welfare function need not prevent (and might, on average, tend to favor) reforms that hold distribution constant, offering gains to everyone. And, regardless of political feasibility, it is useful to distinguish the effects of a distribution-neutral commodity tax reform from purely redistributive effects. The latter may accompany any sort of policy change and can be analyzed independently of any particular commodity tax reform.

As suggested in the Introduction, the present results are especially significant in light of their large qualitative departure from those of Ramsey models that underlie much public economics teaching and analysis of taxation and other government policies. Accordingly, a substantial reorientation of thinking, which seems to have been held in abeyance due to skepticism about the applicability of the Atkinson–Stiglitz theorem, seems to be in order.

This viewpoint is not greatly affected by various other qualifications to [Atkinson and Stiglitz's \(1976\)](#) result that are familiar from their article and from other literature on

¹³ To verify the distributional characteristics of the reform as a whole, observe first that $T^\circ(wl)$ 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 $T^\circ(wl)$ by subtracting a constant, as indicated by expression (5).

optimal commodity taxation in the presence of an optimally determined nonlinear income tax.¹⁴ Consider, for example, the assumption of weak labor separability, which rules out the ability of differential commodity taxation to help 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 uninfluenced 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.) This caveat and most others in the pertinent literatures would seem to have a similar effect on the present analysis as on that which assumes an optimal income tax, so it does not appear useful to pursue such matters further here. However, it should be emphasized that these sorts of qualifications are largely orthogonal to and thus do not restore basic principles of Ramsey taxation, such as the inverse-elasticity rule or the desirability of taxing luxuries when distributive concerns are introduced into the basic model. Instead, when there is an income tax, the relevant benchmark for analysis—the focal point from which adjustments should be considered—is Atkinson and Stiglitz’s uniform commodity taxation result.¹⁵

Acknowledgments

I am grateful to James Hines, Steven Shavell, participants at the NBER public economics program meeting, and the referees for comments and to the John M. Olin Center for Law, Economics, and Business at Harvard University for financial support. This article relates to an in-progress book, tentatively entitled *Taxation and Redistribution*, which develops many applications of the results herein.

Appendix

Proof of Lemma 2'

As noted in the text, 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

$$T^o(wl) + \sum \tau_i^* x_i(wl) > T(wl) + \sum \tau_i x_i(wl), \text{ for all } wl. \quad (A1)$$

Because the second term on the left side of (A1) has x_i instead of x_i^o , the left side is not total revenue under the intermediate regime. However, this expression can still prove

¹⁴ See, for example, Boadway and Pestieau (2003), Cremer et al. (2001), Marchand et al. (2003), Mirrlees (1976), Naito (1999), and Saez (2002, 2004).

¹⁵ For further elaboration, in the context of applying these ideas to the analysis of public goods and regulation, see Kaplow (1996, 2004), and for an extension of the present model to the case of externalities, see Kaplow (2005).

helpful, specifically, if we substitute $\alpha\tau_i$ for τ_i^* and integrate each side over the population of types to yield

$$\begin{aligned} & \int T^\circ(wl)f(w)dw + \alpha \sum \tau_i \int x_i(wl)f(w)dw > \\ & \int T(wl)f(w)dw + \sum \tau_i \int x_i(wl)f(w)dw, \text{ or} \\ & \frac{1}{1-\alpha} \left[\int T^\circ(wl)f(w)dw - \int T(wl)f(w)dw \right] > \sum \tau_i \int x_i(wl)f(w)dw. \end{aligned} \quad (\text{A2})$$

Expression (A2) provides information about the relationship between commodity tax revenue in the initial regime and income tax revenue in the two regimes.

Next, we can obtain information about the relationship between commodity tax revenue in the intermediate regime and income tax revenue in the two regimes, using an approach similar to that used to derive (A2). First, observe that, for any individual 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

$$\sum (p_i + \tau_i)x_i^\circ(wl) > \sum (p_i + \tau_i)x_i(wl), \text{ for all } wl. \quad (\text{A3})$$

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

$$T^\circ(wl) + \sum (p_i + \tau_i^*)x_i^\circ(wl) = T(wl) + \sum (p_i + \tau_i)x_i(wl), \text{ for all } wl. \quad (\text{A4})$$

That is, for a given level of income, all of it must be spent on income tax payments and on commodities (producer prices and commodity taxes) in each regime. Solving (A4) for $\sum p_i x_i^\circ$ and substituting this into the left side of (A3) gives us

$$\begin{aligned} & \sum (p_i + \tau_i)x_i(wl) + T(wl) - T^\circ(wl) + \sum \tau_i x_i^\circ(wl) - \sum \tau_i^* x_i^\circ(wl) > \\ & \sum (p_i + \tau_i)x_i(wl), \text{ for all } wl. \end{aligned} \quad (\text{A5})$$

Simplifying and rearranging terms, we have

$$\sum (\tau_i - \tau_i^*)x_i^\circ(wl) > T^\circ(wl) - T(wl), \text{ for all } wl. \quad (\text{A6})$$

The intuition behind expression (A6) 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 this part of the argument, on the left side of (A6) we can substitute for τ_i using the fact that $\tau_i^* = \alpha\tau_i$, integrate each side over the population of types, and multiply both sides by $\alpha/(1-\alpha)$ to yield

$$\sum \tau_i^* \int x_i^\circ(w) f(w) dw > \frac{\alpha}{1-\alpha} \left[\int T^\circ(w) f(w) dw - \int T(w) f(w) dw \right]. \quad (A7)$$

Finally, we can combine inequalities (A2) and (A7), specifically by adding the left sides and adding the right sides, and then rearrange terms to produce the result that

$$\begin{aligned} \int T^\circ(w) f(w) dw + \sum \tau_i^* \int x_i^\circ(w) f(w) dw > \\ \int T(w) f(w) dw + \sum \tau_i \int x_i(w) f(w) dw. \end{aligned} \quad (A8)$$

The left side is total tax revenue in the intermediate regime and the right side is total tax revenue in the original regime, which proves the result. \square

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.
- Boadway, R., Pestieau, P., 2003. Indirect taxation and redistribution: The scope of the Atkinson–Stiglitz theorem. In: Arnott, R., Greenwald, B., Kanbur, R., Nalebuff, B. (Eds.), *Economics for an Imperfect World: Essays in Honor of Joseph E. Stiglitz*. MIT Press, Cambridge, MA, pp. 387–403.
- 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* 18:4, 159–175.
- Kaplow, L., 2005. Optimal control of externalities in the presence of income taxation. Working paper.
- Konishi, H., 1995. A Pareto-improving commodity tax reform under a smooth nonlinear income tax. *Journal of Public Economics* 56, 413–446.
- Laroque, G., 2005. Indirect taxation is harmful under separability and taste homogeneity: A simple proof. *Economics Letters* 87, 141–144.
- 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.
- Ordovery, J.A., Phelps, E.S., 1979. The concept of optimal taxation in the overlapping-generations model of capital and wealth. *Journal of Public Economics* 12, 1–26.

- 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.